IRMA Lectures in Mathematics and Theoretical Physics 10

Edited by Vladimir G. Turaev

Institut de Recherche Mathématique Avancée
Université Louis Pasteur et CNRS
7 rue René Descartes
67084 Strasbourg Cedex
France

IRMA Lectures in Mathematics and Theoretical Physics
Edited by Vladimir G. Turaev

This series is devoted to the publication of research monographs, lecture notes, and other materials arising from programs of the Institut de Recherche Mathématique Avancée (Strasbourg, France). The goal is to promote recent advances in mathematics and theoretical physics and to make them accessible to wide circles of mathematicians, physicists, and students of these disciplines.

Previously published in this series:

1 Deformation Quantization, *Gilles Halbout* (Ed.)
2 Locally Compact Quantum Groups and Groupoids, *Leonid Vainerman* (Ed.)
3 From Combinatorics to Dynamical Systems, *Frédéric Fauvet and Claude Mitschi* (Eds.)
4 Three courses on Partial Differential Equations, *Eric Sonnendrücker* (Ed.)
5 Infinite Dimensional Groups and Manifolds, *Tilman Wurzbacher* (Ed.)
6 *Athanase Papadopoulos*, Metric Spaces, Convexity and Nonpositive Curvature
7 Numerical Methods for Hyperbolic and Kinetic Problems, *Stéphane Cordier, Thierry Goudon, Michaël Gutnic and Eric Sonnendrücker (Eds.)*
8 AdS-CFT Correspondence: Einstein Metrics and Their Conformal Boundaries, *Oliver Biquard* (Ed.)
9 Differential Equations and Quantum Groups, *D. Bertrand, B. Enriquez, C. Mitschi, C. Sabbah and R. Schaefke* (Eds.)
10 Physics and Number Theory, *Louise Nyssen* (Ed.)

Volumes 1–5 are available from Walter de Gruyter (www.degruyter.de)

Physics and Number Theory

Louise Nyssen
Editor

European Mathematical Society

Editor:

Louise Nyssen
Institut de Recherche Mathématique Avancée
Université Louis Pasteur et CNRS
7 Rue René Descartes
67084 Strasbourg Cedex
France

2000 Mathematics Subject Classification : 11F; 11L, 11M, 81T, 52C, 68R15

ISBN 978-3-03719-028-9

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie;
detailed bibliographic data are available in the Internet at http://dnb.ddb.de.

Contact address:

European Mathematical Society Publishing House
Seminar for Applied Mathematics
ETH-Zentrum FLI C4
CH-8092 Zürich
Switzerland

Phone: +41 (0)44 632 34 36
Email: info@ems-ph.org
Homepage: www.ems-ph.org

Typeset using the author's TEX files: I. Zimmermann, Freiburg
Printed in Germany

9 8 7 6 5 4 3 2 1

Preface

A single book is certainly not enough to describe the rich and historical relationships between physics and number theory. This volume presents a selection of problems which are currently in full development and inspire the research of many people. All the papers begin with a survey which will make it possible even for non-specialists to understand them and will give an idea of the great variety of subjects and techniques in this frontier area.

The first paper, "The phase of oscillations and prime numbers: classical and quantum", by Michel Planat, is an example of the strong connection between physics and mathematics. It starts from a concrete problem and brings into play an impressive variety of mathematical techniques, especially in number theory. The paper provides an accessible introduction to the problem of phase-locking in oscillating systems, both at a classical level and at a quantum level. The mathematical formulation of the different aspects of this problem requires numerous tools: first, you see how prime numbers appear, together with continuous fractions and the Mangoldt function. Then, come some hyperbolic geometry and the Riemann ζ-function. On the quantum side, roots of unity and Ramanujan sums are related to noise in oscillations, and when discussing phase in quantum information, the author uses Bost and Connes KMS states, Galois rings and fields along with some finite projective geometry.

Next there are two papers about crystallography. From a physical point of view, a crystal is a solid having an essentially discrete diffraction diagram. It can be periodic or not. From a mathematical point of view, lattices in $\mathbb{R}^n$ are good tools to describe periodic crystals, but not aperiodic ones. Very little is known about aperiodic crystals, apart from the so called *quasicrystals*, whose diffraction diagrams present some regularity: they are invariant under dilatation by a factor that may be irrational. They are well described by some discrete sets called *cut-and-project* sets, which are a generalisation of lattices. In his paper "On Self-Similar Finitely Generated Uniformly Discrete (SFU-)Sets and Sphere Packings", Jean-Louis Verger-Gaugry is interested in cut-and-project sets in $\mathbb{R}^n$. The first part of the paper is a survey of the link between the geometry of numbers and aperiodic crystals in physics, from the mathematical point of view. In the second part, the author proves some new results about the distances between the points of cut-and-project sets. By considering each point as the centre of a sphere, one gets a sphere packing problem which is hopefully a good model for atom packing. In "Nested quasicrystalline discretisation of the line", Jean-Pierre Gazeau, Zuzana Masáková, and Edita Pelantová focus on cut-and-project sets obtained from a square lattice in $\mathbb{R}^2$, with the idea of constructing aperiodic wavelets. They review the geometrical properties of such sets, their combinatorial properties from the point of view of language theory and their relation to nonstandard numeration systems based

on θ when the cut-and-project set is self-similar for an irrational scaling factor θ. Finally, they provide an algorithm which generates cut-and-project sets.

In "Hopf algebras in renormalization theory: locality and Dyson–Schwinger equations from Hochschild cohomology", Christoph Bergbauer and Dirk Kreimer reformulate a problem from quantum field theory in algebraic terms. In the first part, the authors explain how, by constructing an appropriate Hopf algebra structure on rooted trees, one gets an algebraic formulation of the renormalization process for Feynman graphs. After an introductory overview of those Hopf algebras, they give numerous examples of their generalizations. In the second part, they recall the definition of the Hochschild cohomology associated to such Hopf algebras. A physical interpretation of the cocycles leads to the Dyson–Schwinger equations, which describe the loop expansion of Green function in a recursive way. This method is a straightforward alternative to the usual one. The paper ends up with a description of the Dyson–Schwinger equation from different points of view: it provides transcendental numbers, it is a tool to define a generating function for the polylogarithm and it is the equation of motion for a renormalizable quantum field theory.

"Fonction ζ et matrices aléatoires" by Emmanuel Royer, is an extensive survey on the zeroes of L-functions. The analytic properties of the Riemann ζ-function describe the behaviour of prime numbers. More generally, one can attach an L-function to certain arithmetic or geometric objects, and thus obtain some information on them by analytic methods. The functions described in this paper are the Riemann ζ-function and L-functions attached to Dirichlet characters, to automorphic representations, or to modular forms. The Generalised Riemann Hypothesis predicts that all the zeroes of those L-functions lie on the line $\mathrm{Re}(z) = \frac{1}{2}$, and other important conjectures relate their statistical behaviour along this line to the statistical behaviour of the spectrum of large unitary matrices.

The paper by Philippe Michel, "Some recent applications of Kloostermania", is devoted to Kloosterman sums which are a special kind of algebraic exponential sums. They where first developed to bound the number of representations of a large integer n by a diagonal quaternary definite quadratic form

$$n = ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2.$$

They eventually turned out to be one of those fascinating objects with two faces: the Petersson–Kuznetsov trace formula relates sums of Kloostermans sums, data of an arithmetico-geometric nature, to the Fourier coefficients of modular forms, data of a spectral nature. First written by Kuznetsov for $\mathrm{SL}_2(\mathbb{Z})$ the formula was extended by Deshouillers and Iwaniec to arbitrary congruence subgroups. The use of this connection in both directions is a powerful tool in analytic number theory. There are many applications. For example, one can get good estimates of linear combinations of Fourier coefficents, or get an upper bound for the dimension of the space of weight one modular forms, or solve many instances of the subconvexity problem. The last application presented in the paper is that, using the Petersson–Kuznetzov trace formula,

one can refine the error term in Weyl's law and that these refinements can be interpreted in terms of Quantum Chaos.

Finally, Ariane Mézard's "Introduction à la correspondance de Langlands locale", is a survey of the local Langlands correspondence. When K is a non archimedian local field, and W_K its Weil group, the local class field isomorphism

$$W_K^{ab} \simeq K^*$$

can be interpreted as a bijection between continuous irreducible representations of $K^* = \mathrm{GL}_1(K)$ and one dimensional continuous representation of W_K. The Langlands conjecture is a generalisation of this isomorphism in higher dimension: for any $n \geq 1$ there should be an isomorphism between isomorphism classes of continuous irreducible admissible representations of $\mathrm{GL}_n(K)$ and isomorphism classes of n-dimensional Φ-semi-simple continuous representations of the Weil–Deligne group W'_K. In her paper, Ariane Mézard studies the bijection between isomorphism classes of irreducible admissible representations of $\mathrm{GL}_2(\mathbb{Q}_p)$ over $\overline{\mathbb{Q}}_\ell$, and isomorphism classes of some representations of $W_{\mathbb{Q}_p}$ in $\mathrm{GL}_2(\overline{\mathbb{Q}}_\ell)$. She explains the main steps of the proof in the ℓ-adic case (this means $\ell \neq p$, in which case the conjecture is proved for any n). She also points out where and why problems arise in the p-adic case ($\ell = p$). She describes the new objects and explains the strategy for $n = 2$. The p-adic conjecture is still open.

I would like to take the opportunity to thank the referees for their selfless work and numerous constructive remarks, Marcus Slupinski for help with English, and to express special gratitude to Vladimir Turaev whose constant help during the preparation of the volume was a great support.

Strasbourg, August 2006 *Louise Nyssen*

Table of Contents

The phase of oscillations and prime numbers: classical and quantum

*Michel Planat**

Institut FEMTO-ST, Departement LPMO,
32 Avenue de l'Observatoire, 25044 Besançon Cedex, France
e-mail: planat@lpmo.edu

Abstract. An overview of the paradigm of phase-locking at the nonlinear, geometric, quantum and finite field level is attempted. It is applied to finite resolution measurements, the $1/f$ noise in a communication receiver, and to quantum information concepts. The Mangoldt function of prime number theory is found to be the generator of low frequency noise in the coupling coefficient between oscillators, in the scattering coefficient, and in quantum statistical states. On the other hand, phase operators defined on finite fields (and rings) control the quantum information primitives. Huyghens coupled pendulums, the Adler equation, the Arnold map, continued fraction expansions, discrete Möbius transformations, Ford circles, coherent and squeezed states, Ramanujan and Weil sums, the Riemann zeta function, Bost and Connes KMS states, Galois fields and rings and their additive characters are used successfully to unveil the phase dynamics.

1 Introduction

The interleaving of frequencies and phases of electronic oscillators interacting in nonlinear circuits follows arithmetical rules. Continued fraction expansions, prime number decompositions and related number theoretical concepts were successfully used to account for the experimental effects in mixers and phase-locked loops [43], [48]. We also made use of these tools within the field of quantum optics emphasizing the hidden connection between phase-locking and cyclotomy [49]. As a matter of fact the understanding of phase effects in devices is so much entangled that it will reveal useful to check several clues relying on differential and discrete nonlinear equations, Fourier analysis, hyperbolic geometry, Galois fields and indeed quantum mechanics. But over all the fundamental issue is the modelling of finite resolution measurements and how it puts a constraint on the performance, how it generates the fluctuations in oscillating circuits, how optimal classical and quantum measurements can be realized.

In Section 2.1 we report on the early history of phase-locking about classical observations of coupled mechanical pendulums and electronic oscillators. Nonlinear

*Thanks to my collaborators Serge Perrine, Metod Saniga and Haret Rosu.

continuous and discrete generic models are introduced emphasizing the case of homodyne detection in a phase sensitive communication receiver.[1] It is shown that low frequency noise happens due to low pass filtering and the finite resolution of frequency counts. A phenomenological model relating the coupling coefficient to prime numbers is developed.

In Section 2.2 the justification of the model is given by using the hyperbolic geometry of the half-plane applied to the low pass filtering. In Section 3 the phase itself is taken to be discrete. Discrete quantum optics is related to quantum phase-locking, the arithmetic of $1/f$ noise and the famous Bost and Connes quantum statistical approach to the Riemann hypothesis. Finally in Section 4 still another connection is found between quantum phase operators and prime numbers within the context of mutually unbiased bases, which are used for optimal quantum information protocols such as quantum cryptography and quantum state tomography. Their Galois fields and rings and finite projective geometry play a prominent role in the Fourier expansion.

2 Phase-locking effects in classical oscillators

2.1 Classical phase-locking: from Huyghens to the prime numbers

> Being obliged to stay in my room for several days and also occupied in making observations on my two newly made clocks, I have noticed a remarkable effect which no one could have ever thought of. It is that these two clocks hanging next to one another separated by one or two feet keep an agreement so exact that the pendulums invariably oscillate together without variation. After admiring this for a while, I finally figured out that it occurs through a kind of sympathy: mixing up the swings of the pendulums, I have found that within a half hour always return to consonance and remain so constantly afterwards as long as I let them go. I then separated them, hanging one at the end of the room and the other fifteen feet away, and noticed that in a day there was five seconds difference between them. Consequently, their earlier agreement must in my opinion have been caused by an imperceptible agitation of the air produced by the motion of the pendulums.

The citation is taken from [61]. The authors remind a later letter by Huyghens that the coupling mechanism was in fact a small vibration transmitted through the wall, and not movement of air:

> Lord Rayleigh (1907) made similar observations about two driven tuning forks coupled by vibrations transmitted through the table on which both forks sat … Locking in triode circuits was explained by Van der Pol (1927) who included in the equation for the triode oscillator an external electromotive force as given in

$$\frac{d^2v}{dt^2} - \frac{d}{dt}(gv - \beta' v^3) + \omega^2 v = \omega_0^2 V_0 \sin \omega_0 t, \tag{2.1}$$

[1] In the homodyne method, the reflected signal is mixed with the transmitted signal. This results in a signal from the receiver which is proportional to the phase difference between the two signals.

where g is the linear net gain (i.e. the gain in excess of losses), β' the saturation coefficient, and ω is the resonance frequency in the absence of dissipation or gain. He showed that when an external electromotive force is included, of frequency ω_0, and tuned close to the oscillator frequency ω, the oscillator suddenly jumped to the external frequency. It is important to note that the beat note between the two frequencies vanishes not because the two frequencies vanish, not because the triode stops oscillating, but because it oscillates at the external frequency.

We can show the locking effect by utilizing the slowly varying amplitude approach, including a slowly varying phase Φ and oscillation at the external frequency ω_0 and amplitude V

$$\frac{d\Phi}{dt} + K \sin \Phi = \omega - \omega_0 = \omega_{\mathrm{LF}}, \tag{2.2}$$

where we use ω_{LF} for the detuning term and $K = \omega_0 V_0 / V$ for the locking coefficient [61].

The regime just described is the so-called injection locking regime, also found in injection-locked lasers. Equation (2.2) is the so-called Adler equation of electronics [1].

One way to synthesize (2.2) is thanks to the phase-locked loop (or PLL)[2] of a communication receiver. The receiver is designed to compare the information carrying external oscillator (RF) to a local oscillator (LO) of about the same high frequency through a nonlinear mixing element. For narrow band demodulation one uses a discriminator of which the role is first to differentiate the signal, that is to convert frequency modulation (FM) to amplitude modulation (AM) and second to detect its low frequency envelope: this is called baseband filtering. For more general FM demodulation one uses a low pass filter instead of the discriminator to remove the high frequency signals generated after the mixer. In the closed loop operation a voltage controlled LO (or VCO) is used to track the frequency of the RF. Phase modulation is frequently used for digital signals because low bit error rates can be obtained despite poor signal to noise ratio in comparison to frequency modulation [31].

[2] A phase-locked loop (PLL) is an electronic circuit with a voltage- or current-driven oscillator that is constantly adjusted to match in phase (and thus lock on) the frequency of an input signal. In addition to stabilizing a particular communications channel (keeping it set to a particular frequency), a PLL can be used to generate a signal, to modulate or demodulate a signal, to reconstitute a signal with less noise, or to multiply or divide a frequency. PLLs are frequently used in wireless communication, particularly where signals are carried using frequency modulation (FM) or phase modulation (PM). PLLs can also be used in amplitude modulation (AM). PLLs are more commonly used for digital data transmission, but can also be designed for analog information. Phase-locked loop devices are more commonly manufactured as integrated circuits (ICs) although discrete circuits are used for microwave.

A PLL consists of a voltage-controlled oscillator (VCO) that is tuned using a special semiconductor diode called a varactor. The VCO is initially tuned to a frequency close to the desired receiving or transmitting frequency. A circuit called a phase comparator causes the VCO to seek and lock onto the desired frequency, based on the output of a crystal-controlled reference oscillator. This works by means of a feedback scheme. If the VCO frequency departs from the selected crystal reference frequency, the phase comparator produces an error voltage that is applied to the varactor, bringing the VCO back to the reference frequency. The PLL, VCO, reference oscillator, and phase comparator together comprise a frequency synthesizer. Wireless equipment that uses this type of frequency control is said to be frequency-synthesized.

Since a PLL requires a certain amount of time to lock on the frequency of an incoming signal, the intelligence on the signal (voice, video, or data) can be obtained directly from the waveform of the measured error voltage, which will reflect exactly the modulated information on the signal.

Let us consider a type of receiver which consists in a mixer, in the form of a balanced Schottky diode bridge[3] and a low pass filter. If f_0 and f are the frequencies of the RF and the LO, and $\theta(t)$ and $\psi(t)$ their respective phases, the set mixer and filter essentially behaves as a phase detector of sensitivity u_0 (in Volts/rad.), that is the instantaneous voltage at the output is the sinus of the phase difference at the inputs

$$u(t) = u_0 \sin(\theta(t) - \psi(t)). \tag{2.3}$$

The nonlinear dynamics of the set-up in the closed loop configuration is well described by introducing the phase difference $\Phi(t) = \theta(t) - \psi(t)$. Using $\dot{\theta} = \omega_0$ and $\dot{\psi}(t) = \omega + Au(t)$, with $\omega_0 = 2\pi f_0$, $\omega = 2\pi f$ and A (in rad. Hz/Volt) as the sensitivity of the VCO, one recovers Adler's equation (2.2) with the open loop gain $K = u_0 A$.

Equation (2.2) is integrable but its solution looks complex [18]. If the frequency shift ω_{LF} does not exceed the open loop gain K, the average frequency $\langle\dot{\Phi}\rangle$ vanishes after a finite time and reaches the stable steady state $\Phi(\infty) = 2l\pi + \sin^{-1}(\omega_{\mathrm{LF}}/K)$, l integer. In this phase-tracking range of width $2K$ the RF and the LO are also frequency-locked. Outside the mode-locking zone there is a sech (hyperbolic secant) shape beat signal of frequency

$$\tilde{\omega}_{\mathrm{LF}} = \langle\dot{\Phi}(t)\rangle = (\omega_{\mathrm{LF}}^2 - K^2)^{1/2}. \tag{2.4}$$

The sech shape signal and the nonlinear dependance on parameters ω_{LF} and K are actually found in experiments [18], [48]. In addition the frequency ω_{LF} is fluctuating (see Figure 1). It can be characterized by the Allan variance $\sigma^2(\tau)$ which is the mean squared value of the relative frequency deviation between adjacent samples in the time series, averaged over an integration time τ. Close to the phase-locked zone the Allan deviation (which the square root of Allan variance) is

$$\sigma(\tau) = \frac{\sigma_0 K}{\tilde{\omega}_{\mathrm{LF}}}, \tag{2.5}$$

where σ_0 is a residual frequency deviation depending of the quality of input oscillators and that of the phase detector. Allan deviation is found independent of τ which is a signature of a $1/f$ frequency noise of power spectral density $S(f) = \sigma/(2\ln 2 f)$. One way to predict the dependence (2.5) is to use differentiation of (2.4) with respect

[3]The Schottky diode (named after German physicist Walter H. Schottky) is a semiconductor diode with a low forward voltage drop and a very fast switching action. A typical application is discharge-protection for solar cells connected to lead-acid batteries. While standard silicon diodes have a forward voltage drop of about 0.6 volts, Schottky diodes have a drop of only about 0.3 volts. This is due to the higher current density in the Schottky diode.

A Schottky diode uses a metal-semiconductor junction as a Schottky barrier (instead of a semiconductor-semiconductor junction as in conventional diodes). This Schottky barrier results in both very fast switching times and low forward voltage drop.

It is often said that the Schottky diode is a "majority carrier" semiconductor device. This means that if the semiconductor body is doped N-type, only the N-type carriers (mobile electrons) play a significant role in normal operation of the device. No slow, random recombination of N- and P-type carriers is involved, so this diode can cease conduction faster than an ordinary PN rectifier diode. This property in turn allows a smaller device area, which also makes for a faster transition. Therefore broad-area Schottky diodes are useful in switch-mode power converters which operate at frequencies approaching 1 MHz. Small-area Schottky diodes are the heart of RF detectors and mixers, which often operate up to 5 GHz.

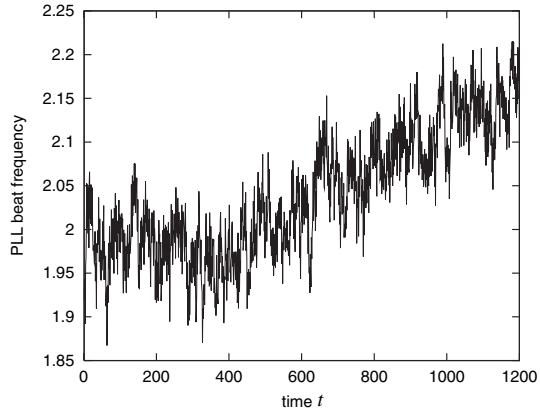


Figure 1. Fluctuating counts of the beat frequency (in Hz) close to the phase-locked zone. The inputs are quartz oscillators at 10 MHz. The power spectrum has a pure $1/f$ dependance.

to the frequency shift $\tilde{\omega}_{\rm LF}$ so that

$$\delta\tilde{\omega}_{\rm LF} = \delta\omega_{\rm LF}(1 + K^2/\tilde{\omega}_{\rm LF}^2)^{1/2}. \tag{2.6}$$

Relation (2.6) is defined outside the mode-locked zone $|\omega_{\rm LF}| > K$; close to it, if the effective beat note $\tilde{\omega}_{\rm LF} \leq K$, the square root term is about $K/\tilde{\omega}_{\rm LF}$. If one identifies $\delta\omega_{\rm LF}/\tilde{\omega}_{\rm LF}$ with a bare Allan deviation σ_0 and $\delta\tilde{\omega}_{\rm LF}/\tilde{\omega}_{\rm LF}$ with a magnified Allan deviation σ one explains the experimental result (2.5). One can conclude that either the phase-locked loop set-up behaves as a microscope of an underlying flicker floor σ_0, or the $1/f$ noise is some dynamical property of the PLL. In the past we looked at a possible low dimensional structure of the time series and found a stable embedding dimension lower or equal to 4 [47]. But at that time the dynamical model of $1/f$ noise still remained elusive.

Adler's model presupposes a fundamental interaction $\omega_{\rm LF} = |\omega_0 - \omega(t)|$ in the mixing of the two input oscillators. But the practical operation of the phase detector involves harmonic interactions of the form $\omega_{\rm LF} = |p\omega_0 - q\omega(t)| \leq \omega_c = 2\pi f_c$, where p and q are integers and f_c is the cut-off frequency of the low pass filter. This can be rewritten by introducing the frequency ratios $\nu = \frac{\omega(t)}{\omega_0}$ and $\mu = \frac{\omega_{\rm LF}}{\omega_0}$ as $\mu = q\left|\nu - \frac{p}{q}\right|$. This form suggests that the aim of the receiver is to select such couples (p, q) which realize a "good" approximation of the "real" number ν. There is a mathematical concept which precisely does that: the diophantine approximator. It selects such couples p_i and q_i, coprime to each other, i.e. with greatest common divisor $(p_i, q_i) = 1$, from the continued fraction expansion of ν truncated at the index i

$$\nu = \{a_0; a_1, a_2, \ldots, a_i\} = a_0 + 1/(a_1 + 1/(a_2 + 1/\cdots + 1/a_i)) = \frac{p_i}{q_i}. \tag{2.7}$$

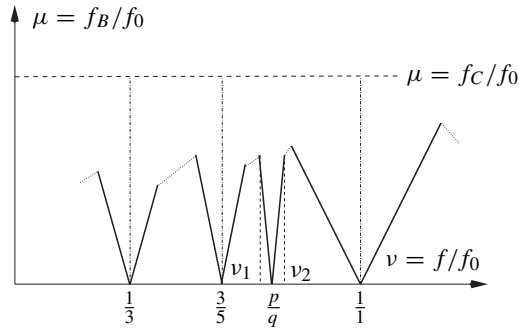


Figure 2. The intermodulation spectrum from homodyne detection.

The diophantine approximation satisfies

$$\left|\nu - \frac{p_i}{q_i}\right| \leq \frac{1}{a_{i+1}q_i^2}. \tag{2.8}$$

The fraction $\frac{p_i}{q_i}$ is a so-called convergent and the a_i''s are called partial quotients. The approximation is truncated at the index i just before the partial quotient a_{i+1}. It should be observed that diophantine approximations are different from decimal approximations $\frac{c_i}{d_i}$ for which one gets $\left|\nu - \frac{c_i}{d_i}\right| \leq \frac{1}{d_i}$. It was shown [43] using the filtering condition that a_{i+1} identifies with a very simple expression

$$a_{i+1} = \left[\frac{f_0}{f_c q_i}\right], \tag{2.9}$$

where [ ] denotes the integer part. For example if one chooses $f_0 = 10\,\text{MHz}$ and $f_c = 300\,\text{kHz}$, the fundamental basin $\frac{p_i}{q_i} = \frac{1}{1}$ will be truncated if $a_{i+1} \geq 33$ and the basin $\frac{p_i}{q_i} = \frac{3}{5}$ will be truncated if $a_{i+1} \geq 6$. The resulting full spectrum is a superposition of V-shape basins of which the edges are located at

$$\begin{aligned} \nu_1 &= \{a_0; , a_1, a_2, \ldots, a_i, a_{i+1}\}, \\ \nu_2 &= \{a_0; a_1, a_2, \ldots, a_{i-1}, 1, a_{i+1}\}, \end{aligned} \tag{2.10}$$

where the partial expansion before a_{i+1} corresponds to the two possible continued fractions of the rational number $\frac{p_i}{q_i}$ (see Figure 2). The basin of number $\nu = \frac{3}{5} = \{0; 1, 1, 2\}$ extends to $\nu_1 = \{0; 1, 1, 2, 33\} = \frac{19}{32} \simeq 0.594$ and $\nu_2 = \{0; 1, 1, 1, 1, 33\} = \frac{31}{34} \simeq 0.618$. For a reference oscillator with $f_0 = 10$ MHz this corresponds to a frequency bandwidth $(0.618 - 0.594).10^7$ MHZ $= 240$ kHz.

With these arithmetical rules in mind one can now tackle the difficult task to account for phase-locking of the whole set of harmonics. The differential equation for the phase shift $\dot{\Phi}(t; q_i, p_i)$ at the harmonic (p_i, q_i) corresponding to the beat frequency

$$\omega_{\text{LF}} = |p_i\omega_0 - q_i\omega(t)| \tag{2.11}$$

can be obtained as

$$\dot{\Phi}(t; q_i, p_i) + q_i H(P) \sum_{r_i, s_i} K(r_i, s_i) \times \sin\left(\frac{s_i}{q_i}\Phi(t; q_i, p_i) - \frac{\omega_0 t}{q_i}(q_i r_i - p_i s_i) + \Phi_0(r_i, s_i)\right) = \omega_{\mathrm{LF}}(p_i, q_i). \tag{2.12}$$

The notation $K(r_i, s_i)$ means the effective gain at the harmonic r_i/s_i, $\Phi_0(r_i, s_i)$ is the reference angle and $H(P)$, where the operator $P = \frac{d}{dt}$ is the open loop transfer function. Solving (2.12) is a difficult task. Let us observe that the RF signal at frequency ω_0 acts as a periodic perturbation of the Adler model of the PLL. If one neglects harmonic interactions, (2.12) may be simplified to the standard Arnold map model

$$\Phi_{n+1} = \Phi_n + 2\pi\Omega - c\ \sin\Phi_n, \tag{2.13}$$

where $\Omega = \frac{\omega}{\omega_0}$ is the bare frequency ratio and $c = \frac{K}{\omega_0}$. Such a nonlinear map is studied by introducing the winding number $\nu = \lim_{n\to\infty}(\Phi_n - \Phi_0)/(2\pi n)$. The limit exists everywhere as long as $c < 1$, the curve ν versus ω is a devil's staircase with steps attached to rational values $\Omega = \frac{p_i}{q_i}$ and width increasing with the coupling coefficient c. The phase-locking zones may overlap if $c > 1$ leading to chaos from quasi-periodicity [15].

The Arnold map is also a relevant model of a short Josephson junction shunted by a strong resistance R and driven by a periodic current of frequency ω_0 and amplitude I_0. Steps are found at the driving voltages $V_r = RI_0 = r(\hbar\omega_0/2e)$, r a rational number. Fundamental resonances $r = n$, n integer, have been used to achieve a voltage standard of relative uncertainty 10^{-7}.

To appreciate the impact of harmonics on the coupling coefficient one may observe that each harmonic of denominator q_i creates the same noise contribution $\delta\omega_{\mathrm{LF}} = q_i\delta\omega(t)$. There are $\phi(q_i)$ of them, where $\phi(q_i)$ is the Euler totient function, that is the number of integers less or equal to q_i and prime to it; the average coupling coefficient is thus expected to be $1/\phi(q_i)$. In [48] a more refined model is developed based on the properties of prime numbers. It is based on defining a coupling coefficient as $c^* = c\Lambda(n; q_i, p_i)$ with $\Lambda(n; q_i, p_i)$ a generalized Mangoldt function. The Mangoldt function is defined as

$$\Lambda(n) = \Lambda(n; 1, 1) = \begin{cases} \ln b & \text{if } n = b^k,\ b \text{ a prime,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.14}$$

The generalized Mangoldt function attached to the resonance $\frac{p_i}{q_i}$ adds the restriction that n should also be congruent to $p_i \pmod{q_i}$. The important result of that analysis is to exhibit a fluctuating average coefficient as follows

$$c^*_{\mathrm{av}}/c = \frac{1}{t}\sum_{n=1}^{t}\Lambda(n; q_i, p_i) = \frac{1}{\phi(q_i)} + \epsilon(t), \tag{2.15}$$

with $\epsilon(t) = O(t^{-1/2}\ln^2(t))$ which is known to be a good estimate as long as $q_i < \sqrt{t}$ [43]. The average coupling coefficient shows the expected dependance on q_i. In addition there is an arithmetical noise $\epsilon(t)$ with a low frequency dependance of the power spectrum reminding $1/f$ noise. Although that stage of the theory is not the last word of the story, it is quite satisfactory that this approach, based on phase-locking of the full set of harmonics, is accounting for the main aspects of $1/f$ noise found in experiments [43].

2.2 Hyperbolic phase-locking

> The concept of an automorphic function is natural generalization of that of a periodic function. Furthermore, an automorphic form is a generalisation of the exponential function $e(z) = \exp(2i\pi z)$.

The citation is from Iwaniec's book [28].

As shown in the previous section the homodyne detector behaves as a diophantine approximator of the frequency ratio ν of input oscillators. The approach is very satisfactory in explaining frequency-locking effects in the open loop, but the recourse to nonlinear differential equations is necessary for the case of the closed loop. Now we attempt to develop a pure arithmetical frame to account for the phase-locking effects as well. This is done by normalizing the beat frequency with respect to the low pass cut-off frequency as $y = \frac{\omega_{\rm LF}}{\omega_c}$ instead of the reference frequency. This is suggested by the geometry of continued fraction expansions which resorts to new mathematical concepts such as Ford circles, the hyperbolic half-plane, Möbius transformations, hyperbolic Laplace equation, ...

The hyperbolic half-plane is defined as the set $\mathbb{H}$ of complex numbers z such that $\Im z > 0$. It is a rich mathematical object first studied by the mathematician Henri Poincaré at the end of nineteenth century. The geometry of $\mathbb{H}$ is related to continued fraction expansions and it is also a natural frame to study the prime numbers. It will be used also for studying phase-locking effects. We put[4].

$$z = \nu + iy, \quad \nu = \frac{\omega}{\omega_0}, \quad 0 < y = \frac{\omega_{\rm LF}}{\omega_c} < 1. \tag{2.16}$$

It is clear that $y > 0$ is a condition which is imposed by counting measurements. The second condition $y < 1$ results from the low pass filtering and will reveal important in our introduction of Ford circles below. Let us start with the continued fraction expansion (2.7) which is rewritten as $\{a_0; a_1, \ldots, a_i\} = a_0 + \frac{1}{\{a_1; a_2, \ldots, a_i\}} = \frac{p_i}{q_i}$.

[4]The imaginary number i such that $i^2 = -1$ should not be confused with the index i in integers p_i, q_i and related integers.

By induction $p_0 = a_0$, $q_0 = 1$, $\{a_0; a_1\} = a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{p_1}{q_1}$, so that

$$\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} p_1 & p_0 \\ q_1 & q_0 \end{bmatrix},$$
$$\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_i & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} p_i & p_{i-1} \\ q_i & q_{i-1} \end{bmatrix}. \tag{2.17}$$

Taking determinants and using a property of continued fractions one gets $p_i q_{i-1} - p_{i-1} q_i = (-1)^{i-1}$. This suggests to associate to the convergents $\frac{p_i}{q_i}$ a group which defines Möbius transformations

$$z \to z' = \gamma(z) = \frac{p_i z + p_{i-1}}{q_i z + q_{i-1}}, \quad p_i q_{i-1} - p_{i-1} q_i = 1. \tag{2.18}$$

A Möbius transform maps a point $z \in \mathbb{H}$, with $\mathbb{H}$ the upper half-plane $\Im z > 0$ into a point $z' \in \mathbb{H}$. Möbius transforms form a group Γ called the modular group. It is a discontinuous group that is used to tesselate $\mathbb{H}$ by copies of a fundamental domain $\mathbb{F}$ under the group action. The fundamental domain of Γ (or modular surface) is defined as $\mathbb{F} = \{z \in \mathbb{H} : |z| \geq 1, \ |\nu| \leq \frac{1}{2}\}$, and the family of domains $\{\gamma(\mathbb{F}), \gamma \in \Gamma\}$ induces a tesselation of $\mathbb{H}$ [28].

We are also interested by the Ford circles [44]: they are defined by the images of the filtering line $z = \nu + i$ under all Möbius transformations (2.18). Rewriting (2.18) as $q_i z' - p_i = -1/(q_i z + q_{i-1})$ and inserting $z = \nu + i$ in that expression one immediately gets

$$\left| z' - \left(\frac{p_i}{q_i} + \frac{i}{2q_i^2} \right) \right| = \frac{1}{2q_i^2}. \tag{2.19}$$

which are circles of radius $\frac{1}{2q_i^2}$ centered at $\frac{p_i}{q_i} + \frac{i}{2q_i^2}$. To each $\frac{p_i}{q_i}$ a Ford circle in the upper-half plane can be attached, which is tangent to the real axis at $\nu = \frac{p_i}{q_i}$. Ford circles never intersect: they are tangent to each other if and only if they belong to fractions which are adjacent in the Farey sequence $\frac{0}{1} < \cdots \frac{p_1}{q_1} < \frac{p_1+p_2}{q_1+q_2} < \frac{q_1}{q_2} \cdots < \frac{1}{1}$ [54]. Ford circles can be considered a complex plane view of continued fractions.

Ford circles are related to an hyperbolic noise model: the Laplace equation for $\mathbb{H}$. To see this one first observe that $\mathbb{H}$ carries a non-Euclidean metric $dz = (d\nu^2 + dy^2)^{1/2}/y$. This is easily shown by using $\Im\gamma(z) = \frac{y}{|q_i z + q_{i-1}|^2}$ and $\frac{d\gamma(z)}{dz} = \frac{1}{(q_i z + q_{i-1})^2}$, since $\left|\frac{d\gamma(z)}{dz}\right| = \frac{\Im\gamma(z)}{y}$ under all group actions.

The invariance of the metric is inherited by the non-Euclidean Laplacian

$$\Delta = y^2 \left(\frac{\partial^2}{\partial \nu^2} + \frac{\partial^2}{\partial y^2} \right). \tag{2.20}$$

We will be concerned by eigenfunctions $\Psi_s(z)$ of (2.20) as our dynamical model of phase-locking. The complex parameter s has been introduced as a label for the eigenvalues. The relevant wave equation is

$$(\Delta + \lambda)\Psi_s(z) = 0. \tag{2.21}$$

The solutions of (2.21) are called automorphic forms.

They are several ways of finding eigenfunctions $\Psi_s(z)$ for a given eigenvalue λ. We will first consider elementary power law solutions, then we will show that there is a method to generate a lot of eigenvalues out of a fixed $\Psi_s(z)$ by shifting to $\Psi_s(\gamma(z))$, and still more by averaging over selected $\gamma(z)$ in Γ.

The simplest solutions of (2.21) are horizontal waves of eigenvalue $\lambda = s(s-1)$ in the form of power laws

$$\Psi_s(z) = \begin{cases} y^s, \\ y^{1-s}. \end{cases} \tag{2.22}$$

Any other wave which satisfies (2.21) can be searched by shifting z to $\gamma(z)$ given in (2.18) so that $y = \Im(z)$ goes to $\Im(\gamma(z))$ with the result

$$\Psi_s(z) = (\Im\gamma(z))^s = \frac{y^s}{|q_i z + q_{i-1}|^{2s}}. \tag{2.23}$$

The wavefronts for that solution are obtained by assigning to the quantity $y/(q_i z+q_{i-1})$ some constant value c. As a result one gets the equation $(x + q_{i-1}/q_i)^2 + (y - 1/(2q_i^2c^2))^2 = 1/(4c^4q_i^4)$ of circles tangent to the real axis at $x = -q_{i-1}/q_i$ and of radius $1/(2c^2q_i^2)$. At $c = 1$ such a circle is obtained from the Ford circle of index i by an horizontal slip from $\nu = p_i/q_i$ to q_{i-1}/q_i. It contracts ($c < 1$) or expands ($c > 1$) remaining tangent to the real axis.

A more general solution of (2.21) is obtained by summing solutions (2.23) over γ transformations in Γ taken over the right coset $\Gamma_\infty \backslash \Gamma$ where Γ_∞ is the subgroup of translations $z \to z + n$, n integer. One gets [32]

$$\Psi_s(z) = y^s\left(1 + \sum_i \frac{1}{|q_i z + q_{i-1}|^{2s}}\right), \tag{2.24}$$

where the summation index i means that the summation should be applied to all Farey fractions $\frac{p_i}{q_i}$.

The whole solution (2.24) can be decomposed into three contributions, the horizontal wave y^s is incident onto the filtering line, the reflected wave $S(s)y^{1-s}$ is scattered with a scattering coefficient $S(s)$ of modulus 1, whereas the remaining part $T_s(y, \nu)$ is a complex superposition of waves depending on y and the harmonics of $\exp(2i\pi\nu)$, but goes to zero away from the real line. More precisely one obtains $S(s) = A(s)Z(s)$, $Z(s) = \frac{\zeta(2s-1)}{\zeta(2s)}$, $A(s) = \frac{\Gamma(1/2)\Gamma(s-1/2)}{\Gamma(s)}$, $\Gamma(s)$ is the Gamma function and $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ is the Riemann zeta function. There is much to say about the singularities of the Riemann zeta function. It is analytic only on that part of the complex plane where $\Re s > 1$. But in the scattering coefficient it appears in the extended Riemann zeta function $\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$ which is analytic over the whole complex plane. The scattering coefficient equals the quotient $S(s) = \xi(2s-1)/\xi(2s)$. One knows from the theory of zeta function that $s = \frac{1}{2}$ is an axis of symmetry for $\zeta(s)$, and according to the Riemann hypothesis all non-trivial zeros lie on that axis. They show a very random distribution. An interesting case is thus to compute the scattering

coefficient $S(s)$ along the critical line $s = \frac{1}{2} + ik$ in which case the superposition $T_s(y, \nu)$ also vanishes. As a result one obtains

$$Z(1/2 + ik) \simeq \exp(2i\kappa(k)) \text{ with } \kappa'(k) = \frac{d \ln Z(s)}{ds} \text{ at } s = \frac{1}{2} + ik. \qquad (2.25)$$

In that expression we removed the smooth part $A(k)$ in the scattering coefficient so that its modulus is no longer 1, but its phase variability is still well described by the exponential factor above (see Figure 3). The aim of this replacement is to get an explicit link of the counting function $\kappa'(k)$ to the Mangoldt function $\Lambda(n)$ already accounted in (2.14) as an effective coupling coefficient in the Arnold map (2.13). It is known that the logarithmic derivative of $\zeta(s)$ is $-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n\geq 1} \frac{\Lambda(n)}{n^s}$. It is also easy to check that if we denote $Z(s) = B(2s-1)$, $B(s) = \frac{\zeta(s)}{\zeta(s+1)}$, one gets $-\frac{B'(s)}{B(s)} = \sum_{n\geq 1} \frac{b(n)}{n^s}$ with $b(n) = \frac{\Lambda(n)\phi(n)}{n}$ the modified Mangoldt function.

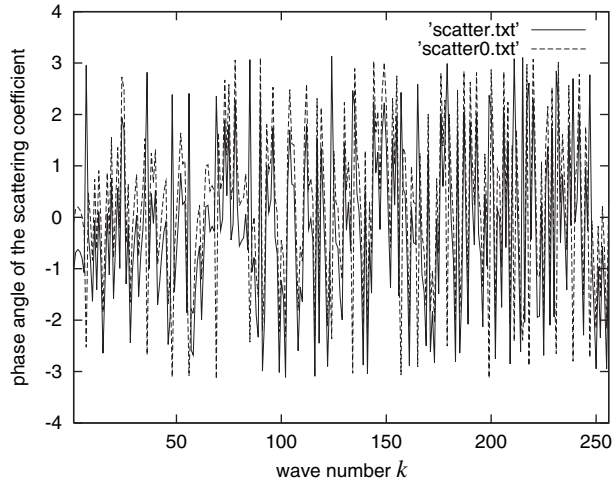


Figure 3. The phase angle $\kappa(k)$ for the scattering of noise waves on the modular surface. Plain lines: exact phase factor. Dotted lines: approximation based on the quotient $Z(1/2+ik)$ of two Riemann zeta functions at $2ik$ and $2ik+1$.

It is now clear that the scattering of waves in the hyperbolic plane has much to share with the phase-locking model formulated in Section 2.1. One argument in favor of that new frame is that the average modified Mangoldt function satisfies $\frac{1}{t}\sum_{n\geq 1} b(n) = 1 + \epsilon_B(t)$, where $\epsilon_B(t)$ is an error term of power spectral density equal to $1/f^{2G}$ with $G \simeq 0.618$ the Golden Ratio [44]. The approach is also related to hyperbolic models of quantum chaos [8].

3 Quantum phase-locking in quantum optics

3.1 Quantum phase states

> Apparently Dirac was the first to attempt a definition of a phase operator by means of an operator amplitude and phase decomposition. As we have discussed, with a complex c-number $a = Re^{i\Phi}$ one obtains the phase via $e^{i\Phi} = a/R$. Similarly, he sought to decompose the annihilation operator a into amplitude and phase components … After a brief calculation we obtain a relation indicating that the number operator N and phase operator Φ are canonically conjugate
>
> $$[N, \Phi] = 1. \tag{3.1}$$
>
> The equation immediately leads to a number-phase uncertainty relation which is often seen
>
> $$\delta N \delta \Phi \geq 1/2. \tag{3.2}$$
>
> However, all of the previous development founders upon closer examination.

This is taken from [34], a comprehensive review of the quantum phase problem. See also [62].

Let us start with the familiar Fock states of the quantized electromagnetic field (the photon occupation states) $|n\rangle$ which live in an infinite dimensional Hilbert space. They are orthogonal to each other: $\langle n|m\rangle = \delta_{mn}$, where δ_{mn} is the Dirac symbol. The states form a complete set: $\sum_{n=0}^{\infty} |n\rangle\langle n| = 1$.

The annihilation operator $\hat{a}$ removes one photon from the electromagnetic field

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad n = 1, 2, \ldots. \tag{3.3}$$

Similarly the creation operator $\hat{a}^\dagger$ adds one photon: $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, $n = 0, 1, \ldots$. These operators follow the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$. The operator $N = \hat{a}^\dagger\hat{a}$ represents the particle number operator and satisfies the eigenvalue equation $N|n\rangle = n|n\rangle$.

The algebra can be simplified by removing the normalization factor $\sqrt{n}$ in (3.3) defining thus the shift operator E as

$$E|n\rangle = |n-1\rangle, \quad n = 1, 2, \ldots. \tag{3.4}$$

Similarly $E^\dagger|n\rangle = |n+1\rangle$, $n = 0, 1, \ldots$. The operator E has been known as the (exponential) phase operator

$$E = e^{i\Phi} = (N+1)^{-1/2}\hat{a} = \sum_{n=0}^{\infty} |n\rangle\langle n+1|. \tag{3.5}$$

The problem with this definition is that the operator E is only one-sided unitary since $EE^\dagger = 1$, $E^\dagger E = 1 - |0\rangle\langle 0|$ due to the vacuum-state projector $|0\rangle\langle 0|$. The problem of defining a Hermitian quantum phase operator was already encountered by the founder

of quantum field theory P. A. M. Dirac [34]. The eigenvectors of E have been known as the Susskind–Glogower phase states [64]; they satisfy the eigenvalue equation

$$E|\psi\rangle = e^{i\psi}|\psi\rangle \quad \text{with} \quad |\Psi\rangle = \sum_{n=0}^{\infty} e^{in\psi}|n\rangle. \tag{3.6}$$

They are non-orthogonal to each other and form an overcomplete basis which solves the identity operator since $\frac{1}{2\pi}\int_{-\pi}^{\pi} d\psi|\psi\rangle\langle\psi| = 1$. When properly normalized the phase states $|\psi\rangle$ in (3.6) are related to the SU(1, 1) Perelomov coherent states [65], [41], [38], [17]. The operator $\cos\Phi = \frac{1}{2}(E + E^{\dagger})$ is used in the theory of Cooper pair box with a very thin junction when the junction energy $E_J \cos\Phi$ is higher than the electrostatic energy [11].

Further progress in the definition of phase operator was obtained by Pegg and Barnett [34], [40] in the context of a finite Hilbert space. From now on we consider a finite basis $|n\rangle$ of number states, with the index $n \in \mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$, the ring of integers modulo q. The phase states are now defined from the discrete Fourier transform (or more precisely from the quantum Fourier transform since the superposition is on Fock states not on real numbers)

$$|\theta_p\rangle = \frac{1}{\sqrt{q}} \sum_{n=0}^{q-1} \exp\left(2i\pi \frac{p}{q} n\right)|n\rangle. \tag{3.7}$$

The states are eigenstates of the Hermitian phase operator Θ_q such that $\Theta_q|\theta_p\rangle = \theta_p|\theta_p\rangle$ where $\Theta_q = \sum_{p=0}^{q-1} \theta_p|\theta_p\rangle\langle\theta_p|$ and $\theta_p = \theta_0 + 2\pi p/q$ with θ_0 a reference angle. It is implicit in the definition (3.7) that the Hilbert space is of finite dimension q. The states $|\theta_p\rangle$ form an orthonormal set and in addition the projector over the subspace of phase states is $\sum_{p=0}^{q-1} |\theta_p\rangle\langle\theta_p| = 1_q$, where 1_q is the unity operator. Given a state $|F\rangle$ one can write a probability distribution $|\langle\theta_p|F\rangle|^2$ which may be used to compute various moments, e.g. expectation values, variances. The key element of the formalism is that first the calculations are done in the subspace of dimension q, then the limit $q \to \infty$ is taken.

3.2 Quantum phase-locking

We are now in a position to define a quantum phase-locking operator. Our viewpoint has much to share with the classical phase-locking problem as soon as one reinterprets the fraction $\frac{p}{q}$ in (3.7) as arising from the resonant interaction between two oscillators and the dimension q as a number which defines the resolution of the experiment. From now on we emphasize such phase states $|\theta'_p\rangle$ which satisfy phase-locking properties and we impose the coprimality condition

$$(p, q) = 1. \tag{3.8}$$

The quantum phase-locking operator is defined as

$$\Theta_q^{\text{lock}} = \sum_p \theta_p |\theta_p'\rangle\langle\theta_p'|, \tag{3.9}$$

with $\theta_p = 2\pi \frac{p}{q}$ as the scalar coefficient[5] and the notation p means summation from 0 to $q-1$ with $(p,q)=1$. Using (3.7) and (3.8) in (3.9) one obtains

$$\Theta_q^{\text{lock}} = \frac{1}{q} \sum_{n,l} c_q(n-l)|n\rangle\langle l|, \tag{3.10}$$

where the range of values of n, l is from 0 to $\phi(q)$, and $\phi(q)$ is the Euler totient function (the number of inversible elements in the ring $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$). The coefficients in front of the outer products $|n\rangle\langle l|$ are the so-called Ramanujan sums

$$c_q(n) = \sum_p \exp\left(2i\pi \frac{p}{q} n\right) = \frac{\mu(q_1)\phi(q)}{\phi(q_1)} \quad \text{with} \quad q_1 = q/(q,n). \tag{3.11}$$

In the above equation $\mu(q)$ is the Möbius function, which is 0 if the prime number decomposition of q contains a square, 1 if $q=1$ and $(-1)^k$ if q is the product of k distinct primes. Ramanujan sums are relative integers which are quasi-periodic versus n with quasi-period $\phi(q)$, and aperiodic versus q with a type of variability imposed by the Möbius function. Ramanujan sums have been used for signal processing of low frequency noise [52], [50]. With the Ramanujan sum expansion the modified Mangoldt function introduced at the end of Section 2.1 is the dual of Möbius function

$$b(n) = \frac{\phi(n)}{n}\Lambda(n) = \sum_{q\geq 1} \frac{\mu(q)}{\phi(q)} c_q(n). \tag{3.12}$$

This illustrates that many "interesting" arithmetical functions carry the structure of prime numbers. We already mentioned the relation $d\ln\zeta(s)/ds = \sum_{n\geq 1} \frac{\Lambda(n)}{n^s}$, but there is also the relation $1/\zeta(s) = \sum_{n\geq 1} \frac{\mu(n)}{n^s}$. There is a well-known formulation of the Riemann hypothesis arising from the summatory Möbius function $\sum_{n=1}^{t} \mu(n) = O(t^{1/2+\epsilon})$, whatever ϵ [43].

Given a state $|\beta\rangle$ one can calculate the expectation value of the quantum phase-locking operator as

$$\langle\Theta_q^{\text{lock}}\rangle = \sum_p \theta_p \langle\theta_p'|\beta\rangle^2. \tag{3.13}$$

If one uses the finite form of Susskind–Glogower phase states (3.6) and a real parameter β

$$|\beta\rangle = \frac{1}{\sqrt{q}} \sum_{n=0}^{q-1} \exp(in\beta)|n\rangle, \tag{3.14}$$

the expectation value of the locked phase becomes

[5]The reference angle θ_0 introduced in [40] is irrelevant in our model.

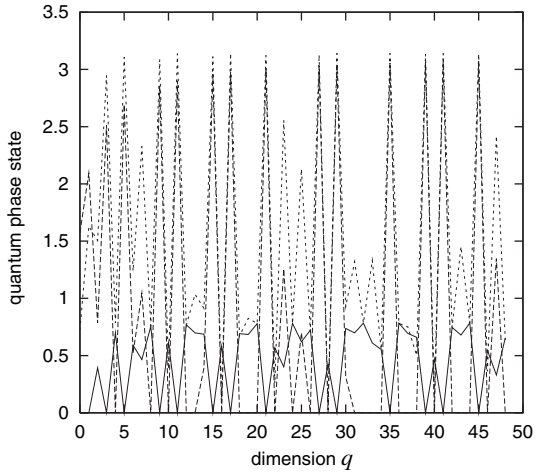


Figure 4. Oscillations in the expectation value (3.15) of the locked phase at $\beta = 1$ (dotted line) and their squeezing at $\beta = 0$ (plain line). The brokenhearted line which touches the horizontal axis is $\pi\Lambda(q)/\ln q$.

$$\langle \Theta_q^{\text{lock}} \rangle = \frac{\pi}{q^2} \sum_{n,l} c_q(l-n) \exp(i\beta(n-l)). \tag{3.15}$$

For $\beta = 1$ it is found that $\langle \Theta_q^{\text{lock}} \rangle$ has the more pronounced peaks at those values of q which are powers of a prime number. It can be approximated by the normalized Mangoldt function $\pi\Lambda(q)/\ln q$, as is shown in Figure 4. For $\beta = 0$ the expectation value of $\langle \Theta_q^{\text{lock}} \rangle$ is much lower. The parameter β can be used to minimize the phase uncertainty well below the classical value [49].

3.3 Primitive roots and $1/f$ noise

Let us consider the ring $\mathbb{Z}_q$. An important point about this set is that it is endowed with cyclic properties. If $q = p$, a prime number, then the period is $p - 1$; in any case there are cyclic subgroups in $\mathbb{Z}_q$. This can be shown by the examination of the algebraic equation

$$a^\alpha = 1 \pmod{q}. \tag{3.16}$$

Let us ask the question: what is the cycle of largest period in $\mathbb{Z}_q$? To find it one should look at the primitive roots which are defined as the solution of (3.16) such that the equation is wrong for any $1 \leq \alpha < q - 1$ and true only for $\alpha = q - 1$. If q is a prime number p the largest period is thus $\phi(p) = p - 1$. For instance Table 1 gives the cycle for $q = 7$. If $q = 2$, or 4, or $q = p^r$, a power of a prime number > 2, or $q = 2p^r$, twice the power of a prime number > 2, then a primitive root exists, and the largest cycle in the group is $\phi(q)$. For example $a = 2$ and $q = 3^2$ leads to the period $\phi(9) = 6 < q - 1 = 8$, as it is shown in Table 2.

Table 1. $(\mathbb{Z}/7\mathbb{Z})^*$ is a cyclic group of order $\phi(7) = 6$.

α	1	2	3	4	5	6	7	8
3^α	3	2	6	4	5	1	3	2

Table 2. $(\mathbb{Z}/3^2\mathbb{Z})^*$, is a cyclic group of order $\phi(9) = 6$.

α	1	2	3	4	5	6	7	8
2^α	2	4	8	7	5	1	2	4

Otherwise there is no primitive root. The period of the largest cycle in $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$ can still be calculated and is called the Carmichael Lambda function $\lambda(q)$. It is shown in Table 3 for the case $a = 3$ and $q = 8$. It is $\lambda(8) = 2 < \phi(8) = 4 < 8 - 1 = 7$.

Table 3. $(\mathbb{Z}/8\mathbb{Z})^*$ has a largest cyclic group of order $\lambda(8) = 2$.

α	1	2	3	4	5	6	7	8
3^α	3	1	3	1	3	1	3	1

Figure 5 shows the properly normalized period for the cycles in $\mathbb{Z}_q$. Its fractal character can be appreciated by looking at the corresponding power spectral density (FFT) shown in Figure 6. It has the form of a $1/f^\gamma$ noise, with $\gamma = 0.70$. For a more refined link between primitive roots, cyclotomy and Ramanujan sums see also [36].

Let us remind that the algebraic equation (3.16) is used as a numerical trap door for public key encryption (the RSA cryptosystem) [63], p. 125. While it is easy to multiply 1.000-digit numbers within a time of few milliseconds with the help of a modern classical computer, the opposite task of factorizing is impossible in a reasonable time. This impossibility would be lifted on a quantum computer thanks to Schor's algorithm of finding periods efficiently in $\mathbb{Z}_q$ (see Section 3.4).

3.4 Quantum computation and the Bost–Connes algebra

The quantum operator corresponding to (3.16) is another shift μ_a in the space of number states, where $a \in \mathbb{Z}_q$ and $(a, n) = 1$. It is defined as

$$\mu_a |n> = |an \pmod q\rangle, \tag{3.17}$$

It is multiplicative: $\mu_k \mu_l = \mu_{kl}$ and $\mu_k^* \mu_k = 1$, $k, l \in \mathbb{N}$. The eigenvalues and eigenvectors of the operator (3.17) allow to define the order $r = \mathrm{ord}_q(a)$ of $a \pmod q$ which is the smallest exponent such that (3.16) is satisfied. When the order is $r = \phi(q)$ then a s a primitive root. This results from Euler's theorem which

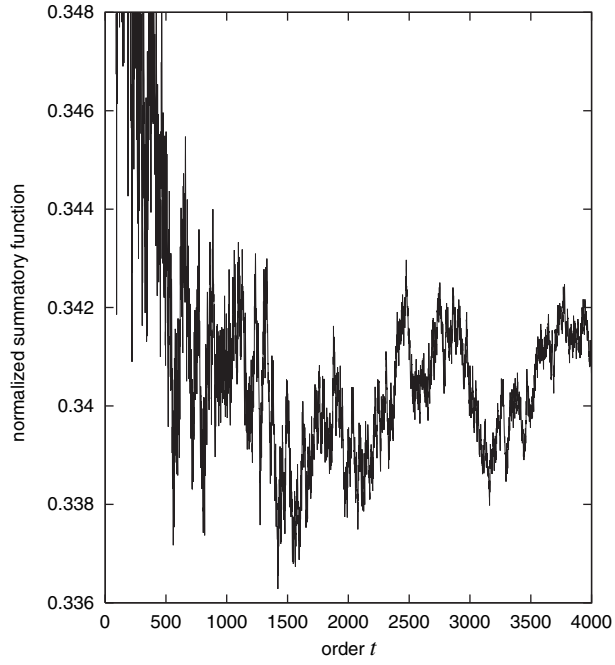


Figure 5. Normalized Carmichael lambda function: $(\sum_1^t \lambda(n))/t^{1.90}$.

states that $a^{\phi(q)} = 1 \pmod q$. If the order is the same for any a such that $(a, q) = 1$ it is also called a universal exponent. Thus $\phi(q)$ is a universal exponent. It is the smallest one if a is a primitive root, otherwise there is a smaller one which is Carmichael's lambda function introduced above. Let us summarize

$$ord_q(a) \leq \lambda(q) \leq \phi(q) \leq q - 1. \tag{3.18}$$

The eigenvalues of operator (3.17) are of the form $\exp\left(2i\pi\frac{k}{r}\right)$ and the corresponding eigenvectors are given as a quantum Fourier transform [21], p. 22.

$$|u_k\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} \exp\Big(- 2i\pi \frac{kj}{r}\Big)|a^j \bmod q\rangle. \tag{3.19}$$

We observe that (3.19) refers to the subspace of dimension r (the least period) from the powers of a in (3.16). One sees that the u_k's acquire the status of phase states in the cyclic subspace of $\mathbb{Z}_q$ of dimension r.

The operator μ_a is used in the context of quantum computation in relation to Schor's algorithm [21]. The goal is to factor integers efficiently (in polynomial time instead of the exponential time needed on the classical computer). It is in fact related to the efficient estimation of the period r taken from the eigenvalue $\exp\left(2i\pi\frac{k}{r}\right)$, and to the efficient implementation of quantum Fourier transforms on the quantum computer.

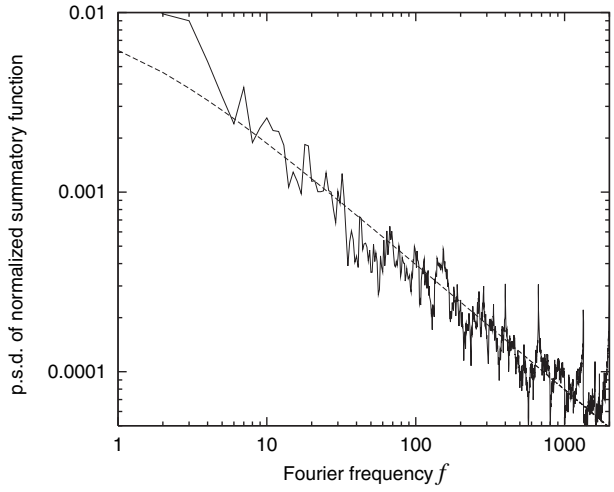


Figure 6. FFT of the normalized Carmichael lambda function. The straight line has slope -0.70.

The shift operator (3.17) which quantizes (3.16) plays the role of multiplication in the language of operators. The addition operator is easier to define

$$e_p|n\rangle = \exp\left(2i\pi\frac{p}{q}n\right)|n\rangle. \tag{3.20}$$

It encodes the individuals in the quantum Fourier transform (3.7). These individuals are eigenvalues of the operator e_p. One gets $e_0 = 1$, $e_p^* = e_{-p}$, $e_l e_m = e_{l+m}$. Both operators μ_a and e_p form an algebra $A = (\mu_a, e_p)$. It was used by Bost and Connes [10], [14] to build a remarkable quantum statistical model undergoing a phase transition.

3.5 The "quantum Fechner law" and its thermodynamics

Let us now make use of the concepts of quantum statistical mechanics [46], [10]. Given an observable Hermitian operator M and a Hamiltonian H_0 one has the evolution $\sigma_t(M)$ versus time t

$$\sigma_t(M) = e^{itH_0} M e^{-itH_0}, \tag{3.21}$$

and the expectation value of M at inverse temperature $\beta = \frac{1}{kT}$ is the unique Gibbs state

$$\text{Gibbs}(M) = \text{Trace}(Me^{-\beta H_0})/\,\text{Trace}(e^{-\beta H_0}). \tag{3.22}$$

For the more general case of an algebra of observables A, the Gibbs state is replaced by the so-called Kubo–Martin–Schwinger (or KMS$_\beta$) state. One introduces a 1-parameter group σ_t of automorphisms of A so that for any $t \in R$ and $x \in A$

$$\sigma_t(x) = e^{itH_0} x e^{-itH_0}, \tag{3.23}$$

but the equilibrium state remains unique only if some conditions regarding the evolution of the operators $x \in A$ are satisfied: the so-called KMS conditions [10]. It happens quite often that there is a unique equilibrium state above a critical temperature T_c and a coexistence of many equilibrium states at low temperature $T < T_c$. There is thus a spontaneous symmetry breaking at the critical temperature T_c. A simple example is the phase diagram for a ferromagnet. At temperature larger than $T \simeq 10^3$ K the disorder dominates and the thermal equilibrium state is unique, while for $T < T_c$ the individual magnets tend to align each other, which in the three-dimensional space $\mathbb{R}^3$ of the ferromagnet yields a set of extremal equilibrium phases parametrized by the symmetry group SO(3) of rotations in $\mathbb{R}^3$.

In the Bost–Connes approach the Hamiltonian operator is given as the logarithm of the number operator $H_0 = \ln N$, i.e. it is a kind of "quantum Fechner law" [46] since the "sensor" operator H_0 is the logarithm of the "physical" operator N, in contrast to the ordinary Hamiltonian $H_0 = N + \frac{1}{2}$ of the harmonic oscillator. The dynamical system (A, σ_t) is defined by its action on the Fock states

$$H_0|n\rangle = \ln n|n\rangle, \quad n \geq 1. \tag{3.24}$$

Using the relations $e^{-\beta H_0}|n\rangle = e^{-\beta \ln n}|n\rangle = n^{-\beta}|n\rangle$, it follows that the partition function of the model at the inverse temperature β is

$$\text{Trace}(e^{-\beta H_0}) = \sum_{n=1}^{\infty} n^{-\beta} = \zeta(\beta), \quad \Re\beta > 1, \tag{3.25}$$

where $\zeta(\beta)$ is the Riemann zeta function introduced at the end of Section 2.1. Applying (3.23) to the Hamiltonian H_0 in (3.24) one gets

$$\sigma_t(\mu_a) = a^{it}\mu_a \quad \text{and} \quad \sigma_t(e_p) = e_p. \tag{3.26}$$

One observes that under the action of σ_t the additive phase operator (3.20) is invariant, while the multiplicative shift operator (3.17) oscillates in time at angular frequency $\ln a$. It is found in [10], [14] that the pole $\beta = 1$ of the Riemann zeta function separates two dynamical regimes.

In the high temperature regime $0 < \beta \leq 1$ there is a unique KMS_β state of the dynamical system (A, σ_t). The expectation value $\psi_\beta\left(\frac{p}{q}\right)$ for the irreducible fractions $\frac{p}{q}$, $(p, q) = 1$ is

$$\psi_\beta(\frac{p}{q}) = \prod_{\substack{b \text{ prime} \\ k_b \neq 0}} b^{-k_b\beta} \frac{1 - b^{\beta-1}}{1 - b^{-1}}, \tag{3.27}$$

and the k_b are the exponents in the unique prime decomposition of the denominator q

$$q = \prod_{b \text{ prime}} b^{k_b}. \tag{3.28}$$

In the low temperature regime $\beta > 1$ things become more involved. The KMS_β state is no longer unique as it is in the case of the low temperature ferromagnet. The

symmetry group G of the dynamical system (A, σ_t) is a subgroup of the automorphisms of A which commutes with σ_t

$$w \circ \sigma_t = \sigma_t \circ w, \quad \text{for all } w \in G,\ t \in \mathbb{R}, \tag{3.29}$$

where $\circ$ means the composition law in G. When the temperature is high the system is disordered enough that there is no interaction between its constituents and the equilibrium state of the system does not see the action of the symmetry group G, which explains why the thermal state is unique. At lower temperature $T < T_c$ the constituents of the system may interact. The symmetry group G then permutes a family of extremal KMS_β states generating the possible states of the system after phase transition. The symmetry group for the algebra $A = (\mu_a, e_p)$ and the "quantum Fechner Hamiltonian" H_0 in (3.24) is the Galois group $G = \mathrm{Gal}(\mathbb{Q}^{\mathrm{cycl}}/\mathbb{Q})$ of the cyclotomic extension $\mathbb{Q}^{\mathrm{cycl}}$ of the field of rational numbers $\mathbb{Q}$. The field $\mathbb{Q}^{\mathrm{cycl}}$ is the generated over $\mathbb{Q}$ by all the roots of unity. The Galois group G is the automorphism group of $\mathbb{Q}^{\mathrm{cycl}}$ which preserves the subfield $\mathbb{Q}$. It is isomorphic to $\mathbb{Z}_q^* = (\mathbb{Z}/q\mathbb{Z})^*$ which is the group of integers modulo q with the zero element removed.

One immediately sees that the action of the Galois group G commutes with the 1-parameter group defined in (3.26). Hence G permutes the extremal KMS_β states of (A, σ_t). To describe these thermal states for $\beta > 1$ one starts from the action of operators μ_a and e_p on Fock states as given in (3.17) and (3.20), and one allows the permutation by the symmetry group G. For each $w \in G$, one has a representation π_w of the algebra A so that

$$\begin{aligned} \pi_w(\mu_a)|n\rangle &= |an \pmod q\rangle, \\ \pi_w(e_p)|n\rangle &= w(\exp\left(2i\pi n\frac{p}{q}\right)|n\rangle, \end{aligned} \tag{3.30}$$

and it is shown in [10] that the state

$$\phi_{\beta,w}(x) = \zeta(\beta)^{-1}\,\mathrm{Trace}\left(\pi_w(x)e^{-\beta H_0}\right), \quad x \in A \tag{3.31}$$

is a KMS_β state for (A, σ_t). The action of G on A induces an action on these thermal states which permutes them. For $\beta > 1$ the expectation value acting on the phase operator e_p may be written as $\mathrm{KMS}(e_p) = q^{-\beta}\prod_{\substack{b \text{ divides } q \\ b \text{ prime}}} \frac{1-b^{\beta-1}}{1-b^{-1}}$. But it can be also found that formula (3.27) is valid in the whole range of temperatures $0 \le \beta \le \infty$. Plotting the thermal state (Figure 7) one observes the following asymptotes. In the high temperature limit β=0 the thermal state tends to 1; in the low temperature range $\beta \gg 1$ the thermal state is well approximated by the function $\mu(q)/\phi(q)$, where the Möbius function $\mu(q)$ is 0 if the prime decomposition of q contains a square, 1 if $q = 1$ and $(-1)^k$ if q is the product of k distinct primes. Close to the critical temperature at $\beta = 1$ the fluctuations are squeezed and the thermal state is quite close to $\epsilon\Lambda(q)/q$, when $\epsilon = 1 - \beta \to 0$, where $\Lambda(q)$ is the Mangoldt function defined in (2.14). It is an unexpected surprise that phase fluctuations at the critical region of the quantum perception model resemble those of the classical phenomenological model of time measurements.

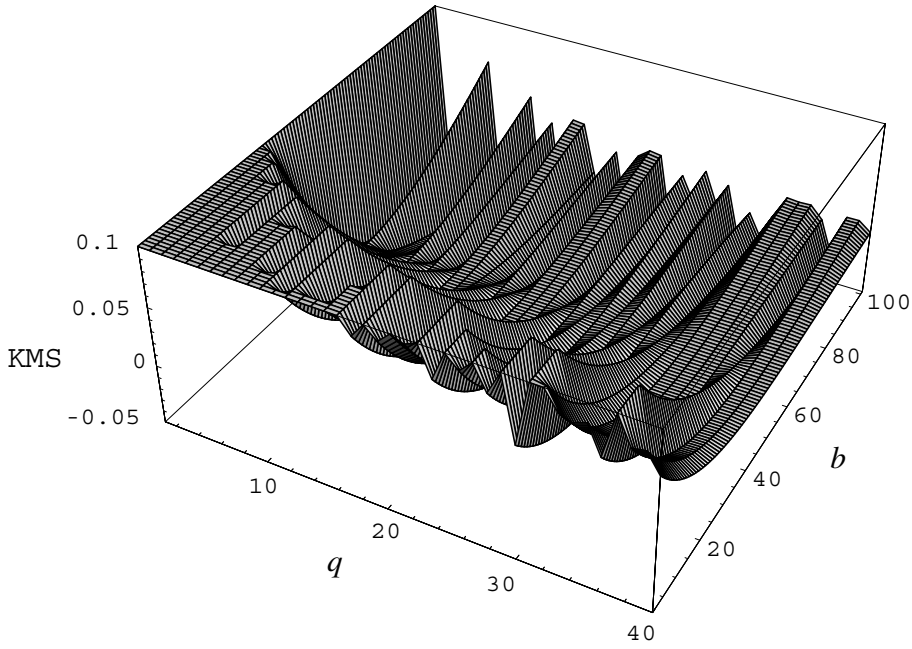


Figure 7. The KMS thermal state as given from (3.27). The inverse temperature scale is from $\beta = 0.5$ to 1.5. The dimension is from $q = 1$ to 40 as shown. Some parts of the graphics are truncated in the high temperature region $\beta < 1$. Thermal fluctuations are clearly reduced at the inverse critical temperature $\beta = 1$.

The normalized Möbius function $\mu(q)/\phi(q)$ is the dual to the Mangoldt function $\Lambda(n)$ when projected onto the basis of Ramanujan sums [45]. Ramanujan sums $c_q(n)$, which are defined as the sums of n^{th} powers of primitive roots of unity, generalize the cosine function, being quasi-periodic in n and aperiodic in q. The thermal states in the transition region can thus be considered as an unfolding through the dimensions q of the unique high temperature thermal state. This also suggests to consider the inverse temperature β and the dimension q of the Hilbert space as complementary to each other in the sense of Heisenberg.

4 Phase in quantum information

4.1 Introduction

As a visionary founder of contemporary physics, Galileo Galilei wrote:

> Philosophy (nature) is written in that great book which ever lies before our eyes. I mean the universe, but we cannot understand it if we do not

> first learn the language and grasp the symbols in which it is written. The book is written in the mathematical language … without whose help it is humanly impossible to comprehend a single word of it, and without which one wanders in vain through a dark labyrinth.

Even without advocating the unity of physics and mathematics, it is becoming a reality that the concepts of quantization invade mathematics after having profoundly changed physics. Problems pertinent to quantum information theory are touching more and more branches of pure mathematics, such as number theory, abstract algebra and projective geometry. This paper focuses on one of the most prominent issues in this respect, namely the construction of sets of mutually unbiased bases (MUBs) in a Hilbert space of finite dimension. An updated list of open problems related to the development of quantum technologies can be, for example, found in [53].

To begin with, one recalls that two different orthonormal bases A and B of a d-dimensional Hilbert space $\mathcal{H}^d$ with metrics $\langle \dots | \dots \rangle$ are called mutually unbiased if and only if $|\langle a|b\rangle| = 1/\sqrt{d}$ for all $a \in A$ and all $b \in B$. An aggregate of mutually unbiased bases is a set of orthonormal bases which are pairwise mutually unbiased. It has been found [70] that the maximum number of such bases cannot be greater than $d+1$. It is also known that this limit is reached if d is a power of prime. Yet, a still unanswered question is if there are non-prime-power values of d for which this bound is attained. It is surmised [70], [30] that the maximum number of such bases, $N(d)$, is equal to $1 + \min(p_i^{e_i})$, the latter quantity being the lowest factor in the prime number decomposition of d, $d = \prod_i p_i^{e_i}$. But, for example, it is still not known [25] whether there are more than three MUBs for $d = 6$, the lowest non-prime-power dimension, although the latest findings of Wootters [69] (and an earlier result of G. Tarry quoted in the last reference) seem to speak in favor of this conjecture.

MUBs have already been recognized to play an important role in quantum information theory. Their main domain of applications is the field of secure quantum key exchange (quantum cryptography). This is because any attempt by an eavesdropper to distinguish between two non-orthogonal quantum states shared by two remote parties will occur at the price of introducing a disturbance into the signal, thus revealing the attack and allowing to reject the corrupted quantum data. Until recently, most quantum cryptography protocols have solely relied, like the original BB84 one, upon 1-qubit technologies[6], i.e. on the lowest non-trivial dimension ($d = 2$), usually the polarization states of a single photon, or other schemes such as the sidebands of phase-modulated light [35]. But security against eavesdropping has lately been found to substantially increase by using all the three bases of qubits, employing higher dimensional states, e.g. qudits [37], [13], or even entanglement-based protocols [20]. Another closely related application of MUBs is so-called quantum state tomography, i.e. the most efficient way to decipher an unknown quantum state [53].

Quantum state recovery and secure quantum key distribution can also be furnished in terms of so-called positive operator valued measures (POVMs) which are sym-

[6]The quantum bit is the smallest unit of information in quantum computing. Qubits hold an exponentially larger amount of information than traditional bits.

metric informationally complete (SIC-POVMs) [55]. These are defined as sets of d^2 normalized vectors a and b such that $|\langle a|b\rangle| = 1/\sqrt{d+1}$, where $a \neq b$, and they are, obviously, very intimately connected with MUBs. Unlike the latter, however, the SIC-POVMs can exist in all finite dimensions and they have already been constructed for $d = 6$ [25]. The intricate link between MUBs and SIC-POVMs has recently been examined by Wootters [69] and acquired an intriguing geometrical footing in the light of the "SPR conjecture" [60] stating that the question of the existence of a set of $d+1$ mutually unbiased bases in a d-dimensional Hilbert space if d differs from a power of a prime number is equivalent to the problem of whether there exist projective planes whose order d is not a power of a prime number.

The section is organized as follows. In Sections 4.2 and 4.3 the construction of a maximal set of MUBs in dimension $d = p^m$, p being a prime, as a quantum Fourier transform acting on a Galois field (p odd) and a Galois ring GR(4^m) ($p = 2$) is discussed. This puts in perspective the earlier formulas by [70] and [30], respectively. The case of non-prime-power dimensions is briefly examined in Section 4.4. In Section 4.5, we focus on our recent conjecture on the equivalence of two problems: the surmised nonexistence of projective planes whose order is not a power of a prime and the suspected non-existence of a complete set of MUBs in Hilbert spaces of non-prime-power dimensions. The geometry of qubits is discussed and the concept of a lifted Fano plane is introduced. Finally, an intricate relationship between MUBs and maximal entanglement is emphasized, which promises to shed fresh light on newly emerging concepts such as the distillation of mixed states and bound entanglement. The exposition of the theory should be self-containing. Yet, the interested reader may find it helpful to consult some introductory texts on quantum theory in a finite Hilbert space and its relation to Fourier transforms and phase space methods, e.g., the review by A. Vourdas [66], or our more advanced text [51].

4.2 MUBs, quantum Fourier transforms and Galois fields

In this subsection we shall examine a close connection between MUBs and Fourier transforms. Let us consider an orthogonal computational basis

$$B_0 = (|0\rangle, |1\rangle, \ldots, |n\rangle, \ldots, |d-1\rangle)$$

with indices n in the ring $\mathbb{Z}_d$ of integers modulo d. There is a dual basis which is defined by the quantum Fourier transform

$$|\theta_k\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} \omega_d^{kn} |n\rangle, \tag{4.1}$$

where $k \in \mathbb{Z}_d$, $\omega_d = \exp\left(\frac{2i\pi}{d}\right)$ and $i^2 = -1$. In the context of quantum optics this Fourier transform relates Fock states $|k\rangle$ of light to the so-called phase states $|\theta_k\rangle$. The properties of the quantum phase operator underlying this construction have extensively been studied and found to be linked to prime number theory [50].

Let us start with $d = 2$, i.e. the case of qubits where $\omega = -1$ and so

$$|\theta_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |\theta_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{4.2}$$

These two vectors can also be obtained by applying the Hadamard matrix $H = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ to the basis $(|0\rangle, |1\rangle)$. Note that the two orthogonal bases $B_0 = (|0\rangle, |1\rangle)$ and $B_1 = (|\theta_0\rangle, |\theta_1\rangle)$ are mutually unbiased. The third base $B_2 = (|\psi_0\rangle, |\psi_1\rangle)$ which is mutually unbiased to both B_0 and B_1 is obtained from H by the pre-action of a $\pi/2$ rotation $S = \left[\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right]$, so that $HS = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & i \\ 1 & -i \end{smallmatrix}\right]$. The three matrices (I, H, HS) thus generate the three mutually unbiased bases. These matrices are also important for two qubits gates in quantum computation [37].

The above-outlined strategy for finding MUBs for qubits contrasts with that used by a majority of authors. The eigenvectors of Pauli spin matrices $\sigma_z = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, $\sigma_x = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, $\sigma_y = \left[\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right]$, where $\sigma_y = i\sigma_x\sigma_z$, are precisely the sought bases B_0, B_1 and B_2. A natural generalization of Pauli operators σ_x and σ_z for an arbitrary dimension d is the Pauli group of shift and clock operators:

$$\begin{aligned} X_d|n\rangle &= |n+1\rangle, \\ Z_d|n\rangle &= \omega_d^n|n\rangle. \end{aligned} \tag{4.3}$$

For a prime dimension $d = p$, it can be shown that the eigenvectors of the unitary operators $(Z_p, X_p, X_pZ_p, \ldots, X_pZ_p^{p-1})$ generate the set of $d+1$ MUBs [3]. A natural question here emerges whether this method can straightforwardly be generalized to any dimension.

To this end in view, let us attempt to rewrite eq. (4.1) in such a way that the exponent of ω_d now acts on the elements of a Galois field $G = \mathrm{GF}(p^m)$, the finite field of characteristic p and cardinality $d = p^m$. Denoting "$\oplus$" and "$\odot$" the two usual operations in the field and replacing ω_d by the root of unity ω_p, we get

$$|\theta_k\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} \omega_p^{k\odot n}|n\rangle. \tag{4.4}$$

Next, we employ the Euclidean division theorem for fields [33], which says that given any two polynomials k and n in G there exists a uniquely determined pair a and b in G such that $k = a \odot n \oplus b$, $\deg b < \deg a$. This allows for the exponent in eq. (4.4), E, to be written as $E = (a \odot n \oplus b) \odot n$. In the case of prime dimension $d = p$, E is an integer. Otherwise E is a polynomial and eq. (4.4) generalizes to

$$|\theta_b^a\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} \omega_p^{\mathrm{tr}[(a\odot n\oplus b)\odot n]}|n\rangle, \tag{4.5}$$

where "tr" stands for the trace of $\mathrm{GF}(p^m)$ down to $\mathrm{GF}(p)$,

$$\mathrm{tr}(E) = E \oplus E^p \oplus \cdots \oplus E^{p^{m-1}}, \quad E \in \mathrm{GF}(p^m). \tag{4.6}$$

In a finite field of odd characteristic p, eq. (4.5) defines the set of d bases, with the index a for the base and the index b for the vector in the base, mutually unbiased to each other and to the computational base B_0 as well. In a slightly different form this equation was first derived by Wootters [70]. Its nice short elucidation, based on Weil sums, is due to Klappenecker et al. [30]. Another more tricky derivation still in the spirit of Fourier transforms and claimed to also apply to the case of characteristic 2 was found by Durt [19]. A very recent approach based on the Weyl operators in the L^2-space over Galois fields is also worth mentioning [39].

As already pointed out in [70], the reason why (4.5) defines the complete set of MUBs relies on the field theoretical formula for the norm $\left| \sum_{n=0}^{d-1} \omega_p^{\mathrm{tr}[(a\odot n\oplus b)\odot n]} |n\rangle \right| = p^{1/2}$, with $a \neq 0$ and p being an odd prime. This method, however, fails for characteristic two where $\left| \sum_{n=0}^{d-1} \omega_2^{\mathrm{tr}[(a\odot n\oplus b)\odot n]} |n\rangle \right| = 0$ for any a, b. As shown in Section 3 below, here one has to use Galois rings instead of Galois fields to find a complete set of MUBs.

A closer inspection of (4.5) reveals an intricate relation between MUBs and quantum phase operators. It is known [50] that the Fourier basis $|\theta_k\rangle$ can be derived in terms of the eigenvectors of a quantum phase operator with eigenvalues θ_k and given by $\Theta_d = \sum_{k=0}^{d-1} \theta_k |\theta_k\rangle\langle\theta_k|$. Similarly, using well known properties of the field trace, one can show that each base of index a can be associated with a quantum phase operator

$$\Theta_d^a = \sum_{b=0}^{d-1} \theta_b^a |\theta_b^a\rangle\langle\theta_b^a|, \tag{4.7}$$

with eigenvectors $|\theta_b^a\rangle$ and eigenvalues $\theta_b^a = 2\pi \frac{b}{q}$; the latter may thus be called an "MUB operator". For still a more explicit relation of quantum phase states to the additive Fourier transform over Galois fields (and rings) and to the study of phase fluctuations see [51].

4.3 MUBs for even characteristic from Galois rings

Our next goal is to find a Fourier transform formulation of MUBs in characteristic 2. Eq. (4.3), as it stands, is in principle valid for any power of a prime, $d = p^m$, thus also for 2^m-dits, and one may therefore be tempted to connect the Galois field algebra and generalized Pauli operators (3) by constructing discrete vector spaces over the Galois field [24]. For the one qubit case we already know that the eigenvectors of Pauli matrices σ_z, σ_x and $\sigma_x\sigma_z$ define the three MUBs. Passing to the quartit (i.e., 4-dit) case, one finds that the operators of the following tensorial products $\sigma_z \otimes \sigma_x$, $\sigma_z \otimes \sigma_x\sigma_z$ and $\sigma_x\sigma_z \otimes \sigma_z$ are associated to translations, i.e. to a single line in the corresponding vector space, and they define a unique basis represented by their simultaneous eigenvectors. Since there are $4+1$ lines in this discrete vector space, there are also $4+1$ MUBs. Other geometrically inspired derivations based on the tensorial decomposition of operators in the Pauli group can be found in [3], [56], [42].

Now, let us try adjusting eq. (4.4) for the case of characteristic two. Instead of the Euclidean division in the field $\mathrm{GF}(2^m)$, it is necessary to consider a decomposition in the Galois ring $\mathrm{GR}(4^m)$ (defined below) so that the relevant root of unity in the Fourier formula now reads $\omega_4 = \exp(2i\pi/4) = i$. For qubits $\mathrm{GR}(4) = \mathbb{Z}_4$, and since any number k in $\mathbb{Z}_4$ can be written as $k = a \oplus 2 \odot b$, eq. (4.4) transforms into

$$|\theta_b^a\rangle = \frac{1}{\sqrt{2}} \sum_{n=0}^{1} i^{(a\oplus 2\odot b)\odot n} |n\rangle, \tag{4.8}$$

where $\oplus$ and $\odot$ now act in $\mathbb{Z}_4$. We note that the bases are identical to those obtained earlier from eq. (4.1), i.e. $B_1 = (|\theta_0^0\rangle, |\theta_1^0\rangle)$ and $B_2 = (|\theta_0^1\rangle, |\theta_1^1\rangle)$.

To generalize further this formula one needs to introduce some abstract algebra. First one recalls that the Galois field $\mathrm{GF}(p^m)$ is the field of polynomials defined as the residue class ring $\mathbb{Z}_p[x]/(q(x))$ of the ring of polynomials $\mathbb{Z}_p[x]$ by a primitive polynomial of order m over $\mathbb{Z}_p = \mathrm{GF}(p)$. By definition, this primitive element $\alpha = q(x)$ has the property to be irreducible over the base field $\mathrm{GF}(p)$, i.e. it cannot be factored into products of lesser-degree polynomials; it is also primitive over $\mathrm{GF}(p)$ of order $p-1$ in the sense that it generates any non-zero element of $\mathrm{GF}(p)$ by a power sequence $(\alpha^1, \alpha^2, \ldots, \alpha^{p-1} = 1)$ and in addition all of its roots are in the extension field $\mathrm{GF}(p^m)$. There is at least one primitive polynomial for any extension field $\mathrm{GF}(p^m)$. For $p = 2$ and $m = 2$, 3 and 4 they are, for example, of the form $q(x) = x^2 + x + 1$, $x^3 + x + 1$ and $x^4 + x + 1$, respectively.

A Galois ring $\mathrm{GR}(4^m)$ of order m is a ring of polynomials which is an extension of $\mathbb{Z}_4$ of degree m containing an r-th root of unity [26], [67]. Let $h_2(x) \in \mathbb{Z}_2[x]$ be a primitive irreducible polynomial of degree m. There is a unique monic polynomial $h(x) \in \mathbb{Z}_4[x]$ of degree m such that $h_2(x) = h(x) \pmod 2$ and $h(x) \pmod 4$ divides $x^r - 1$, where $r = 2^m - 1$. The polynomial $h(x)$ is the basic primitive polynomial and defines the Galois ring $\mathrm{GR}(4^m) = \mathbb{Z}_4/(h(x))$ of cardinality 4^m. This ring can be found as follows. Let $h_2(x) = e(x) - d(x)$, where $e(x)$ contains only even powers and $d(x)$ only odd powers; then $h(x^2) = \pm(e^2(x) - d^2(x))$. For $m = 2$, 3 and 4 one gets $h(x) = x^2 + x + 1$, $x^3 + 2x^2 + x - 1$ and $x^4 + 2x^2 - x + 1$, respectively.

Any non-zero element of $\mathrm{GF}(p^m)$ can be expressed in terms of a single primitive element. This is no longer true in $\mathrm{GR}(4^m)$, which contains zero divisors. But in the latter case there exists a non-zero element ξ of order $2^m - 1$ which is a root of the basic primitive polynomial $h(x)$. Any element $\beta \in \mathrm{GR}(4^m)$ can be uniquely determined in the form $\beta = a \oplus 2 \odot b$, where a and b belong to the so-called Teichmüller set $\mathbb{T}_m = (0, 1, \xi, \ldots, \xi^{2^m-2})$. Moreover, one finds that $a = \beta^{2^m}$. We can also define the trace to the base ring $\mathbb{Z}_4$ by the map

$$\mathrm{tr}(\beta) = \sum_{k=0}^{m-1} \sigma^k(\beta), \tag{4.9}$$

where the summation runs over $\mathrm{GR}(4^m)$ and the Frobenius automorphism σ reads

$$\sigma(a \oplus 2 \odot b) = a^2 \oplus 2 \odot b^2, \tag{4.10}$$

with $a^2 \equiv a \odot a$. Using the 2-adic decomposition of k in the exponent of (4.4) and the above-given trace map, we finally get

$$|\theta_b^a\rangle = \frac{1}{\sqrt{2^m}} \sum_{n=0}^{2^m-1} i^{\mathrm{tr}[(a \oplus 2 \odot b) \odot n]} |n\rangle. \tag{4.11}$$

The last expression gives a set of $d = 2^m$ bases with index a for the base and index b for the vectors in the base, mutually unbiased to each other and to the computational base B_0 [30].

Let us apply this formula to the case of quartits. In $\mathrm{GR}(4^2) = \mathbb{Z}_4[x]/(x^2 + x + 1)$ the Teichmüller set reads $\mathbb{T}_2 = (0, 1, x, 3 + 3x)$; the 16 elements $a \oplus 2 \odot b$ with a and b in $\mathbb{T}_2$ are shown in the following matrix

$$\begin{bmatrix} 0 & 2 & 2x & 2+2x \\ 1 & 3 & 1+2x & 3+2x \\ x & 2+x & 3x & 2+3x \\ 3+3x & 1+3x & 3+x & 1+x \end{bmatrix}.$$

Extracting the Teichmüller decomposition $(a \oplus 2 \odot b) \odot n = a' \oplus 2 \odot b'$ and calculating the exponent $\mathrm{tr}(a' \oplus 2 \odot b') = a' \oplus 2 \odot b' \oplus a'^2 \oplus 2 \odot b'^2$ one gets the four MUBs:

$$\begin{aligned} B_1 &= (1/2)\{(1,1,1,1), (1,1,-1,-1), (1,-1,-1,1), (1,-1,1,-1)\}, \\ B_2 &= (1/2)\{(1,-1,-i,-i), (1,-1,i,i), (1,1,i,-i), (1,1,-i,i)\}, \\ B_3 &= (1/2)\{(1,-i,-i,-1), (1,-i,i,1), (1,i,i,-1), (1,i,-i,1)\}, \\ B_4 &= (1/2)\{(1,-i,-1,-i), (1,-i,1,i), (1,i,1,-i), (1,i,-1,i)\}. \end{aligned} \tag{4.12}$$

The case of 8-dits can be examined in a similar fashion, with the ring $\mathrm{GR}(4^3) = \mathbb{Z}_4[x]/(x^3 + 2x^2 + x - 1)$ and Teichmüller set featuring the following eight elements: $\mathbb{T}_2 = \{0, 1, x, x^2, 1 + 3x + 2x^2, 2 + 3x + 3x^2, 3 + 3x + x^2, 1 + 2x + x^2\}$.

4.4 MUBs for non-prime-power dimensions

For $d = 6$, the lowest non-prime-power (n-p-p) case, one constructs a set of three MUBs as follows. One takes the three MUBs in $d = 2$, viz.

$$B_0^{(1)} = (|0\rangle, |1\rangle), \quad B_1^{(1)} = (|\theta_0\rangle, |\theta_1\rangle), \quad B_2^{(1)} = (|\psi_0\rangle, |\psi_1\rangle), \tag{4.13}$$

or, in matrix form, $B_0^{(1)}$=I_2, $B_1^{(1)}$=H and $B_2^{(1)}$=HS, and the first three MUBs in $d = 3$, viz.

$$B_0^{(2)} = (|0\rangle, |1\rangle, |2\rangle), \quad B_1^{(2)} = (|u_0\rangle, |u_1\rangle, |u_2\rangle), \quad B_2^{(2)} = (|v_0\rangle, |v_1\rangle, |v_2\rangle), \tag{4.14}$$

or, in a more convenient form,

$$B_0^{(2)} = I_3, \quad B_1^{(2)} = 1/\sqrt{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega_3 & \overline{\omega}_3 \\ 1 & \overline{\omega}_3 & \omega_3 \end{bmatrix}, \quad B_2^{(2)} = 1/\sqrt{3}\begin{bmatrix} 1 & \omega_3 & \omega_3 \\ 1 & \overline{\omega}_3 & 1 \\ 1 & 1 & \overline{\omega}_3 \end{bmatrix},$$

and extracts the expressions for three MUBs in $d = 6$ from the rows of the following tensorial product matrices $C_0 = B_0^{(1)} \otimes B_0^{(2)} = I_6$, $C_1 = B_1^{(1)} \otimes B_1^{(2)}$ and $C_2 = B_2^{(1)} \otimes B_2^{(2)}$. This construction can easily be generalized to any n-p-p dimension [30], [72]. One considers the prime number decomposition $d = \prod_{i=1}^{r} p_i^{e_i}$, takes its smallest factor $\tilde{m} = \min_i(p_i^{e_i})$, and gets $\tilde{m} + 1$ MUBs from the tensorial product $B^{(k)} = \otimes_{i=1}^{r} B_i^{(k)}$, $(k = 0, \ldots, \tilde{m})$.

At this point it is instructive to enlighten the above-described construction of MUBs by confining ourselves to the Galois ring in $d = 6$. Let us take the latter as the quotient $\mathrm{GR}(6^2) = \mathbb{Z}_6[x]/(x^2 + 3x + 1)$ of polynomials over $\mathbb{Z}_6$ by a polynomial irreducible over both $\mathbb{Z}_2$ and $\mathbb{Z}_3$. $\mathrm{GR}(6^2)$ has 36 elements. The notion of Teichmüller set can be generalized to the so-called Sylow decomposition [2]. Any element $\beta \in \mathrm{GR}(6)$ can be uniquely determined in the form $\beta = a \oplus b$, where a and b are in the Sylow subgroups S_a and S_b. These can be defined as $S_a = \{x \in \mathrm{GR}(6) : 2x = 0\}$ and $S_b = \{x \in \mathrm{GR}(6) : 3x = 0\}$, i.e.

$$\begin{aligned} S_a &= \{0, 3, 3x, 3 + 3x\}, \\ S_b &= \{0, 2, 4, 2x, 4x, 2 + 2x, 2 + 4x, 4 + 2x, 4 + 4x\}. \end{aligned} \tag{4.15}$$

Since the quotient polynomial is irreducible, one observes that S_a and S_b themselves are finite fields, being isomorphic to GF(4) and GF(9), respectively. One can therefore express the ring in dimension 6 as the direct product $\mathrm{GF}(4) \otimes \mathrm{GF}(9) = \mathrm{GR}(6)$. Can this property be useful to construct MUBs themselves, or does it merely represents a constraint on the maximum number of MUBs? One construction of MUBs for d=6 was based on the tensorial product of MUBs in dimension 2 and 3, respectively. But the three MUBs in dimension two do not follow from the four elements of GF(4), but from the four elements of $\mathrm{GR}(4^1) = \mathbb{Z}_4$. On the other hand, the four MUBs in d=3 follow from the three elements of $\mathrm{GF}(3) = \mathbb{Z}_3$. So the decomposition of GR(6) as a product of two fields appears to be irrelevant to the topic of MUBs. Moreover, it was shown that complete sets of MUBs in n-p-p dimensions cannot be constructed using a majority of generalizations of known formulas for finite rings [2]. This, however, should not deter us from looking at other possible constructions. For example, using the properties of sets of mutually orthogonal Latin squares, it has recently been shown that from a particular square of dimension 26^2 it is, in principle, possible to construct at least 6 MUBs, while the construction based on the prime number decomposition determines only $\min_i(p_i^{e_i}) + 1 = 2^2 + 1 = 5$ of them [68].

4.5 MUBs and finite projective planes

An intriguing similarity between mutually unbiased measurements and finite projective geometry has recently been noticed [60]. Let us find the minimum number of different measurements we need to determine uniquely the state of an ensemble of identical d-state particles. The density matrix of such an ensemble, being Hermitian and of unit trace, is specified by $(2d^2/2) - 1 = d^2 - 1$ real parameters. When one performs a non-degenerate orthogonal measurement on each of many copies of such a system one eventually obtains $d - 1$ real numbers (the probabilities of all but one of the d possible outcomes). The minimum number of different measurements needed to determine the state uniquely is thus $(d^2 - 1)/(d - 1) = d + 1$ [70], [24].

It is striking that the identical expression can be found within the context of finite projective geometry. A finite projective plane is an incidence structure consisting of points and lines such that any two points lie on just one line, any two lines pass through just one point, and there exist four points, no three of them on a line [7]. From these properties it readily follows that for any finite projective plane there exists an integer d with the properties that any line contains exactly $d + 1$ points, any point is the meet of exactly $d + 1$ lines, and the number of points is the same as the number of lines, namely $d^2 + d + 1$. This integer d is called the order of the projective plane. The most striking issue here is that the order of known finite projective planes is a power of prime. The question of which other integers occur as orders of finite projective planes remains one of the most challenging problems of contemporary mathematics. The only "no-go" theorem known so far in this respect is the Bruck–Ryser theorem [12] saying that there is no projective plane of order d if $d - 1$ or $d - 2$ is divisible by 4 and d is not the sum of two squares. For the first few non-prime-power numbers, this theorem rules out finite projective planes of order 6, 14, 21, 22, 30 and 33. Moreover, using massive computer calculations, it was proved that there is no projective plane of order ten. It is surmised that the order of any projective plane is a power of a prime.

It is conjectured [60] that the question of the existence of a set of $d + 1$ mutually unbiased bases in a d-dimensional Hilbert space if d differs from a power of a prime number is identical with the problem of whether there exist projective planes whose order d is not a power of a prime number.

4.5.1 GF(8) and the Fano plane. The smallest projective plane, also called the Fano plane, is obviously the one for $d = 2$; it contains 7 points and 7 lines, any line contains 3 points and each point is on 3 lines. It comprises a 3-dimensional vector space over the field GF(2), each point being a triple (g_1, g_2, g_3), excluding (0,0,0), where $g_i \in \mathrm{GF}(2) = \{0, 1\}$ [7]. The points of this plane can also be represented in terms of the non-zero elements of the Galois field $G = \mathrm{GF}(2^3)$.

To see this, we recall that this field is isomorphic to $\mathbb{Z}_2(x)/(\alpha)$ with the polynomial $\alpha = p(x) = x^3 + x + 1$ irreducible in GF(2). It is well known that there are three useful representations of the elements of GF(8) as shown in Table 4 [7], [27], [4].

Table 4. Representations of the elements of the Galois field GF(8).

as powers of α	as polynomials	as 3-tuples in $\mathbb{Z}_2^3$
0	0	(0,0,0)
1	1	(0,0,1)
α	α	(0,1,0)
α^2	α^2	(1,0,0)
α^3	$1+\alpha$	(0,1,1)
α^4	$\alpha+\alpha^2$	(1,1,0)
α^5	$1+\alpha+\alpha^2$	(1,1,1)
α^6	$1+\alpha^2$	(1,0,1)

The first representation emphasizes the fact that $G^* = G - \{0\}$ is a multiplicative cyclic group of order 7, for $\alpha^7 = 1$. The second representation is obtained from the first by calculating modulo the primitive polynomial α. Finally, the 3-tuple representation is obtained from the coefficients of the three powers $x^0 = 1$, $x^1 = x$ and x^2. Taking these 3-tuples as the points of a 3-dimensional vector space, we recover the Fano plane, see Figure 8.

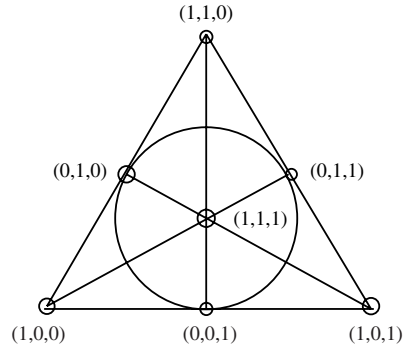


Figure 8. The Fano plane.

4.5.2 The lifted Fano plane in GR(4^3). We already know from Section 4.3 that the relevant object for 2^m-dits is not the Galois field GF(2^m), but rather the Galois ring GR(4^m). It is therefore important to have a look at the geometry in the space $A = \mathrm{GR}(4^3)$. For a ring, the concept of a vector space must be replaced by that of a module. The largest cycle in A is the set $\mathbb{T}_3^* = \mathbb{T}_3 - \{0\}$ (see Section 3), and each

element of $\mathbb{T}_3^*$ can be represented in the same way as in the case of a Galois field. This is summarized in Table 5.

Table 5. Representations of the elements of the cyclic group in the Galois ring GR(4^3).

as powers of ξ	as polynomials	as 3-tuples in $\mathbb{Z}_4^3$	as 3-tuples in $\mathbb{Z}_2^3$
0	0	(0,0,0)	(0,0,0)
1	1	(0,0,1)	(0,0,1)
ξ	ξ	(0,1,0)	(0,1,0)
ξ^2	ξ^2	(1,0,0)	(1,0,0)
ξ^3	$1+3\xi+2\xi^2$	(2,3,1)	(0,1,1)
ξ^4	$2+3\xi+3\xi^2$	(3,3,2)	(1,1,0)
ξ^5	$3+3\xi+\xi^2$	(1,3,3)	(1,1,1)
ξ^6	$1+2\xi+\xi^2$	(1,2,1)	(1,0,1)

Any polynomial $h(x)$ in $\mathbb{T}_3^*$ (column 2) is uniquely projected as a polynomial $h_2(x) = h(x) \pmod 2$ in GF(8), which results in the 3-tuple representation in $\mathbb{Z}_2^3$ (column 4). Vice versa, any polynomial in GF(8) has a unique lift in $\mathbb{T}_3^*$. Since the geometrical structure we are looking at is combinatorial and does not depend on particular coordinates, it follows that the lifted Fano plane in $\mathbb{T}_3^*$ is still the Fano plane up to isomorphism. So the Fano geometry is inherent in the geometry of qubits, but we needed a special coordinatization in order to be able to see that. For a more recent hint on this subject see [58] and [59].

4.6 MUBs of maximally entangled states

The above-discussed methods of constructing MUBs can straightforwardly be used for recognizing orthogonal bases of maximally entangled states, of which some can be mutually unbiased. Following the methodology outlined in Sections 4.2 and 4.3, let us consider a set of generalized Bell states defined as a two particle quantum Fourier transform [13], [22]

$$|\mathcal{B}_{h,k}\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} \omega_d^{kn} |n, n+h\rangle, \tag{4.16}$$

where $|n, n+h\rangle$ denotes the two-particle state $|n\rangle$, $|n+h\rangle$ and the operation $n+h$ is performed modulo d. These states are both orthonormal, $\langle \mathcal{B}_{h,k} | \mathcal{B}_{h',k'} \rangle = \delta_{hh'}\delta_{kk'}$, and maximally entangled, $\mathrm{trace}_2 \, |\mathcal{B}_{h,k}\rangle\langle\mathcal{B}_{h,k}| = \frac{1}{d} I_d$, where trace_2 means the partial trace over the second qudit [37]. If one restricts to the case of 2-qubits, one recovers

the well-known representation of Bell states

$$(|\mathcal{B}_{0,0}\rangle, |\mathcal{B}_{0,1}\rangle) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle, |00\rangle - |11\rangle),$$
$$(|\mathcal{B}_{1,0}\rangle, |\mathcal{B}_{1,1}\rangle) = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle, |01\rangle - |10\rangle),$$

where a more compact notation $|00\rangle = |0, 0\rangle$, $|01\rangle = |0, 1\rangle$, ... is employed. Let us first focus on 2-qubits starting from eq. (4.8). Paralleling of what we did in Section 4.3, one first identifies kn in (4.16) as the multiplication $k \odot n$ of polynomials in GR(4) and then makes use of Teichmüller decomposition $k = a \oplus 2 \odot b$. This leads to a set of 4 bases ($h, a = 0, 1$) of two vectors ($b = 0, 1$), namely

$$|\mathcal{B}^a_{h,b}\rangle = \frac{1}{\sqrt{2}} \sum_{n=0}^{1} i^{(a\oplus 2\odot b)\odot n} |n, n \oplus h\rangle. \tag{4.17}$$

Casting the last equation into its matrix form (safe for the proportionality factor),

$$\begin{bmatrix} (|00\rangle + |11\rangle, |00\rangle - |11\rangle); & (|01\rangle + |10\rangle, |01\rangle - |10\rangle) \\ (|00\rangle + i|11\rangle, |00\rangle - i|11\rangle); & (|01\rangle + i|10\rangle, |01\rangle) - i|10\rangle) \end{bmatrix}, \tag{4.18}$$

one finds that two bases in one column are mutually unbiased, while vectors in two bases on the same line are orthogonal to each other.

Eq. (4.17) can easily be extended to maximally entangled two-particle sets of 2^m-dits by applying, as in eq. (4.11), the Frobenius map (4.9) to the base field $\mathbb{Z}_4$

$$|\mathcal{B}^a_{h,b}\rangle = \frac{1}{\sqrt{2^m}} \sum_{n=0}^{2^m-1} i^{\mathrm{tr}[(a\oplus 2\odot b)\odot n]} |n, n \oplus h\rangle. \tag{4.19}$$

For 2-particle sets of quartits, using eqs. (4.12) and (4.19), one thus gets 4 sets $(|\mathcal{B}^a_{h,b}\rangle,\ h = 0, \ldots, 3)$ of 4 MUBs ($a = 0, \ldots, 3$):

$$\{(|00\rangle + |11\rangle + |22\rangle + |33\rangle, |00\rangle + |11\rangle - |22\rangle - |33\rangle,$$
$$|00\rangle - |11\rangle - |22\rangle + |33\rangle, |00\rangle - |11\rangle + |22\rangle - |33\rangle);$$
$$(|00\rangle - |11\rangle - i|22\rangle - i|33\rangle, |00\rangle - |11\rangle + i|22\rangle + i|33\rangle,$$
$$|00\rangle + |11\rangle + i|22\rangle - i|33\rangle, |00\rangle + |11\rangle - i|22\rangle + i|33\rangle);\ \ldots\}$$
$$\{(|01\rangle + |12\rangle + |23\rangle + |30\rangle, |01\rangle + |12\rangle - |23\rangle - |30\rangle,$$

$$
\begin{aligned}
&|01\rangle - |12\rangle - |23\rangle + |30\rangle, |01\rangle - |12\rangle + |23\rangle - |30\rangle);\\
&(|01\rangle - |12\rangle - i|23\rangle - i|30\rangle, |01\rangle - |12\rangle + i|23\rangle + i|30\rangle,\\
&|01\rangle + |12\rangle + i|23\rangle - i|30\rangle, |01\rangle + |12\rangle - i|23\rangle + i|30\rangle);\ \ldots\}\\
&\{(|02\rangle + |13\rangle + |20\rangle + |31\rangle, |02\rangle + |13\rangle - |20\rangle - |31\rangle,\\
&|02\rangle - |13\rangle - |20\rangle + |31\rangle, |02\rangle - |13\rangle + |20\rangle - |31\rangle);\ \ldots\}\\
&\{(|03\rangle + |10\rangle + |21\rangle + |32\rangle, |03\rangle + |10\rangle - |21\rangle - |32\rangle,\\
&|03\rangle - |10\rangle - |21\rangle + |32\rangle, |03\rangle - |10\rangle + |21\rangle - |32\rangle);\ \ldots\},
\end{aligned}
\tag{4.20}
$$

where, for the sake of brevity, we omitted the normalization factor $(1/2)$. Within each set, the four bases are mutually unbiased, as in (4.12), while the vectors of the bases from different sets are orthogonal.

Turning now to odd characteristic, i.e. to $d = p^m$ with p an odd prime, we can similarly extend Wootters formula (4.5) to the generalized Bell states

$$
|\mathcal{B}^a_{h,b}\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} \omega_d^{\mathrm{tr}[(a\odot n\oplus b)\odot n]} |n, n\oplus h\rangle, \tag{4.21}
$$

where the trace is defined by eq. (4.6). A list of the generalized Bell states of qutrits for the base $a = 0$ can be found in [23], the work that relies on a coherent state formulation of entanglement. In general, for d a power of a prime, starting from (4.16) or (4.21) one obtains d^2 bases of d maximally entangled states. Each set of the d bases (with h fixed) has the property of mutual unbiasedness.

Eq. (4.16) can be used, without any substantial restriction, to find d bases ($h = 0, \ldots, d-1$) of maximally entangled states in any composite dimension $d = \prod_{i=1}^r p_i^{e_i}$. Or one can also follow the strategy of Section 4.4 to get $\tilde{m} = \min_i(p_i^{e_i})$ sets of mutually unbiased bases of maximally entangled states. In $d = 6$, for example, one expects that two such sets of d bases can be constructed. Using the tensorial products in Section 4.4, one indeed finds the two 2×6 sets (with the $1/\sqrt{6}$ factor omitted)

$$
\begin{aligned}
&\{(|00\rangle + |11\rangle + |22\rangle + |33\rangle + |44\rangle + |55\rangle,\\
&|00\rangle + \omega_3|11\rangle + \overline{\omega}_3|22\rangle + |33\rangle + \omega_3|44\rangle + \overline{\omega}_3|55\rangle,\\
&|00\rangle + \overline{\omega}_3|11\rangle + \omega_3|22\rangle + |33\rangle + \overline{\omega}_3|44\rangle + \omega_3|55\rangle,\\
&|00\rangle + |11\rangle + |22\rangle - |33\rangle - |44\rangle - |55\rangle,\\
&|00\rangle + \omega_3|11\rangle + \overline{\omega}_3|22\rangle - |33\rangle - \omega_3|44\rangle - \overline{\omega}_3|55\rangle,\\
&|00\rangle + \overline{\omega}_3|11\rangle + \omega_3|22\rangle - |33\rangle - \overline{\omega}_3|44\rangle - \omega_3|55\rangle);\\
&(|00\rangle + \omega_3|11\rangle + \omega_3|22\rangle + i|33\rangle + i\omega_3|44\rangle + i\omega_3|55\rangle,\\
&|00\rangle + \overline{\omega}_3|11\rangle + |22\rangle + i|33\rangle + i\overline{\omega}_3|44\rangle + i|55\rangle,\\
&|00\rangle + |11\rangle + \overline{\omega}_3|22\rangle + i|33\rangle + i|44\rangle + i\overline{\omega}_3|55\rangle,\\
&|00\rangle + \omega_3|11\rangle + \omega_3|22\rangle - i|33\rangle - i\omega_3|44\rangle - i\omega_3|55\rangle,
\end{aligned}
$$

$$|00\rangle + \overline{\omega}_3|11\rangle + |22\rangle - i|33\rangle - i\overline{\omega}_3|44\rangle - i|55\rangle,$$
$$|00\rangle + |11\rangle + \overline{\omega}_3|22\rangle - i|33\rangle - i|44\rangle - i\overline{\omega}_3|55\rangle);\ \ldots\}$$
$$\vdots$$
$$\{(|01\rangle + |12\rangle + |23\rangle + |34\rangle + |45\rangle + |50\rangle,$$
$$|01\rangle + \omega_3|12\rangle + \overline{\omega}_3|23\rangle + |34\rangle + \omega_3|45\rangle + \overline{\omega}_3|50\rangle,\ \ldots\}.$$

Multipartite entanglement is a key ingredient of many quantum protocols, still needing much work to be properly understood. Sets of orthogonal product states that are unextendible, meaning that no further product states can be found orthogonal to all the existing ones, have recently attracted a lot of attention. These unextendible product bases [16] and their complement [5] certainly deserve reconsideration in terms of the above-outlined theory, which is based on abstract algebra and finite geometry.

The Fourier transform approach implies that mutual unbiasedness and maximal entanglement are complementary aspects in orthogonal quantum measurements. In such measurements the quantum states are encoded in a three-dimensional lattice of indices h (entanglement), a (unbiasedness) and b (dimensionality of Hilbert space). If d is a power of a prime, the lattice is a cube since in this case h, a and b reach their limiting value d. If one forgets about entanglement ($h = 0$), the finite geometry which seems to be of most relevance is that of a finite projective plane. On the other hand, when unbiasedness is not taken into account, as well as for multipartite information tasks when d is not (a power of) a prime, other concepts have been introduced, such as Bell inequalities [71], coherent states [23], entanglement swapping[9], generalized Hopf fibrations [6], topological entanglement [29] and bound entanglement [16], to mention a few.

Lifts of the base ring $\mathbb{Z}_4$ to Galois rings $\mathrm{GR}(4^m)$ of characteristic four have been used in this section to derive the MUBs of m-qubits. Similarly one can use lifts of the base field $\mathbb{Z}_3$ to near-fields $\mathrm{NF}(3^m)$, with a view of revealing a new class of quantum states for the corresponding dimensions. In the simplest case of miniquaternions ($m = 2$) [57] the relevant automorphism (permutation) group is used to get the trace map from $\mathrm{NF}(3^2)$ to $\mathbb{Z}_3$ and to find the corresponding additive characters. This may be related to the unextendible product bases [16]. As nearfields (and nearrings) are of great importance for coordinatizing (non-Desarguesian) projective planes [58] and the latter play a crucial role in the classical coding theory, we expect that these structures will also play a similar role in quantum coding and computing.

References

[1] R. Adler, A study of locking phenomena in oscillators, *Proc. IRE* **34** (1946), 351–357; reprinted in *Proc. IEE* 61, 1380-1385 (1973).

[2] C. Archer, There is no generalization of known formulas for mutually unbiased bases, *J. Math. Phys.* 46 (2005), 022106.

[3] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury and F. Vatan, A new proof of the existence of mutually unbiased bases. Quantum computation and quantum cryptography, *Algorithmica* 34 (2002), 512–528.

[4] L. M. Batten, *Combinatorics of finite geometries*, Cambridge University Press, Cambridge 1997.

[5] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin and B. M. Terhal, Unextendible product bases and bound entanglement, *Phys. Rev. Lett.* 82 (1999), 5385–5388.

[6] B. A. Bernevig and H.-D. Chen, Geometry of the three-qubit state, entanglement and division algebras, *J. Phys. A* 36 (2003), 8325–8339.

[7] A. Beutelspacher and U. Rosenbaum, *Projective geometry: from foundations to applications*, Cambridge University Press, Cambridge 1998.

[8] E. Bogomolny, Quantum and Arithmetical Chaos, Lectures given at Les Houches School: Frontiers in Number Theory, Physics and Geometry, March 2003; `nlin.CD/0312061`.

[9] S. Bose, V. Vedral and P. L. Knight, A multiparticle generalization of entanglement swapping, *Phys. Rev. A* 57 (1998), 822–829.

[10] J.-B. Bost and A. Connes, Hecke algebras, type III factors and phase transitions with spontaneous symmetry breaking in number theory, *Selecta Math.* 1 (1995), 411–457.

[11] V. Bouchiat, D. Vion, P. Joyez, D. Estève and M. H. Devoret, Quantum coherence with a single Cooper pair, *Physica Scripta* T76 (1998), 165–170.

[12] R. H. Bruck and H. J. Ryser, The non-existence of certain finite projective planes, *Canad. J. Math.* 1 (1949), 88–93.

[13] N. J. Cerf, M. Bourennane, A. Karlsson and N. Gisin, Security of quantum key distribution using d-level systems, *Phys. Rev. Lett.* 88 (2002), 127902.

[14] P. B. Cohen, Quantum statistical mechanics and number theory, in *Algebraic geometry: Hirzebruch 70* (Warsaw, 1998), Contemp. Math. 241, Amer. Math. Soc., Providence, RI, 1999, 121–128.

[15] P. Cvitanovic, Circle maps: irrationally winding, in *From number theory to physics* (Les Houches, 1989), Springer-Verlag, Berlin 1992, 631–658.

[16] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin and B. M. Terhal, Unextendible product bases, uncompletable product bases and bound entanglement, *Comm. Math. Phys.* 238 (2003), 379–410.

[17] V. V. Dodonov and V. I. Man'ko, *Theory of nonclassical states of light*, Taylor & Francis, London 2003.

[18] S. Dos Santos and M. Planat, $1/f$ spectrum in a nonlinear electronic loop: consequences in frequency standards and signal processing, *Ann. Télécommun.* 53 (1998), 488–494.

[19] T. Durt, If $1 = 2+3$, then $1 = 2.3$: Bell states, finite groups, and mutually unbiased bases, a unifying approach, `arXiv:quant-ph/0401046`.

[20] T. Durt, D. Kaszlikowski, J. L. Chen, and L. C. Kwek, Security of quantum key distributions with entangled qudits, *Phys. Rev. A* 69 (2004), 032313.

[21] A. Ekert, P. Hayden and H. Inamori, Basic concepts in quantum computation, in *Coherent atomic matter waves*, Les Houches Summer School 1999, Springer-Verlag, Berlin 2001, 659–702.

[22] D. I. Fivel, Remarkable Phase oscillations appearing in the lattice dynamics of Einstein-Podolsky-Rosen states, *Phys. Rev. Lett.* 74 (1995), 835–838.

[23] K. Fujii, A relation between coherent states and generalized bell states, `arXiv:quant-ph/0105077`.

[24] K. S. Gibbons, M. J. Hoffman and W. K. Wootters, Discrete phase space based on finite fields, *Phys. Rev. A* 70 (2004), 062101.

[25] M. Grassl, On SIC-POVMs and MUBs in Dimension 6, `arXiv:quant-ph/0406175`.

[26] A. R. Hammons, Jr. , P. V. Kumar, A. R. Calderbank, N. J. A. Sloane and P. Solé, The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes, *IEEE Trans. Inform. Theory* 40 (1994), 301–319.

[27] J. W. P. Hirschfeld, *Projective geometries over finite fields*, Oxford University Press, Oxford 1998.

[28] H. Iwaniec, *Spectral Methods of automorphic functions*, Second edition, Amer. Math. Soc., Providence, RI, 2002.

[29] L. H. Kauffman and S. J. Lomonaco, Quantum entanglement and topological entanglement, *New J. Phys.* 4 (2002), 73.1–73.18.

[30] A. Klappenecker and M. Rötteler, Constructions of mutually unbiased bases, in *Finite fields and applications*, Lecture Notes in Comput. Sci. 2948, Springer, Berlin, 2004, 137–144.

[31] J. Klapper and J. T. Frankle, *Phase-locked and Frequency-Feedback Systems*, Academic Press, New York 1972.

[32] A. Knauf, Number theory, dynamical systems and statistical mechanics, *Rev. Math. Phys.* 11 (1999), 1027–1060.

[33] R. Lidl and G. Pilz, *Applied Abstract Algebra*, Second edition, Springer-Verlag, New York 1998.

[34] R. Lynch, The quantum phase: a critical review, *Physics Reports*, 256 (1995), 367–436.

[35] J. M. Mérolla, Y. Mazurenko, J. P. Goedgebuer and W. T. Rhodes, Single-photon interference in sidebands of phase-modulated light for quantum cryptography, *Phys. Rev. Lett.* 82 (1999), 1656–1659.

[36] P. Moree and H. Hommersom, Value distribution of Ramanujan sums and of cyclotomic polynomial coefficients; `arXiv:math.NT/0307352`.

[37] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge 2000.

[38] M. Novaes and J. P. Gazeau, Multidimensional generalized coherent states, *J. Phys. A* 36 (2003), 199–212.

[39] K. R. Parthasarathy, On estimating the state of a finite level quantum system, `arXiv:quant-ph/0408069`.

[40] D. T. Pegg and S. M. Barnett, Phase properties of the quantized single-mode electromagnetic field, *Phys. Rev. A* 39 (1989), 1665–1675.

[41] A. Perelomov, *Generalized coherent states and their applications*, Springer-Verlag, Berlin 1986.

[42] A. O. Pittenger and M. H. Rubin, Mutually unbiased bases, generalized spin matrices and separability, *Linear Algebra Appl.* 390 (2004), 255–278.

[43] M. Planat, $1/f$ noise, the measurement of time and number theory, *Fluct. and Noise Lett.* 1 (2001), R65–R79.

[44] M. Planat, The hyperbolic geometry of $1/f$ noise in phase-locked loops, in Noise and Fluctuations, ed. J. Sikula, Czech Noise Research Laboratory, Brno 2003, 533–539.

[45] M. Planat, Invitation to the "spooky" quantum phase-locking effect and its link to $1/f$ fluctuations; `arXiv:quant-ph/0310082`.

[46] M. Planat, On the cyclotomic quantum algebra of time perception, *NeuroQuantology* 2 (2004), 292–308.

[47] M. Planat and C. Eckert, On the frequency and amplitude spectrum and the fluctuations at the output of a communication receiver, *IEEE Trans. Ultrason. Ferr. Freq. Contr.*, 47 (2000), 1173–1182.

[48] M. Planat and E. Henry, Arithmetic of $1/f$ noise in a phase-locked loop, *Appl. Phys. Lett.* 80 (2002), 2413–2415.

[49] M. Planat and H. Rosu, Cyclotomy and Ramanujan sums in quantum phase-locking, *Phys. Lett. A* 315 (2003), 1–5.

[50] M. Planat and H. Rosu, The hyperbolic, the arithmetic and the quantum phase, *J. Opt. B Quantum Semiclass. Opt.* 6 (2004), S583–S590.

[51] M. Planat and H. Rosu, Mutually unbiased phase states, phase uncertainties and Gauss sums, *Eur. Phys. J. D* 36 (2005), 133–139.

[52] M. Planat, H. Rosu and S. Perrine, Ramanujan sums for signal processing of low frequency noise, *Phys. Rev. E* 66 (2002), 56128–56134.

[53] Quiprocone website; `http://www.imaph.tu-bs.de/qi/problems`.

[54] H. Rademacher, *Topics in analytic number theory*, Springer Verlag, New-York 1973.

[55] J. M. Renes, R. Blume-Kohout, A. J. Scott and C. M. Caves, Symmetric informationally complete quantum measurements, *J. Math. Phys.* 45 (2004), 2171–2180.

[56] C. Rigetti, R. Mosseri and M. Devoret, Geometric approach to digital quantum information, *Quantum Inf. Process.* 3 (2004), 351–389.

[57] T. G. Room and P. B. Kirkpatrick, *Miniquaternion geometry, an introduction to the study of projective planes*, Cambridge University Press , Cambridge 1971.

[58] M. Saniga and M. Planat, Sets of mutually unbiased bases as arcs in finite projective planes?, *Chaos, Solitons, Fractals* 26 (2005), 1267–1270.

[59] M. Saniga and M. Planat, Hjelmslev geometry of mutually unbiased bases, *J. Phys. A* 39 (2006), 435–440.

[60] M. Saniga, M. Planat and H. Rosu, Mutually unbiased bases and finite projective planes, *J. Opt. B Quantum Semiclass. Opt.* 6 (2004), L19–L20.

[61] M. S. Sargent, M. O. Scully and W. E. Lamb, *Laser physics*, Addison-Wesley, London 1974.

[62] W. P. Schleich and S. M. Barnett (eds.), Quantum phase and phase dependent measurements, *Physica Scripta* T48 (1993), 144 pages.

[63] M. R. Schroeder, *Number theory in science and communication*, Springer-Verlag, Berlin 1999, 16–24.

[64] L. Susskind and J. Glogower, Quantum mechanical phase and time operator, *Physics* 1 (1964), 49–61.

[65] A. Vourdas, Phase states: An analytic approach in the unit disk, *Physica Scripta* T48 (1993), 84–86.

[66] A. Vourdas, Quantum systems with finite Hilbert space, *Rep. Progr. Phys.* 67 (2004), 267–320.

[67] Z. X. Wan, *Quaternary codes*, World Scientific, Singapore 1997.

[68] P. Wocjan and T. Beth, New construction of mutually unbiased bases in square dimensions; `arXiv:quant-ph/0407081`.

[69] W. K. Wootters, Quantum measurements and finite geometry, *Found. Physics* 36 (2006), 112–126.

[70] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, *Ann. Physics* 191 (1989), 363–381.

[71] S. Yu, Z. B. Chen , J. W. Pan and Y. D. Zhang, Classifying N-qubit entanglement via Bell's inequalities, *Phys. Rev. Lett.* 90 (2003), 080401.

[72] G. Zauner, *Quantendesigns-Grundzüge einer nichtkommutativen Designtheorie*, Dissertation, Universität Wien, Wien 1999.

On self-similar finitely generated uniformly discrete (SFU-)sets and sphere packings

Jean-Louis Verger-Gaugry

Institut Fourier - CNRS UMR 5582, Université Grenoble I,
BP 74 - Domaine Universitaire,
38402 - Saint Martin d'Hères, France
email: jlverger@ujf-grenoble.fr

Abstract. The first part of this paper is a survey on links between Geometry of Numbers and aperiodic crystals in Physics, viewed from the mathematical side. In a second part, we prove the existence of a canonical cut-and-project scheme above a (SFU-set) self-similar finitely generated packing of (equal) spheres Λ in $\mathbb{R}^n$ and investigate its consequences, in particular the role played by the Euclidean and inhomogeneous minima of the algebraic number field generated by the self-similarity on the Delone constant of the sphere packing. We discuss the isolation phenomenon. The degree d of this field divides the $\mathbb{Z}$-rank of $\mathbb{Z}[\Lambda - \Lambda]$. We give a lower bound of the Delone constant of a k-thin SFU-set (sphere packing) which arises from a model set or a Meyer set when d is large enough.

Résumé. Cet article donne d'abord un panorama rapide des liens existant entre Géométrie des Nombres et Cristaux Apériodiques en Physique, vus du côté mathématique. Dans une deuxième partie, nous prouvons l'existence d'un schéma de coupe-et-projection au dessus d'un empilement de sphères (égales) de type fini et autosimilaire (ensemble AFU) Λ dans $\mathbb{R}^n$ et regardons ses conséquences, en particulier le rôle joué par les minima inhomogènes et Euclidiens du corps de nombres engendré par l'autosimilarité sur la constante de Delone de l'empilement de sphères. Nous discutons le phénomène d'isolement. Le degré d du corps de nombres divise le $\mathbb{Z}$-rang de $\mathbb{Z}[\Lambda - \Lambda]$. Nous donnons une borne inférieure de la constante de Delone d'un ensemble (empilement de sphères) AFU k-mince qui provient d'un ensemble modèle ou d'un ensemble de Meyer lorsque d est suffisamment grand.

Contents

1 Introduction

The mathematics of uniformly discrete point sets and Delone sets developed recently has at least four different origins: (i) the experimental evidence of nonperiodic states of matter in condensed matter physics, so-called aperiodic crystals, like quasicrystals [4], [55], [62], [68], [103] incommensurate modulated crystals phases [67], [69] and their geometric modelization (cf. Appendix), (ii) works of Delone [36], [37], [42], [97] on geometric crystallography (comparatively, see [58], [83], [90], [101] for a classical mathematical approach of periodic crystals), (iii) works of Meyer on now called cut-and-project sets and Meyer sets [80], [81], [82], [92] (for a modern language of Meyer sets in locally compact abelian groups: [84]), (iv) the theory of self-similar tilings [10], [75], [109] and the use of ergodic theory to understand diffractivity [5], [98], [109]. In particular, the impact on mathematics of the discovery of quasicrystals in 1984 [103], as long-range ordered phases, was outlined by Lagarias [71]. The term *mathematical quasicrystals* [6], [72] was proposed to name these Delone sets which are used as discrete geometrical models of these new states of matter which have particular spectral or diffraction properties; in particular *crystals* are those states for which the spectrum is essentially pure point (see [66], [102] and the Appendix for the new definition of what is a *crystal*, and [34], [57], [63], [64] for spectral/diffraction theory). Delone sets are conceived as natural generalizations of lattices in modern crystallography.

In this note we will briefly review these notions (Section 2) and will consider more generally uniformly discrete sets of $\mathbb{R}^n$, in particular (SFU-) self-similar finitely generated uniformly discrete sets (Definition 2.19). A uniformly discrete set of $\mathbb{R}^n$ of constant $r > 0$ is a packing of (equal) spheres of $\mathbb{R}^n$ of (common) radius $r/2$. There are several advantages to consider uniformly discrete sets instead of Delone sets only: their $\mathbb{R}$-spans may take arbitrary dimensions between 0 and n, while that of a Delone set is only n; they can be finite sets which is forbidden for Delone sets; they may exhibit (spherical) holes of arbitrary size at infinity whereas the size of holes in Delone sets is limited by the Delone constant. For instance, see [7] for a nonclassical example. A classification of uniformly discrete sets, hence of Delone sets, which extends that given in [70], is proposed in Subsection 2.3. Finitely generated uniformly discrete sets of $\mathbb{R}^n$ constitute the largest class on which an address map (Subsection 2.3) can be defined.

The theory of SFU-sets generalizes that of lattice packings of (equal) spheres of $\mathbb{R}^n$ [22], [25], [29], [56], [77], [113] since a lattice is already an SFU-set itself (integers are self-similarities: if $m \in \mathbb{Z}$ and L is a lattice, $mL \subset L$), where lattices are $\mathcal{O}_F$-lattices or not for F an algebraic number field with involution [15], [33], and makes use of algebraic integers of certain types (Subsection 2.4). Self-similar Meyer sets only admit self-similarities which are Pisot or Salem numbers [80], while self-similar finitely generated Delone sets only provide Perron or Lind numbers as self-similarities [70]. It is an open problem to find a criterium which ensures that a given uniformly discrete set admits at least one self-similarity. For a general Delone set symmetries

and in particular inflation symmetries are expected to be rare, especially when the dimension of the ambient space is large. For lattices, Bannai [11] has shown the existence of many unimodular $\mathbb{Z}$-lattices with trivial (point) automorphism group in a given genus of positive definite unimodular $\mathbb{Z}$-lattices of sufficiently large rank (see also [28]).

The existence of cut-and-project schemes above Delone sets is useful to characterize the set of its self-similarities, inflation centers, local clustering, etc. [30], [31], [52], [78]. Given an arbitrary uniformly discrete set it is an open problem whether a cut-and-project scheme lies above it (Subsection 2.1.2). Theorem 1.1 answers in full generality this problem for a given SFU-set $\Lambda \subset \mathbb{R}^n$ (with $\Lambda \in \mathcal{UD}_{fg}$, see Subsection 2.3) with self-similarity λ. The constructions use the Archimedean embeddings of the number field $K := \mathbb{Q}(\lambda)$ generated by the self-similarity λ (Section 3) in a vectorial way, as a product of copies of the étale $\mathbb{R}$-algebra $K_{\mathbb{R}} := K \otimes_{\mathbb{Q}} \mathbb{R}$. Denote by $\Sigma\colon K \to K_{\mathbb{R}}$ the canonical map. The structure of the lattice in the cut-and-project scheme above Λ arises as a consequence of the Jordan invariants of $\mathbb{R}^{\mathrm{rk}\,\Lambda}$ as a $K[X]$-module (from (ii) in Theorem 1.1).

Theorem 1.1. *Let $\Lambda \subset \mathbb{R}^n$, $n \geq 1$, be a uniformly discrete set such that $m := \mathrm{rk}\,\mathbb{Z}[\Lambda - \Lambda] < +\infty$ with $m \geq 1$. Let $\lambda > 1$ be a (affine) self-similarity of Λ, i.e. a real number > 1 such that $\lambda(\Lambda - c) \subset \Lambda - c$ for a certain $c \in \mathbb{R}^n$. Then the following holds:*

(i) *λ is a real algebraic integer of degree $d \geq 1$ and d divides m.*

(ii) *There exist $r = m/d$ $\mathbb{Q}$-linearly independent vectors $w_1, w_2, \ldots, w_r$ in the $\mathbb{Q}(\lambda)$-vector space $\mathbb{Q}[\Lambda - \Lambda]$ such that $\mathbb{Z}[\Lambda - \Lambda]$ is a rank m $\mathbb{Z}$-submodule of the $\mathbb{Z}$-module:*

$$\mathbb{Z}[w_1, \lambda w_1, \ldots, \lambda^{d-1} w_1, w_2, \lambda w_2, \ldots, \lambda^{d-1} w_2, \ldots, w_r, \lambda w_r, \ldots, \lambda^{d-1} w_r].$$

(iii) *For every $\mathbb{Z}$-basis $\{v_1, v_2, \ldots, v_m\}$ of $\mathbb{Z}[\Lambda - \Lambda]$, a matrix relation $\lambda V = MV$ holds, where $V = {}^t[v_1, \ldots, v_m]$ and M is an invertible integral $m \times m$ matrix with characteristic polynomial $\det(XI - M) = (\varphi(X))^{m/d}$ in which $\varphi(X)$ is the minimal polynomial of λ; in particular, $\det M = N_{K/\mathbb{Q}}(\lambda)^{m/d}$, where $N_{K/\mathbb{Q}}(\lambda)$ is the algebraic norm of λ.*

(iv) *There exists a cut-and-project scheme above Λ,*

$$\Big(\coprod_{i=1}^{r} K_{\mathbb{R}} \frac{w_i}{\|w_i\|} \simeq H \times \mathbb{R}[\Lambda], L, \pi, \mathrm{pr}_1\Big)$$

where the lattice $L = \coprod_{i=1}^{r} \Sigma(\mathbb{Z}[\lambda]) \frac{w_i}{\|w_i\|}$ is such that $\mathrm{pr}_1(L) \supset \mathbb{Z}[\Lambda - \Lambda]$, whose internal space H is the product of two spaces

$$H = (R_K \setminus \mathbb{R}[\Lambda]) \times \overline{G},$$

where R_K is the image of $\mathbb{R}[\Lambda]$ in $\coprod_{i=1}^{r} K_{\mathbb{R}} \frac{w_i}{\|w_i\|}$ by the real and imaginary embeddings of K, and $\overline{G}$ the closure in $\coprod_{i=1}^{r} K_{\mathbb{R}} \frac{w_i}{\|w_i\|}$ of the image by Σ of

the space of relations over K between the generators $w_1, \ldots, w_r$. The space $R_K \setminus \mathbb{R}[\Lambda]$ is called the shadow space of Λ. This cut-and-project scheme is endowed with an Euclidean structure given by a real trace-like symmetric bilinear form for which R_K and $\overline{G}$ are orthogonal.

The *central cluster* of the basis $(\lambda^j w_i)_{i=1,\ldots,r,\ j=0,\ldots,d-1}$ is by definition the set $\{w_1, w_2, \ldots, w_r\}$. Note that some vectors in a central cluster may be $\mathbb{R}$-linearly dependent. When $w_1, w_2, \ldots, w_r$ have identical norms and constitute orbits (i.e. F-clusters) under the action of a finite group, say F, constructions in (iv) in Theorem 1.1 can be deduced from [32]. It is easy to check that

$$r = 1 \text{ in Theorem 1.1} \Longrightarrow \text{the } \mathbb{R}\text{-span of } \Lambda \text{ is one-dimensional.}$$

The converse is not always true: in Subsection 2.5 we give the example of the sets $\mathbb{Z}_\beta$ of beta-integers [12] on the line for which open problems exist.

Conversely, what is the set of all self-similarities obtained by Theorem 1.1 (i)? The answer is simple. If $\beta > 1$ is an algebraic integer there exists at least one SFU-set S admitting β as self-similarity,

$$S = \{\ldots, -\beta^2, -\beta, -1, 0, +1, \beta, \beta^2, \ldots\},$$

on the line, which is of finite type and satisfies $\beta S \subset S$ by construction.

Corollary 1.2. *If Λ is a self-similar finitely generated sphere packing in $\mathbb{R}^n$, with self-similarity λ, such that $r = 1$, i.e. for which the degree d of λ equals the rank m of $\mathbb{Z}[\Lambda - \Lambda]$, then $\mathbb{Z}[\Lambda - \Lambda]$ is the projection of a sublattice of finite index of an ideal lattice of $K = \mathbb{Q}(\lambda)$ on the cut-and-project scheme above Λ, of index an integer multiple of $(\mathcal{O}_K : \mathbb{Z}[\lambda])$.*

Theorem 1.1 gives a framework for constructing aperiodic (equal) sphere packings $\mathcal{B}(\Lambda)$ for which local arrangements, for instance like t-designs [9], can be computed from a lattice in higher dimension above Λ. In Corollary 1.2 the notion "Arakelov divisor", meaning that the embedding of $\mathbb{Z}[\Lambda - \Lambda]$ into the cut-and-project scheme is given with an Euclidean structure, could be substituted for "ideal lattice" by the one-to-one correspondence given in [99] (see also Neukirch [88]).

Dense sphere packings of $\mathbb{R}^n$ are of general interest [20], [25], [29], [54], [56], [86], [113]. For a sphere packing whose set of centers is a Delone set Λ, which is a uniformly discrete set of constant $r > 0$, of Delone constant $R(\Lambda) := \sup_{z\in\mathbb{R}^n} \inf_{\mu\in\Lambda} \|z - \mu\|$, the density $\delta(\mathcal{B}(\Lambda))$ of $\mathcal{B}(\Lambda)$ satisfies $\delta(\mathcal{B}(\Lambda)) \geq (2R(\Lambda)/r)^{-n}$ [86]. If Λ is only a uniformly discrete set, no equivalent formula for bounding from below the density $\delta(\mathcal{B}(\Lambda))$ exists in general. However, this is not the case for uniformly discrete sets of $\mathbb{R}^n$, called *pseudo-Delone sets*, which behave in some sense like Delone sets (Subsection 5.1).

Theorem 1.3. *Let Λ be a uniformly discrete set of $\mathbb{R}^n$ of constant $r > 0$, which is pseudo-Delone of pseudo-Delone constant $R(\Lambda)$. Then the density of the sphere*

packing $\mathcal{B}(\Lambda)$ (*of common radius* $r/2$) *whose set of centers is* Λ *satisfies*

$$\delta(\mathcal{B}(\Lambda)) \geq \left(\frac{2R(\Lambda)}{r}\right)^{-n}. \tag{1.1}$$

Theorem 1.3 shows that it is important to obtain interesting lower bounds of the Delone constant (or pseudo-Delone constant) $R(\Lambda)$ to control dense sphere packings, in particular, sphere packings whose set of centers is an SFU-set.

In Section 6 we comment on the two origins of the (pseudo-) Delone constant of a sphere packing whose set of centers is an SFU-set: the first one is related to the geometrical properties of the central cluster $\{w_1, w_2, \ldots, w_r\}$ as given by Theorem 1.1 (ii), the second one is of purely arithmetical nature; it comes from the Euclidean and inhomogeneous minima associated with a sublattice of a product of ideal lattices ([14], [26], [27]), in bijection with $\mathbb{Z}[\Lambda - \Lambda]$, in the cut-and-project scheme given by Theorem 1.1 (iv). Only the case $r = 1$ is reported in Section 6.

Theorem 1.4. *Let* $\Lambda \subset \mathbb{R}^n$, $n \geq 1$, *be an SFU-set which is either a model set or a Meyer set in the cut-and-project scheme defined by Theorem* 1.1 (iv) *with* $r = 1$, Ω *as window and lattice* L' *such that* $\mathrm{pr}_1(L') = \mathbb{Z}[\Lambda - \Lambda]$.

Assume that the self-similarity λ *is of degree* $d \geq 3$, *that* $K = \mathbb{Q}(\lambda)$ *has unit rank* > 1 *and is not a CM-field. Then, if* Λ *is* k*-thin,* $k \geq 2$, *its Delone constant* $R(\Lambda)$ *satisfies*

$$R(\Lambda) \geq \sqrt{d}\left(M(K)^{2/d} - M_k(K)^{2/d}\right)^{\frac{1}{2}} > 0, \tag{1.2}$$

where $M(K)$ *resp.* $M_k(K)$ *is the Euclidean minimum resp. the* k*-th Euclidean minimum of* K.

The space-filling condition $m/d = r \geq n$ for couples of values $\{(d, m)\}$ (with the notations of Theorem 1.1) is necessary to construct dense sphere packings of $\mathbb{R}^n$.

The $75^{ièmes}$ *Rencontres between Mathematicians and Physicists* held at IRMA, Strasbourg, on the subject "Number Theory and Physics" have offered to the author the opportunity of writing this brief note, initially conceived as a short survey, on the relationships between sphere packings, the mathematics of aperiodic crystals, algebraic number theory and numeration in base β, where β is an algebraic integer > 1.

2 Uniformly discrete sets and Delone sets

2.1 Definitions and topology

Let us define *uniformly discrete sets* and *Delone sets* in two different contexts: in the metric case when the ambient space is a metric space which is σ-compact and locally compact, like $\mathbb{R}^n$, and when the ambient space is $\mathbb{R}^n$ with a cut-and-project scheme that lies above it with a locally compact abelian group as internal space.

2.1.1 Metric case. Let (H, δ) be a σ-compact and locally compact metric space with infinite diameter (for δ). A discrete subset Λ of H is said to be *uniformly discrete* if there exists a real number $r > 0$ such that

$$x, y \in \Lambda,\ x \neq y \text{ implies } \delta(x, y) \geq r.$$

A uniformly discrete set is either the empty set, or a subset $\{x\}$ of H reduced to one element, or, if it contains at least two points, they satisfy such an inequality. If r is equal to the minimal interpoint distance

$$\inf\{\delta(x, y) \mid x, y \in H,\ x \neq y\}$$

(when $|\Lambda| \geq 2$) Λ is said to be a uniformly discrete set *of constant* r. The space of uniformly discrete sets of constant $r > 0$ of (H, δ) is denoted by $\mathcal{UD}(H, \delta)_r$. It is the space $\mathrm{SS}(H, \delta)_r$ of systems of equal spheres (or space of sphere packings) of radius $r/2$ of (H, δ): $\Lambda = (a_i)_{i\in\mathbb{N}} \in \mathcal{UD}(H, \delta)_r$ is the set of sphere centers of

$$\mathcal{B}(\Lambda) = \{B(a_i, r/2) \mid i \in \mathbb{N}\} \in \mathrm{SS}(H, \delta)_r$$

where $B(z, t)$ denotes generically the closed ball centered at $z \in H$ of radius $t > 0$.

An element $\Lambda \in \bigcup_{r>0} \mathcal{UD}(H, \delta)_r$ is said to be a *Delone set* if there exists $R > 0$ such that, for all $z \in H$, there exists an element $\lambda \in \Lambda$ with $\delta(z, \lambda) \leq R$ (relative denseness property). Then a Delone set is never empty. If Λ is a Delone set, then

$$R(\Lambda) := \sup_{z\in H} \inf_{\lambda\in\Lambda} \delta(z, \lambda) \tag{2.1}$$

is called the *Delone constant* of Λ. In [86] the range of values of the ratio $R(\Lambda)/r$ in the case $H = \mathbb{R}^n$, $n \geq 1$, is shown to be the continuum

$$\left[\frac{\sqrt{2}}{2}\sqrt{\frac{n}{n+1}},\ +\infty\right). \tag{2.2}$$

In the context of lattices which are $\mathcal{O}_K$-modules, with K a number field, the "ambient space" is obtained in a canonical way via the real and complex embeddings of K and the Delone constant is reminiscent of the Euclidean minimum or inhomogeneous minimum, with possible isolated values instead of the continuum (2.2) (Section 6).

The Delone constant of Λ is the maximal circumradius of all its Voronoi cells. If $\Lambda \in \mathcal{UD}(H, \delta)_r$ is a Delone set of Delone constant R, the discrete set Λ is also called an (r, R)-system [97]. Let $X(H, \delta)_{r,R} \subset \mathcal{UD}(H, \delta)_r$ be the subset of uniformly discrete sets of constant r which are Delone sets of H of Delone constant $\leq R$.

Theorem 2.1. *Let (H, δ) be a σ-compact and locally compact metric space for which* $\mathrm{diam}(H)$ *is infinite. Then, for all $r > 0$, $\mathcal{UD}(H, \delta)_r$ can be endowed with a metric d such that the topological space $(\mathcal{UD}(H, \delta)_r, d)$ is compact and such that the Hausdorff metric on $\mathcal{UD}(H, \delta)_{r,f}$ is compatible with the restriction of the topology of $(\mathcal{UD}(H, \delta)_r, d)$ to $\mathcal{UD}(H, \delta)_{r,f}$. For all $R > 0$ the subspace $X(H, \delta)_{r,R}$ is closed.*

In [87] several (classes of equivalent) metrics on H are constructed. In such constructions a base point, say $\alpha \in \mathbb{R}^n$, is required. When $H = \mathbb{R}^n$, endowed with the Euclidean norm $\|\cdot\|$, the topology on $\mathcal{UD}(\mathbb{R}^n, \|\cdot\|)_r$, $r > 0$, is expressed by "unique local pairings of points in big balls centered at the base point α" as follows (Proposition 3.6 in [87]). Let $r = 1$, the general case being the same.

Proposition 2.2. *Let $\Lambda, \Lambda' \in \mathcal{UD}(\mathbb{R}^n, \|\cdot\|)_1$ with Λ and Λ' nonempty. Let $l = \inf\{\|t - \alpha\| \mid t \in \Lambda\} < +\infty$ and $\varepsilon \in (0, (1 + 2l)^{-1})$. Assume $d(\Lambda, \Lambda') < \varepsilon$. Then, for all $\lambda \in \Lambda$ such that $\|\lambda - \alpha\| < \frac{1-\varepsilon}{2\varepsilon}$,*

(i) *there exists a unique $\lambda' \in \Lambda'$ such that $\|\lambda - \lambda'\| < \frac{1}{2}$,*
(ii) *this pairing (λ, λ') satisfies the inequality $\|\lambda - \lambda'\| \leq (\frac{1}{2} + \|\lambda - \alpha\|)\varepsilon$.*

2.1.2 Cut-and-project schemes above uniformly discrete sets. A locally compact abelian (lca) group is an abelian group G endowed with a topology for which G is a Hausdorff space, each point admits a compact neighbourhood, and such that the mapping $G \times G \to G$, $(x, y) \to x - y$ is continuous. In the sequel we will denote additively the additive law of G so that 0 is the neutral element of G.

Definition 2.3. Let G be a lca group.

(i) A subset Λ of G is *uniformly discrete* if there exists an open neighbourhood W of 0 so that $(\Lambda - \Lambda) \cap W = \{0\}$;
(ii) a subset Λ of G is *relatively dense* if there exists a compact subset K of G such that $G = \Lambda + K$;
(iii) a Delone set of G is a subset Λ of G which is relatively dense and uniformly discrete.

Definition 2.4. A lattice of $\mathbb{R}^n$, $n \geq 1$, is a discrete $\mathbb{Z}$-module of rank n. A lattice in a lca group G is a subgroup L of G such that

(i) L is discrete, i.e. the topology on L induced by that of G is the discrete topology,
(ii) L is cocompact, i.e. G/L is compact.

In the sequel we will only define cut-and-project schemes over uniformly discrete sets Λ which lie in finitely dimensional Euclidean spaces $\mathbb{R}^n$, leaving aside the general case where the ambient space of Λ is an lca group. Such more general constructions can be found in [80], Chap. II, and in [100]. Denote by $\mathcal{L}_n$ the space of (affine) lattices of $\mathbb{R}^n$, $n \geq 1$.

Definition 2.5. A cut-and-project scheme (over $\mathbb{R}^n$) is given by a 4-tuple $(G \times \mathbb{R}^n, L, \pi_1, \pi_2)$ where

(i) $G \times \mathbb{R}^n$ is the direct product of a lca group G and the n-dimensional Euclidean space $\mathbb{R}^n$, $n \geq 1$,
(ii) L is a lattice in $G \times \mathbb{R}^n$

such that the natural projections $\pi_1 : G \times \mathbb{R}^n \to G$ and $\pi_2 : G \times \mathbb{R}^n \to \mathbb{R}^n$ satisfy

(1) the restriction $\pi_2|_L$ of π_2 to L is a bijection from L to $\pi_2(L)$,
(2) the image $\pi_1(L)$ is dense in G.

G is called the *internal space*.

Definition 2.6. Let Λ be a uniformly discrete set in the n-dimensional Euclidean space $\mathbb{R}^n$, $n \geq 1$. A cut-and-project scheme given by the 4-tuple $(G \times \mathbb{R}^n, L, \pi_1, \pi_2)$ is said to lie *above* Λ if there exists $t \in \mathbb{R}^n$ such that

$$\Lambda - t \subset \pi_2(L).$$

Remark 2.7. The need to introduce the translation t in the last definition comes from the fact that the uniformly discrete set Λ does not necessarily contain the base point of the cut-and-project scheme, which is the origin of $\mathbb{R}^n$ and at the same time the origin of G. Being a uniformly discrete set, or a Delone set, is an affine notion in the ambient space $\mathbb{R}^n$, while cut-and-project schemes privilege a base point. For instance the lattice $u + \mathbb{Z}$, $u = 1/2$, of $\mathbb{R}$ admits $(\{0\} \times \mathbb{R}, \mathbb{Z}, 0, \mathrm{Id})$ as cut-and-project scheme above it; the translation t being $1/2$ in this case. If Λ is a Delone set, the translation t can be chosen such that $\|t\| \leq R(\Lambda)$ the Delone constant of Λ.

In the last definition the image $\pi_2(L)$ of the discrete subgroup $L \subset G \times \mathbb{R}^n$ is a $\mathbb{Z}$-module in $\mathbb{R}^n$ (the classical structure of $\mathbb{Z}$-modules in $\mathbb{R}^n$ is given for instance in [38], Theorem 2.3.7).

Cut-and-project sets of $\mathbb{R}^n$, also called model sets, form a particular class of Delone sets.

Definition 2.8. A discrete subset Λ of $\mathbb{R}^n$, $n \geq 1$, is a cut-and-project set, or model set, if there exists a cut-and-project scheme $(G \times \mathbb{R}^n, L, \pi_1, \pi_2)$ over Λ, with G an lca group, and a relatively compact subset Ω of the internal space G, with nonempty interior, such that

$$\Lambda - t = \{\pi_2(\omega) \mid \pi_1(\omega) \in \Omega\},$$

for a certain $t \in \mathbb{R}^n$. The set Ω is called the *window* of the cut-and-project set $\Lambda = \Lambda(\Omega)$.

Model sets which arise from cut-and-schemes $(G \times \mathbb{R}^n, L, \pi_1, \pi_2)$ with an lca group G as internal space do not differ too much from model sets that come from cut-and-project sets where the internal space is $\mathbb{R}^m$, for a certain m, by the following proposition.

Proposition 2.9. *Let $\Lambda(\Omega)$ be a cut-and-project set in the cut-and-project scheme $(G \times \mathbb{R}^n, L, \pi_1, \pi_2)$ where G is an lca group. Then there exists a subgroup of G isomorphic to $\mathbb{R}^m$ for a certain $m \geq 0$, and a model set $\Lambda' \subset \mathbb{R}^n$ having $(\mathbb{R}^m \times \mathbb{R}^n, L', \pi_1, \pi_2)$ as cut-and-project scheme above it such that $\Lambda(\Omega)$ is contained in a finite number of translates of Λ'.*

Proof. Proposition 2.7 in [84]. □

Proposition 2.10. *Suppose that* $\Lambda = \Lambda(\Omega)$ *is a model set in a cut-and-project scheme* $(G \times \mathbb{R}^n, L, \pi_1, \pi_2)$ *where* G *is an lca group. Then we have*:

(i) Λ *is Delone set of* $\mathbb{R}^n$;
(ii) *if* $\Omega \subset \overline{\mathrm{int}(\Omega)}$ (*adherence of its interior*) *and* Ω *generates* G *as a group, the following equality holds*:

$$\mathbb{Z}[\Lambda - \Lambda] = \pi_2(L).$$

Proof. Proposition 2.6 in [84]. □

Remark 2.11. In [84] and [70], the origin implicitly belongs to the Delone set Λ, whereas in the present note we do not assume this minor fact. That is why we refer to $\mathbb{Z}[\Lambda - \Lambda]$ throughout instead of $\mathbb{Z}[\Lambda]$, as in Theorem 1.1 or in Proposition 2.10 for instance.

2.2 Continuity of sphere packings arising from model sets

Let Λ be a model set in $\mathbb{R}^n$, viewed as set of centers of a sphere packing, and consider the cut-and-project scheme $(\mathbb{R}^k \times \mathbb{R}^n, L, \pi_1, \pi_2)$ above Λ which allows the construction of Λ by means of a window $\Omega \subset \mathbb{R}^k$. Let us fix the direct product $\mathbb{R}^k \times \mathbb{R}^n$. Write $\Lambda = \Lambda_L(\Omega)$ and consider how the model set $\Lambda_L(\Omega)$ varies when Ω and L vary continuously.

Let $\mathcal{W}(\mathbb{R}^k)$ be the uniform space of nonempty open relatively compact subsets of $\mathbb{R}^k$ (set of acceptance windows in $\mathbb{R}^k$) whose affine hull is $\mathbb{R}^k$, endowed with the pseudo-metric

$$\Delta_{\mathcal{W}}(\Omega_1, \Omega_2) := \Delta(\overline{\omega}_1, \overline{\omega}_2)$$

where Δ is the Hausdorff metric on the space of nonempty closed subsets of $\mathbb{R}^k$. The space of lattices $\mathcal{L}_{n+k}$ in $\mathbb{R}^{n+k}$ is equipped with the quotient topology of $\mathrm{GL}(n+k, \mathbb{R})/\mathrm{GL}(n+k, \mathbb{Z})$. The metric d built on $\bigcup_{r>0} \mathcal{UD}(\mathbb{R}^{n+k}, \|\cdot\|)_r$ with 0 as base point [87] is compatible with the quotient topology of $\mathrm{GL}(n+k, \mathbb{R})/\mathrm{GL}(n+k, \mathbb{Z})$. Let $\mathcal{UD} = \bigcup_{r>0} \mathcal{UD}(\mathbb{R}^n, \|\cdot\|)_r$ be the space of uniformly discrete subsets of $\mathbb{R}^n$, endowed with the metric d where here an arbitrary base point $\alpha \in \mathbb{R}^n$ is taken (see [87], Theorem 2.1 and Proposition 2.2). Denote by d_α the metric d in this paragraph only. The two origins of the cut-and-project scheme and of $\mathbb{R}^n$ for the construction of the metric d_α on $\mathcal{UD}$ are taken a priori different.

Theorem 2.12. *For any base point* $\alpha \in \mathbb{R}^n$ *the mapping*

$$\mathcal{W}(\mathbb{R}^k) \times \mathcal{L}_{n+k} \to (\mathcal{UD}, d_\alpha) \colon (\Omega, L) \mapsto \Lambda_L(\Omega) \tag{2.3}$$

is continuous.

Proof. Let $\varepsilon > 0$. Let $L_0 \in \mathcal{L}_{n+k}$ and $\Omega_0 \in \mathcal{W}(\mathbb{R}^k)$. Let us show the continuity

at (Ω_0, L_0). Let $t = \|\alpha\| + \frac{1-\varepsilon/2}{\varepsilon}$. Since Ω_0 is open, there exists $\eta_1 > 0$ such that all the sets $\{x \in L \mid \pi_1(x) \in \Omega_0, \|\pi_2(x)\| \leq t\}$ have the same cardinality if L belongs to the open set $\{L \mid d(L, L_0) < \eta_1\}$. Since π_2 is continuous and $\pi_{2_{|L_0}}$ is assumed to be a bijection from L_0 onto $\pi_2(L_0)$, $\pi_{2_{|L}}$ is also a bijection from L onto $\pi_2(L)$ provided $d(L, L_0)$ is small enough. Then, using Proposition 2.2 and invoking the continuity of π_2, there exists $\eta' \leq \eta_1$ such that $d(L, L_0) < \eta'$ implies $d_\alpha(\Lambda_L(\Omega_0), \Lambda_{L_0}(\Omega_0)) < \varepsilon/2$.

The subset $\{x \in L \mid \pi_1(x) \in \Omega, \|\pi_2(x)\| \leq t\}$ of L is such that its projection by π_1 is made of a finite collection of points which lie inside Ω (which is open), and its projection by π_2 is a finite subset of $\pi_2(L)$ which contains $\Lambda_L(\Omega_0) \cap B\big(\alpha, \frac{1-\varepsilon/2}{\varepsilon}\big)$ (see Proposition 2.2). Since the projection mappings π_1 and π_2 are continuous and $\pi_{2_{|L}}$ is a bijection from L onto $\pi_2(L)$, the mapping $L \to \pi_1 \circ (\pi_{2_{|L}})^{-1}$ is continuous on the open set $\{L \mid d(L, L_0) < \eta'\}$. Then there exists $\eta'' > 0$ such that $\Delta_{\mathcal{W}}(\Omega, \Omega_0) < \eta''$ implies $d_\alpha(\Lambda_L(\Omega), \Lambda_L(\Omega_0)) < \varepsilon/2$ (the value of t is chosen according to this last inequality and Proposition 2.2).

Then, as soon as $\Delta_{\mathcal{W}}(\Omega, \Omega_0) < \eta''$ and $d(L, L_0) < \eta'$ hold, we have

$$d_\alpha(\Lambda_L(\Omega), \Lambda_{L_0}(\Omega_0)) \leq d_\alpha(\Lambda_L(\Omega), \Lambda_L(\Omega_0)) + d_\alpha(\Lambda_L(\Omega_0), \Lambda_{L_0}(\Omega_0)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

We deduce the continuity of (2.3). □

Note that the assumption "open" for windows in $\mathcal{W}(\mathbb{R}^k)$ is essential to obtain the continuity in Theorem 2.12. If we consider a collection of model sets parametrized by a sequence of windows which are not necessarily open, but with nonempty interiors, then Theorem 2.12 should be applied with the collections of the interiors of the windows.

For deformations of model sets see [8], [19].

2.3 A classification of uniformly discrete sets

Classes of uniformly discrete sets and Delone sets in $\mathbb{R}^n$, and their relative inclusions, are given in Theorem 2.16, following Lagarias [70] for Delone sets. We first define some classes of uniformly discrete sets intrinsically, i.e. without any cut-and-project scheme formalism above them. Then we indicate the definitions of point sets which invoke cut-and-project schemes.

Definition 2.13. Let Λ be a nonempty uniformly discrete set of $\mathbb{R}^n$.

(i) Λ is *finitely generated* if the $\mathbb{Z}$-module

$$\mathbb{Z}[\Lambda - \Lambda] := \Big\{ \sum\nolimits_{\text{finite}} \alpha_i (x_i - y_i) \mid \alpha_i \in \mathbb{Z},\ x_i, y_i \in \Lambda \Big\}$$

is finitely generated, i.e. $\dim_{\mathbb{Q}} \mathbb{Q} \otimes \mathbb{Z}[\Lambda - \Lambda] < +\infty$.

(ii) Λ is *of finite type* if, for all $t > 0$, the intersection

$$(\Lambda - \Lambda) \cap B(0, t)$$

is a finite set.

If Λ is a nonempty finitely generated uniformly discrete set, the *rank of* Λ, denoted by $\mathrm{rk}\,\Lambda$, is by definition the dimension of the $\mathbb{Q}$-vector space $\mathbb{Q} \otimes \mathbb{Z}[\Lambda - \Lambda] = \mathbb{Q}[\Lambda - \Lambda]$. The rank $\mathrm{rk}\,\Lambda$ is an invariant of Λ. Let $c \in \mathbb{R}^n$. The rank of $\mathbb{Z}[\Lambda - c]$ varies with c and may be different from that of $\mathbb{Z}[\Lambda - \Lambda]$. For instance, with $c = 0$ and $\Lambda = \sqrt{2} + \mathbb{Z}$ in $\mathbb{R}$ we have $\mathrm{rk}\,\Lambda = 1$, while the rank of $\mathbb{Z}[\Lambda] = \mathbb{Z}[\Lambda - c]$ equals 2 (the notations $\mathrm{rk}\,\Lambda$ and $\mathrm{rk}\,\mathbb{Z}[\Lambda]$ should not be confused); moreover $\Lambda = \Lambda - c$ and $\mathbb{Z}[\Lambda - \Lambda] = \mathbb{Z}$ are disjoint.

Theorem 2.14 (Lagarias). *Let Λ be a Delone set of finite type of $\mathbb{R}^n$, $n \geq 1$. Then*

$$\mathrm{rk}\,\Lambda \leq |((\Lambda - \Lambda) \cap B(0, 2R(\Lambda)))| < +\infty \tag{2.4}$$

where $R(\Lambda)$ is the Delone constant of Λ.

Proof. Theorem 2.1 in [70]. □

Definition 2.15. Let Λ be a relatively dense discrete subset of $\mathbb{R}^n$. Λ is a Meyer set if one of the following equivalent assertions is satisfied:

(i) $\Lambda - \Lambda$ is uniformly discrete;
(ii) Λ is a Delone set and there exists a finite set $F \subset \mathbb{R}^n$ such that

$$\Lambda - \Lambda \subset \Lambda + F; \tag{2.5}$$

(iii) Λ is a subset of a model set.

Proof. Theorem 9.1 and Proposition 9.2 in [84]. □

Conditions (i) and (ii) in the definition of Meyer sets are given independently of any "cut-and-project scheme above Λ" consideration while condition (iii) asserts the existence of such a cut-and-project scheme above it. In a similar way a (affine) lattice $L \in \mathcal{L}_n$ in $\mathbb{R}^n$ is intrinsically defined in $\mathbb{R}^n$, without any help of cut-and-project schemes, and admits also $(\{0\} \times \mathbb{R}^n, L, 0, \pi_2)$ as cut-and-project scheme above it and is a model set in this cut-and-project scheme. The objectives of Theorem 1.1 consist in showing the existence of general constructions of cut-and-project schemes above SFU-sets in $\mathbb{R}^n$.

Let $n \geq 1$. Denote:

$\mathrm{M}^{(\mathbb{R})}$:= {Model sets in $\mathbb{R}^n$ arising from cut-and-project schemes having an m-dimensional Euclidean space $\mathbb{R}^m$ as internal space}
$\mathrm{M}^{(\mathrm{lcag})}$:= {Model sets in $\mathbb{R}^n$ arising from cut-and-project schemes having a lca group G as internal space}
$\mathcal{M}^{(\mathbb{R})}$:= {Meyer sets in $\mathbb{R}^n$ arising from cut-and-project schemes having a m-dimensional Euclidean space $\mathbb{R}^m$ as internal space}
$\mathcal{M}^{(\mathrm{lcag})}$:= {Meyer sets in $\mathbb{R}^n$ arising from cut-and-project schemes having a lca group G as internal space}

$\mathcal{UD}$:= {Uniformly discrete sets in $\mathbb{R}^n$}
$\mathcal{UD}_{\rm fg}$:= {Finitely generated uniformly discrete sets in $\mathbb{R}^n$ }
$\subset \cup_{r>0}\mathcal{UD}(\mathbb{R}^n, \|\cdot\|)_r$
$\mathcal{UD}_{\rm ft}$:= {Uniformly discrete sets of finite type in $\mathbb{R}^n$}
$X_{\rm fg}$:= {Finitely generated Delone sets in $\mathbb{R}^n$}
$X_{\rm ft}$:= {Delone sets of finite type in $\mathbb{R}^n$}

Theorem 2.16. *The following inclusions hold*:

$$\begin{array}{ccccccccc} & & & & \mathcal{UD}_{\rm ft} & \subset & \mathcal{UD}_{\rm fg} & & \\ & & & & \cup & & \cup & & \\ \mathcal{M}^{(\rm lcag)} & \supset & \mathrm{M}^{(\rm lcag)} & \supset & \mathcal{L}_n & \subset & \mathrm{M}^{(\mathbb{R})} & \subset & \mathcal{M}^{(\mathbb{R})} \subset X_{\rm ft} \subset X_{\rm fg} \end{array} \tag{2.6}$$

Proof. Theorem 9.1 in [84], Theorem 2.1 and Theorem 3.1 in [70]. □

Definition 2.17. Let $\Lambda \in \mathcal{UD}_{fg}$. If $\{e_1, e_2, \ldots, e_v\}$ is a $\mathbb{Z}$-basis of $\mathbb{Z}[\Lambda]$, i.e. $\mathbb{Z}[\Lambda] = \mathbb{Z}[e_1, e_2, \ldots, e_v]$, then the address map $\varphi : \mathbb{Z}[\Lambda] \to \mathbb{Z}^v$ of Λ associated to this basis is by definition

$$\varphi\Big(\sum_{i=1}^{v} m_i e_i\Big) = (m_1, m_2, \ldots, m_v).$$

In Section 3 we will mainly use address maps of difference sets $\Lambda - \Lambda$ for the elements Λ of $\mathcal{UD}_{fg}$.

2.4 Algebraic integers, inflation centers and self-similarities

Given Λ a uniformly discrete set of $\mathbb{R}^n$, an (affine) *self-similarity* of Λ is by definition a real number $\lambda > 1$ such that

$$\lambda(\Lambda - c) \subset \Lambda - c \tag{2.7}$$

for a certain point c in $\mathbb{R}^n$ (note that c need not belong to Λ). A point $c \in \mathbb{R}^n$ for which (2.7) occurs for a certain $\lambda > 1$ is called an *inflation center* of Λ. The concept of self-similarity is an affine notion and λ depends upon c. Denote by

$$\mathcal{C}(\Lambda) := \{c \mid \text{there exists } \lambda > 1 \text{ such that } \lambda(\Lambda - c) \subset \Lambda - c\} \tag{2.8}$$

the set of inflation centers of Λ and by

$$\mathcal{S}(c) := \{\lambda > 1 \mid \lambda(\Lambda - c) \subset \Lambda - c\} \quad \text{for } c \in \mathcal{C}(\Lambda) \tag{2.9}$$

the set of self-similarities associated with the point c.

Proposition 2.18. *Let $\Lambda = t + \mathbb{Z}$ be an (affine) lattice of $\mathbb{R}$ with $t \in [0, 1)$. Then*:

(i) $\mathcal{C}(\Lambda) = t + \mathbb{Q}$,

(ii) $$\mathcal{S}(c) = \begin{cases} \mathbb{N} \setminus \{0\} & \text{if } c \in t + \mathbb{Z}, \\ (1 - q\mathbb{Z}) \cap \mathbb{N} \setminus \{0\} & \text{if } c \in \mathcal{C}(\Lambda),\ c = t + \frac{p}{q}, \text{ with } p, q \\ & \text{relatively prime } (p \in \mathbb{Z},\ q \geq 2). \end{cases}$$

The set of inflation centers of Λ of given (point) density $1/(2q)$ of self-similarities, with q an integer ≥ 2, is exactly the uniformly discrete set

$$t + \left\{ \frac{p}{q} \mid p \in \mathbb{Z},\ \gcd(p, q) = \pm 1 \right\}.$$

Proof. Routine, with the following definition of the (*point*) *density of self-similarities* of an inflation center $c \in \mathcal{C}(\Lambda)$:

$$\operatorname{dens}(\mathcal{S}(c)) = \limsup_{t \to \infty} \frac{1}{2t} \#(\mathcal{S}(c) \cap (1, t]). \tag{2.10}$$

□

In the general case, given a uniformly discrete set, the characterization of $\mathcal{C}(\Lambda)$ and $\mathcal{S}(c)$ with $c \in \mathcal{C}(\Lambda)$ remains an open problem, even for Delone sets; see [30], [31] and [78] for Penrose tilings and sets $\mathbb{Z}_\beta$ of β-integers [52], with β a quadratic Pisot number. At least $\mathcal{C}(\Lambda)$ is expected to be far from being everywhere dense as in Proposition 2.18.

Definition 2.19. A uniformly discrete of $\mathbb{R}^n$, $n \geq 1$, is a self-similar finitely generated uniformly discrete (SFU-)set if it is finitely generated and admits at least one (affine) self-similarity.

Although a uniformly discrete set of $\mathbb{R}^n$, $n \geq 1$, may be finite, let us observe that a nonempty self-similar finitely generated uniformly discrete (SFU-)set in $\mathbb{R}^n$, $n \geq 1$, is always infinite.

Definition 2.20. Let $\lambda > 1$ be a real algebraic integer. Denote by $\lambda^{(i)}$ its conjugates. We say that λ is

(i) a Pisot number if all its conjugates $\lambda^{(i)}$ satisfy $|\lambda^{(i)}| < 1$,
(ii) a Salem number if all its conjugates $\lambda^{(i)}$ satisfy $|\lambda^{(i)}| \leq 1$, with at least one on the unit circle,
(iii) a Perron number if all its conjugates $\lambda^{(i)}$ satisfy $|\lambda^{(i)}| < \lambda$,
(iv) a Lind number if all its conjugates $\lambda^{(i)}$ satisfy $|\lambda^{(i)}| \leq \lambda$, with at least one on the circle $\{|z| = \lambda\}$.

Theorem 2.21 (Meyer). *Let $\Lambda \subset \mathbb{R}^n$, $n \geq 1$, be a Meyer set. If Λ is an SFU-set, then all its self-similarities are Pisot or Salem numbers.*

Proof. See Theorem 6 in Meyer [82]. □

If $\beta > 1$ is a Pisot number or a Parry number, then the sets $\mathbb{Z}_\beta$ are Meyer sets that admit by construction the self-similarity β [52]. However, there exist many Meyer sets which have no self-similarities at all.

Theorem 2.22 (Lagarias). *Let $\Lambda \subset \mathbb{R}^n$, $n \geq 1$, be a Delone set. If Λ is an SFU-set, then all its (affine) self-similarities λ are algebraic integers such that* degree(λ) *divides* rk Λ. *Moreover, if Λ is of finite type, then all the self-similarities are Perron or Lind numbers.*

Proof. See Theorem 4.1 in Lagarias [70]. □

The concept of (affine) self-similarity is extended in a natural way as follows [84].

Definition 2.23. Let Λ be a nonempty uniformly discrete set of $\mathbb{R}^n$. A self-similarity of Λ is given by a triple (c, λ, Q) where λ is a real number > 1 and Q an element of the orthogonal group $\mathrm{O}(n, \mathbb{R})$ such that

$$\lambda Q(\Lambda - c) \subset \Lambda - c \tag{2.11}$$

for a certain $c \in \mathbb{R}^n$ (note that c belongs or not to Λ).

A point c for which (2.11) holds for a certain pair (λ, Q) is called an *inflation center* of Λ, as in the affine case. Problems on self-similar sets are reported in [93].

2.5 Sets $\mathbb{Z}_\beta$ of beta-integers and Rauzy fractals

Meyer sets $\mathbb{Z}_\beta \subset \mathbb{R}$ of β-integers with β a Pisot number, and their vectorial extension to $\mathbb{R}^n$ – so-called β-grids –, are useful tools for modeling quasicrystals in physics [39], [40], [50], [51]. Indeed, Penrose tilings in the plane and in space play a fundamental role in this modelling process, with suitable positioning of atoms in the tiles. Gazeau [51] has observed that Penrose tilings can easily be deduced from τ-grids, where $\tau = \frac{1+\sqrt{5}}{2}$ is the golden mean, a quadratic Pisot number. Therefore it is natural to extend the constructions of Penrose tilings and β-grids with Pisot numbers β (or more generally with algebraic integers) of higher degree which could be used in the objective of providing possibly new models of aperiodic crystals in crystallography to physicists.

Let us recall, in a general way, the mathematical construction of $\mathbb{Z}_\beta$ on the line and its properties when $\beta > 1$ is a real number, and some open questions related to them when β is in particular an algebraic integer. We refer to [47] [48], [91], [95] and [12] for an overview on recent studies on numeration and its applications.

2.5.1 Construction and properties. For a real number $x \in \mathbb{R}$ the integer part of x will be denoted by $\lfloor x \rfloor$ and its fractional part by $\{x\} = x - \lfloor x \rfloor$. The smallest integer larger than or equal to x will be denoted by $\lceil x \rceil$. For $\beta > 1$ a real number and

$z \in [0,1]$ we denote by $T_\beta(z) = \beta z \pmod 1$ the β-transform on $[0,1]$ associated with β, and iteratively, for all integers $j \geq 0$, $T_\beta^{j+1}(z) := T_\beta(T_\beta^j(z))$, where by convention $T_\beta^0 = \mathrm{Id}$.

Let $\beta > 1$ be a real number. A beta-representation (or β-representation, or representation in base β) of a real number $x \geq 0$ is given by an infinite sequence $(x_i)_{i\geq 0}$ and an integer $k \in \mathbb{Z}$ such that $x = \sum_{i=0}^{+\infty} x_i \beta^{-i+k}$, where the digits x_i belong to a given alphabet ($\subset \mathbb{N}$). Among all the beta-representations of a real number $x \geq 0$, $x \neq 1$, there exists a particular one called Rényi β-expansion, which is obtained through the greedy algorithm [47], [48]: in this case k satisfies $\beta^k \leq x < \beta^{k+1}$ and the digits

$$x_i := \left\lfloor \beta T_\beta^i\Big(\frac{x}{\beta^{k+1}}\Big)\right\rfloor, \quad i = 0, 1, 2, \ldots, \tag{2.12}$$

belong to the finite canonical alphabet $\mathbb{A}_\beta := \{0, 1, 2, \ldots, \lceil \beta - 1\rceil\}$. If β is an integer, then $\mathbb{A}_\beta := \{0, 1, 2, \ldots, \beta - 1\}$; if β is not an integer, then $\mathbb{A}_\beta := \{0, 1, 2, \ldots, \lfloor \beta \rfloor\}$. We denote by

$$\langle x\rangle_\beta := x_0 x_1 x_2 \ldots x_k . x_{k+1} x_{k+2} \ldots \tag{2.13}$$

the couple formed by the string of digits $x_0 x_1 x_2 \ldots x_k x_{k+1} x_{k+2} \ldots$ and the position of the dot, which is at the kth position (between x_k and x_{k+1}). By definition the integer part (in base β) of x is $\sum_{i=0}^{k} x_i \beta^{-i+k}$ and its fractional part (in base β) is $\sum_{i=k+1}^{+\infty} x_i \beta^{-i+k}$. If a Rényi β-expansion ends in infinitely many zeros, it is said to be finite and the ending zeros are omitted. If it is periodic after a certain rank, it is said to be eventually periodic (the period is the smallest finite string of digits possible, assumed not to be a string of zeros).

There is a particular Rényi β-expansion which plays an important role in the theory, which is the Rényi β-expansion of 1, denoted by $d_\beta(1)$ and defined as follows: since $\beta^0 \leq 1 < \beta$, the value $T_\beta(1/\beta)$ is here set (by convention) to 1. Then using (2.12) for all $i \geq 1$, we obtain: $t_1 = \lfloor \beta \rfloor$, $t_2 = \lfloor \beta\{\beta\}\rfloor$, $t_3 = \lfloor \beta\{\beta\{\beta\}\}\rfloor$, etc. The equality $d_\beta(1) = 0.t_1 t_2 t_3 \ldots$ corresponds to $1 = \sum_{i=1}^{+\infty} t_i \beta^{-i}$.

Definition 2.24. A real number $\beta > 1$ such that $d_\beta(1)$ is finite or eventually periodic is called a *beta-number* or more recently *a Parry number* (this new name appears in [40]). In particular, it is called a *simple beta-number* or *a simple Parry number* (after [40]) when $d_\beta(1)$ is finite.

Beta-numbers (Parry numbers) are algebraic integers [91] and all their conjugates lie within a compact subset which looks like a fractal in the complex plane [44], [108]. The conjugates of Parry numbers are all bounded above in modulus by the golden mean $\frac{1}{2}(1 + \sqrt{5})$ [44], [108].

Definition 2.25. The set

$$\mathbb{Z}_\beta := \{x \in \mathbb{R} \mid |x| \text{ is equal to its integer part in base } \beta\}$$

is called set of beta-integers, or set of β-integers, or set of integers in base β.

By construction, the set $\mathbb{Z}_\beta$ is discrete, relatively dense and locally finite (its intersection with any interval of the line is finite), self-similar, with β as self-similarity (with inflation center the origin), and symmetrical with respect to the origin: $\beta\mathbb{Z}_\beta \subset \mathbb{Z}_\beta$, $\mathbb{Z}_\beta = -\mathbb{Z}_\beta$. Its complete set of self-similarities is unknown. Thurston [110] showed that it is uniformly discrete, hence a Delone set when β is a Pisot number. From [52], for β a Pisot number, the relatively dense set $\mathbb{Z}_\beta \cap \mathbb{R}^+$ is finitely generated over $\mathbb{N}$.

Theorem 2.26. *If β is a Pisot number, the Delone set $\mathbb{Z}_\beta$ is a Meyer set which is an SFU-set in $\mathbb{R}$.*

Proof. See [24], [52]. For the definition of an SFU-set see Definition 2.19. It is an SFU-set since $\beta\mathbb{Z}_\beta \subset \mathbb{Z}_\beta$. □

Open problem ($\mathbf{P}_1$). What is the class of real numbers $\beta > 1$ for which $\mathbb{Z}_\beta$ is uniformly discrete or, equivalently, a Delone set ?

We know that this class contains Pisot numbers ([21], [52]) and beta-numbers. It also contains some Salem numbers [23]. It is unknown whether it contains all Salem numbers and all Perron numbers [111]. Problem (P_1) is linked to the specification of the β-shift [21], [52], [111].

The set $\mathbb{Z}_\beta$ contains $\{0, \pm 1\}$ and all the polynomials in β for which the coefficients are given by the equations (2.12). Parry [91] showed that the knowledge of $d_\beta(1)$ suffices to exhaust all the possibilities of such polynomials by the so-called "Conditions of Parry (CP_β)". Let us recall them. Let $(c_i)_{i\geq 1} \in \mathbb{A}_\beta^{\mathbb{N}}$ be the following sequence:

$$c_1c_2c_3\ldots = \begin{cases} t_1t_2t_3\ldots & \text{if } d_\beta(1) = 0.t_1t_2\ldots \text{ is infinite,} \\ (t_1t_2\ldots t_{m-1}(t_m-1))^\omega & \text{if } d_\beta(1) \text{ is finite and equal to } 0.t_1t_2\ldots t_m, \end{cases} \tag{2.14}$$

where $(\ )^\omega$ means that the word within $(\ )$ is indefinitely repeated. When the degree of β is ≥ 2, we have $c_1 = t_1 = \lfloor\beta\rfloor$. Then the polynomial $\sum_{i=0}^{v} y_i\beta^{-i+v} \geq 0$, with $v \geq 0$, $y_i \in \mathbb{Z}$ arbitrary, belongs to $\mathbb{Z}_\beta^+ := \mathbb{Z}_\beta \cap \mathbb{R}^+$ if and only if $y_i \in \mathbb{A}_\beta$ and the following $v+1$ inequalities are satisfied:

$$\begin{aligned} (\mathrm{CP}_\beta)\colon (y_j, y_{j+1}, y_{j+2}, \ldots, y_{v-1}, y_v, 0, 0, 0, \ldots) \prec (c_1, c_2, c_3, \ldots), \\ \text{for all } j = 0, 1, 2, \ldots, v, \end{aligned} \tag{2.15}$$

where "$\prec$" means lexicographically smaller. For a negative polynomial we consider the above criterium applied to its opposite. The conditions of Parry (CP_β) sieve the elements of the ring $\mathbb{Z}[\beta]$ in the number field $\mathbb{Q}(\beta)$.

The set $\mathbb{Z}_\beta$ can be viewed as the set of vertices of the tiling $\mathcal{T}_\beta$ of the real line for which the tiles are the closed intervals whose extremities are two successive β-integers. When β is a Pisot number, the number of (non-congruent) tiles in $\mathcal{T}_\beta$ is finite [110]. If V is a tile of $\mathcal{T}_\beta$ we denote by $l(V)$ its length. If β is a Pisot number and $d_\beta(1)$ is finite, say $d_\beta(1) = 0.t_1t_2\ldots t_m$, then the set of the lengths of the tiles of $\mathcal{T}_\beta$

is exactly

$$\{T_\beta^i(1) \mid 0 \le i \le m-1\}$$
$$= \{1, \beta - t_1, \beta^2 - t_1\beta - t_2, \dots, \beta^{m-1} - t_1\beta^{m-2} - t_2\beta^{m-3} - \cdots - t_{m-1}\}.$$

If β is a Pisot number with $d_\beta(1)$ eventually periodic, say

$$d_\beta(1) = 0.t_1t_2\dots t_m(t_{m+1}t_{m+2}\dots t_{m+p})^\omega,$$

then the set of the lengths of the tiles of $\mathcal{T}_\beta$ is exactly

$$\begin{aligned}\{T_\beta^i(1) \mid 0 \le i \le m+p-1\} &= \{1, \beta - t_1, \beta^2 - t_1\beta - t_2, \dots,\\ &\quad \beta^{m-1} - t_1\beta^{m-2} - t_2\beta^{m-3} - \cdots - t_{m-1},\\ &\quad \beta^m - t_1\beta^{m-1} - t_2\beta^{m-2} - \cdots - t_m, \dots,\\ &\quad \beta^{m+p-1} - t_1\beta^{m+p-2} - t_2\beta^{m+p-3} - \cdots - t_{m+p-1}\}.\end{aligned}$$

Hence, when β is a Pisot number, the set $\mathbb{Z}_\beta$ is a Delone set of (sharp) constants (r, R) with $r = \min\{l(V) \mid V \in \mathcal{T}_\beta\} > 0$ and $R = \frac{1}{2}\max\{l(V) \mid V \in \mathcal{T}_\beta\} = \frac{1}{2}$. The tiling $\mathcal{T}_\beta$ can be obtained directly from a substitution system on a finite alphabet which is associated to β in a canonical way [41], [45], [48].

2.5.2 Rauzy fractals and Meyer sets of beta-integers for β a Pisot number. Rauzy fractals were introduced by Rauzy [1], [2], [45] (Chapter 7), [79] [94], [104], [105] to provide geometric interpretations and geometric representations of symbolic dynamical systems, in the general objective of understanding whether substitutive dynamical systems are isomorphic to already known dynamical systems or if they are new. Rauzy [94] generalized the dynamical properties of the Fibonacci substitution [45] to a three-letter alphabet substitution, called Tribonacci substitution or Rauzy substitution, defined by: $1 \to 12$, $2 \to 13$, $3 \to 1$. The incidence matrix of this substitution is

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

its characteristic polynomial is $X^3 - X^2 - X - 1$ with $\beta > 1$ a Pisot number as dominant root and two complex conjugates roots, α and $\overline{\alpha}$, in the unit disc. The incidence matrix admits as eigenspaces in $\mathbb{R}^3$ an expanding one-dimensional direction and a contracting plane [52].

Theorem 2.27 (Rauzy). *The Rauzy fractal generates a self-similar periodic tiling of the plane. The symbolic dynamical system generated by the Tribonacci substitution is measure-theoretically isomorphic to a toral substitution. The Tribonacci substitutive dynamical system has a purely discrete spectrum.*

Proof. See Rauzy [94]. □

Properties of the Rauzy fractal (connectedness, interiors, boundary, etc.) were obtained in [2], [35], [65], [79], [104], [105], [106], [107]. Gazeau and Verger-Gaugry [52] proved that the set $\mathbb{Z}_\beta$ of integers in base β are in close relation with the Rauzy fractal, within the framework of a canonical cut-and-project scheme $(\mathbb{R}^3 = G \times E, L = \mathbb{Z}^3, \pi_1, \pi_2)$ above $\mathbb{Z}_\beta$, where E is a line in $\mathbb{R}^3$ and G is the corresponding internal space (hyperplane): the Rauzy fractal is the adherence of the image of $\mathbb{Z}_\beta$ by the map $\pi_1 \circ (\pi_{2_{|_L}})^{-1}$ in the internal space G. This situation is quite general for Pisot numbers β and the Rauzy fractal appears as a compact canonical window [52]. However, all the points of L are not selected by this window and only some of them which satisfy the conditions of Parry are projected on E, $\mathbb{Z}_\beta$ being a Meyer set [24], [52].

For all Perron numbers β, the construction of the cut-and-project scheme $(\mathbb{R}^d = G \times E, L = \mathbb{Z}^d, \pi_1, \pi_2)$ over $\mathbb{Z}_\beta$, where d is the degree of β, E is a line in $\mathbb{R}^d$ and G is the corresponding internal space (hyperplane), is canonical and does not use the fact that $\mathbb{Z}_\beta$ should be uniformly discrete [52]. If β is a Pisot number then the image of L by $\pi_1 \circ (\pi_{2_{|_L}})^{-1}$ is relatively compact and its adherence is the (geometric) Rauzy fractal. Whether this image is relatively compact for β a general Salem number is not known.

The $\mathbb{Z}$-module $\mathbb{Z}[\mathbb{Z}_\beta - \mathbb{Z}_\beta]$ is finitely generated for β a Pisot number, but it is not known whether this is the case for Perron numbers in general (which are not Pisot numbers).

Open problem ($\mathbf{P}_2$). What is the class of real numbers $\beta > 1$ for which $\mathbb{Z}_\beta$ is uniformly discrete and is not finitely generated, i.e. for which β one has

$$\operatorname{rank} \mathbb{Z}[\mathbb{Z}_\beta - \mathbb{Z}_\beta] = +\infty \ ?$$

3 Proof of Theorem 1.1 – characterization of SFU-sets

(i) This is the same proof as of Theorem 4.1 (i) in [70]. Let $s = \dim_{\mathbb{R}} \mathbb{R}[\Lambda]$ be the dimension of the $\mathbb{R}$-span of Λ (by $\mathbb{R}$-span of Λ we mean the intersection of all the real affine subspaces of $\mathbb{R}^n$ which contain Λ). Then $1 \leq s \leq n$ and $m := \operatorname{rk} \Lambda \geq s$. By definition the $\mathbb{Z}$-module $\mathbb{Z}[\Lambda - \Lambda] := \left\{ \sum_{\text{finite}} a_i(x_i - x_j) \mid a_i \in \mathbb{Z},\ x_i,\ x_j \in \Lambda \right\}$ admits a set of m generators, say $\{v_1, v_2, \ldots, v_m\}$, which are $\mathbb{Q}$-linearly independent (nonzero) vectors of $\mathbb{R}^n$. Then

$$\mathbb{Z}[\Lambda - \Lambda] = \mathbb{Z}[v_1, v_2, \ldots, v_m].$$

If $\lambda > 1$ is a self-similarity of Λ then there exists $c \in \mathbb{R}^n$ such that $\lambda(\Lambda - c) \subset \Lambda - c$. Since $\Lambda - \Lambda = \Lambda - c - (\Lambda - c)$, this implies

$$\lambda \mathbb{Z}[\Lambda - \Lambda] \subset \mathbb{Z}[\Lambda - \Lambda]. \tag{3.1}$$

We deduce that there exist integers $a_{i,j} \in \mathbb{Z}$ such that

$$\lambda v_i = \alpha_{i,1} v_1 + \alpha_{i,2} v_2 + \cdots + \alpha_{i,m} v_m, \quad i = 1, 2, \ldots, m,$$

with $M = (\alpha_{i,j})_{i,j} \in \mathrm{Mat}_m(\mathbb{Z})$ the space of $m \times m$ integral matrices. Hence

$$\lambda V = MV \quad \text{with} \quad V = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}. \tag{3.2}$$

Equivalently the transposed matrix ${}^tM$ is the matrix associated with the $\mathbb{Q}$-linear map which sends $\{v_1, v_2, \ldots, v_m\}$ to the system $\{\lambda v_1, \lambda v_2, \ldots, \lambda v_m\}$ with respect to the $\mathbb{Q}$-free system $\{v_1, v_2, \ldots, v_m\}$. Since the polynomial $h(X) := \det(XI - M) \in \mathbb{Z}[X]$ is monic and cancels at λ, the real number λ is an algebraic integer of degree less than m.

Let d be the degree of λ and

$$\varphi(X) = X^d + a_1 X^{d-1} + a_2 X^{d-2} + \cdots + a_d, \quad a_i \in \mathbb{Z},\ a_d \neq 0,$$

be the minimal polynomial of λ. From (3.2) we deduce $\lambda^j V = M^j V$ for all $j \in \mathbb{N}$. Hence, since $\varphi(\lambda) = 0$,

$$\varphi(M)V = (M^d + a_1 M^{d-1} + a_2 M^{d-2} + \cdots + a_d)V = 0. \tag{3.3}$$

Since $\varphi(M) \in \mathrm{Mat}_m(\mathbb{Z})$ and that the vectors $v_1, v_2, \ldots, v_m$ are $\mathbb{Q}$-linearly independent, we deduce $\varphi(M) = \varphi({}^tM) = 0$. Hence the minimal polynomial $\psi(X) \in \mathbb{Z}[X]$ of the matrix ${}^tM$ divides $\varphi(X)$ in $\mathbb{Z}[X]$. Since $\varphi(X)$ is irreducible over $\mathbb{Q}$, there is equality: $\psi(X) = \varphi(X)$.

Denote by K the number field $\mathbb{Q}(\lambda)$. Equation (3.1) implies that $\mathbb{Z}[\Lambda - \Lambda]$ is a module over the ring $\mathbb{Z}[\lambda]$ and that $\mathbb{Q}[\Lambda - \Lambda]$ is a K-vector space. The ring $\mathbb{Z}[\lambda]$ is a subring of finite index of the ring of integers $\mathcal{O}_K$ of K. The $m \times m$ integral matrix ${}^tM$ corresponds to an endomorphism of $\mathbb{R}^m$, say u, expressed in the canonical basis $\{e_1, e_2, \ldots, e_m\}$. Since $v_1, v_2, \ldots, v_m$ are $\mathbb{Q}$-linearly independent, the base $\{e_1, e_2, \ldots, e_m\}$ of $\mathbb{R}^m$ and the system $\{v_1, v_2, \ldots, v_m\}$ of $\mathbb{R}^n$ can be identified as well as the two $\mathbb{Q}$-vector spaces $\bigoplus_{i=1}^m \mathbb{Q}e_i$ and $\bigoplus_{i=1}^m \mathbb{Q}v_i$. There are two cases: $m = 1$ and $m > 1$. When $m = 1$ then necessarily λ is an integer > 1 and $d = 1$. When $m > 1$ and $d = 1$, then the matrix M is the diagonal matrix λI and λ is an integer > 1. This case occurs for instance for (affine) lattices Λ of $\mathbb{R}^n$. Now if $m > 1$ and $d \geq 2$, then the endomorphism u induces a Jordan decomposition of $\mathbb{R}^m$ as $K[X]$-module as follows (e.g. [46], pp. 295–301). We assume $d \geq 2$ in the sequel. Let

$$A(\psi) = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_d \\ 1 & \ddots & 0 & \vdots & \vdots \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & 0 & -a_2 \\ 0 & \ldots & & 1 & -a_1 \end{pmatrix} \quad \text{resp.} \quad U_d := \begin{pmatrix} 0 & \ldots & 1 \\ \vdots & 0 & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$$

be the $d \times d$ integral matrix whose terms are 0 except the last column which is composed of the coefficients of $\psi(X)$ (up to sign) and the diagonal under the main diagonal, which is composed with 1, respectively the $d \times d$ matrix whose terms are 0 except the term of the first row and the last column, which is 1 (this makes sense since $d > 1$). Then there exists a basis of $\mathbb{R}^m$, say $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m\}$, in which the matrix of the endomorphism u takes the diagonal form

$$\mathcal{J} := \begin{pmatrix} J_{i_1} & 0 & \ldots & 0 \\ 0 & J_{i_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & J_{i_r} \end{pmatrix}$$

where $i_1 + i_2 + \cdots + i_r = m$, $i_1 \geq i_2 \geq \cdots \geq i_r \geq d$, and J_{i_q}, $1 \leq q \leq r$, is the $i_q \times i_q$ integral matrix given by

$$J_{i_q} := \begin{pmatrix} A(\psi) & 0 & \ldots & 0 \\ U_d & A(\psi) & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & U_d & A(\psi) \end{pmatrix}$$

with 0 everywhere except $A(\psi)$ on the main diagonal and U_d on the diagonal under the main diagonal. Since all the diagonal terms of the matrix J_{i_q} are $A(\psi)$, they are identical, and therefore d divides i_q. Consequently d divides $\sum_{q=1}^{r} i_q = m$.

(ii) Let us pull back this Jordan decomposition to the ambient space $\mathbb{R}^n$ of the uniformly discrete set Λ, block by block. Let us consider the first block J_{i_1}, the situation being the same for the others. The system $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{i_1}\}$ satisfies the following relations:
for $1 \leq \beta < d$, $0 \leq \alpha \leq \frac{i_1}{d} - 1$,

$$u(\varepsilon_{\alpha d+\beta}) = \varepsilon_{\alpha d+\beta+1}; \tag{3.4}$$

for $\beta = d$, $0 \leq \alpha < \frac{i_1}{d} - 1$,

$$u(\varepsilon_{(\alpha+1)d}) = \varepsilon_{(\alpha+1)d+1} - a_1\varepsilon_{(\alpha+1)d} - a_2\varepsilon_{(\alpha+1)d-1} - \cdots - a_d\varepsilon_{\alpha d+1}; \tag{3.5}$$

for $\beta = d$, $\alpha = \frac{i_1}{d} - 1$,

$$u(\varepsilon_{i_1}) = -a_1\varepsilon_{i_1} - a_2\varepsilon_{i_1-1} - \cdots - a_d\varepsilon_{i_1-d+1}. \tag{3.6}$$

Now the matrices ${}^tM$ and $\mathcal{J}$ have coefficients in $\mathbb{Q}$ and are such that there exist a $m \times m$ invertible matrix $C \in \mathrm{GL}(m, \mathbb{R})$ such that ${}^tM = CJC^{-1}$. Then (Corollary 2 in [73], Chap. XV, §3) there exists a $m \times m$ invertible matrix C' in $\mathbb{Q}$ such that

$${}^tM = C'JC'^{-1}.$$

The matrix C' is the matrix associated with the linear map which sends $\{e_1, e_2, \ldots, e_m\}$ to $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m\}$ with respect to the basis $\{e_1, e_2, \ldots, e_m\}$. Let

$$\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} = {}^tC'V \quad \text{with} \quad V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}. \tag{3.7}$$

The $\mathbb{Q}$-free system $\{f_1, f_2, \ldots, f_m\}$ of nonzero vectors of $\mathbb{R}^n$, identified with the basis $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m\}$ of $\mathbb{R}^m$, admits the following structure: from (3.2)

$$\lambda V = MV = {}^t(C'JC'^{-1}) \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \Longleftrightarrow {}^tC'^{-1}(\lambda I - {}^tJ){}^tC' \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} = 0, \tag{3.8}$$

hence

$$\lambda \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = {}^tJ \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}. \tag{3.9}$$

From (3.9), considering the first block J_{i_1}, the situation being the same with the other blocks J_{i_q}, $1 \le q \le r$, we deduce (see (3.4), (3.5)):
for $1 \le \beta < d$, $0 \le \alpha \le \frac{i_1}{d} - 1$,

$$f_{\alpha d+\beta+1} = \lambda f_{\alpha d+\beta} = \lambda^\beta f_{\alpha d+1}; \tag{3.10}$$

for $\beta = d$, $0 \le \alpha < \frac{i_1}{d} - 1$,

$$f_{(\alpha+1)d+1} = \lambda f_{(\alpha+1)d} + a_1 f_{(\alpha+1)d} + a_2 f_{(\alpha+1)d-1} + \cdots + a_d f_{\alpha d+1}. \tag{3.11}$$

Let us show that the assumption $d < i_1$ leads to a contradiction. Assume $d < i_1$. Then we would have $f_{d+1} \neq 0$ from (3.7) since C' is invertible. But from (3.10) and (3.11) with $\alpha = 0$ we obtain

$$\begin{aligned} f_{d+1} &= (\lambda + a_1) f_d + a_2 f_{d-1} + \cdots + a_d f_1 \\ &= (\lambda + a_1)\lambda^{d-1} f_1 + a_2 \lambda^{d-2} f_1 + \cdots + a_d f_1 = \varphi(\lambda) f_1 = 0, \end{aligned}$$

which is a contradiction. Hence $d = i_1$. Proceeding now with the other blocks in the same way leads to the equalities $d = i_1 = i_2 = \cdots = i_r$.

The matrix C' belongs to $\mathrm{GL}(m, \mathbb{Q})$. If $C' \in \mathrm{GL}(m, \mathbb{Z})$ we take

$$w_1 = f_1, \ w_2 = f_{d+1}, \ w_3 = f_{2d+1}, \ \ldots, \ w_r = f_{(r-1)d+1}.$$

Then we deduce from (3.7) and (3.10) (and its analogs for the other blocks) that $\bigoplus_{i=1}^m \mathbb{Z} v_i$ and $\bigoplus_{q=1}^r \bigoplus_{i=1}^m \mathbb{Z}\lambda^i w_q$ are isomorphic as $\mathbb{Z}$-modules. We deduce the result in this case. If $C' \in \mathrm{GL}(m, \mathbb{Q}) \setminus \mathrm{GL}(m, \mathbb{Z})$ let us denote by μ the lcm of the m^2

denominators of the coefficients of C'^{-1} and take

$$w_1 = f_1/\mu,\ w_2 = f_{d+1}/\mu,\ w_3 = f_{2d+1}/\mu,\ \ldots,\ w_r = f_{(r-1)d+1}/\mu.$$

The coefficients of $D := \mu\, {}^tC'^{-1}$ are in $\mathbb{Z}$ and relatively prime so that

$$D \begin{bmatrix} f_1/\mu \\ \lambda f_1/\mu \\ \vdots \\ \lambda^{d-1} f_1/\mu \\ \vdots \\ f_r/\mu \\ \lambda f_r/\mu \\ \vdots \\ \lambda^{d-1} f_r/\mu \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}. \tag{3.12}$$

From (3.12), the $\mathbb{Z}$-module $\bigoplus_{i=1}^{m} \mathbb{Z} v_i$ is a $\mathbb{Z}$-submodule of $\bigoplus_{q=1}^{r} \bigoplus_{i=0}^{d-1} \mathbb{Z}\lambda^i w_q$. Hence the result.

(iii) Since $d = i_1 = i_2 = \cdots = i_r$ and $r = \frac{m}{d}$ by (ii), we deduce that $\mathcal{J}$ is the $m \times m$ diagonal matrix for which the diagonal terms are all identical and equal to $A(\psi)$:

$$\mathcal{J} = \begin{pmatrix} A(\psi) & 0 & \ldots & 0 \\ 0 & A(\psi) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & A(\psi) \end{pmatrix}.$$

Since $\det M = \det \mathcal{J}$ we obtain the characteristic polynomial of M:

$$\det(XI_m - M) = \det(XI_d - A(\psi))^{m/d} = (\varphi(X))^{m/d};$$

in particular, $\det M = (\det A(\psi))^{m/d} = N(\lambda)^{m/d}$, where $N(\lambda) = (-1)^d a_d$, the product of the conjugates of λ, is the algebraic norm of λ. This formula is reminiscent of the algebraic norm of an element in a ring extension [73].

(iv) There are two cases: either (iv-1) $w_1, w_2, \ldots, w_r$ are K-linearly independent, or (iv-2) they are K-linearly dependent. In the first case the cut-and-project scheme above Λ will admit an internal space reduced to its *shadow space* (see below and [70]), while, in the second case, the internal space will come from the shadow space and the space of relations over K between the vectors w_i.

First let us fix some notations. Denote by $\mathcal{C}$ the finite set $\{w_1, w_2, \ldots, w_r\}$ and call it the *central cluster* of the basis $(\lambda^j w_i)_{i=1,\ldots,r,\ j=0,1,\ldots,d-1}$. For $i = 1, 2, \ldots, r$, let $\widetilde{w}_i := \|w_i\|^{-1} w_i$. Let $\widetilde{\mathcal{C}} := \{\widetilde{w}_1, \ldots, \widetilde{w}_r\}$ the image of $\mathcal{C}$ on the unit sphere $\mathbb{S}^{n-1}$ of $\mathbb{R}^n$. We have $|(\widetilde{\mathcal{C}})| \leq r$. We assume that the signature of the field $K = \mathbb{Q}(\lambda)$ of degree d is (r_1, r_2). Then $d = r_1 + 2r_2$. Denote by σ_j, $1 \leq j \leq r_1$, the real embeddings of K in $\mathbb{R}$, and by $\sigma_j, \sigma_{r_2+j} = \overline{\sigma}_j$, where $r_1 + 1 \leq j \leq r_1 + r_2$, the imaginary embeddings

of K in $\mathbb{C}$. Assume $\sigma_1(\lambda) = \lambda$. Let Σ be the embedding of K in $\mathbb{R}^{r_1} \times \mathbb{C}^{2r_2}$ defined by

$$\Sigma(\xi) = (\sigma_1(\xi), \sigma_2(\xi), \ldots, \sigma_{r_1+2r_2}(\xi)) \quad \text{for all } \xi \in K.$$

Let $(g_j)_{1\le j\le d}$ be a $\mathbb{Z}$-basis of the ring of integers $\mathcal{O}_K$ of K. We identify the field K to $\mathbb{Q}^d$ via the mapping Ψ defined by $\Psi(z) = \sum_{i=1}^d z_i g_i$ if $z \in \mathbb{Q}^d$, resp. $\mathcal{O}_K$ to $\mathbb{Z}^d$ if $z \in \mathbb{Z}^d$. The composed mapping $\Phi := \Sigma \circ \Psi$ is then extended in a continuous way from $\mathbb{Q}^d$ to $\mathbb{R}^d$ and denoted in the same way:

$$\Phi(z) := \Big(\sum_{i=1}^d z_i \sigma_1(g_i), \sum_{i=1}^d z_i \sigma_2(g_i), \ldots, \sum_{i=1}^d z_i \sigma_{r_1+2r_2}(g_i) \Big) \quad \text{for all } z \in \mathbb{R}^d.$$

Σ is an injective homomorphism for the ring structures while Ψ is $\mathbb{Q}$-vector space isomorphism. Thus Φ is a $\mathbb{R}$-vector space isomorphism from $\mathbb{R}^d$ onto the étale $\mathbb{R}$-vector space

$$K_{\mathbb{R}} := K \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{R}^{r_1} \times \{z \in \mathbb{C}^{2r_2} \mid z_{r_2+j} = \bar{z} z_j \text{ for all } j = 1, 2, \ldots, r_2\}.$$

The $\mathbb{R}$-subspace $\Sigma(\mathcal{O}_K)$ of $K_{\mathbb{R}}$ is a lattice. Let us extend Φ to $\mathbb{C}^d$ as a $\mathbb{C}$-endomorphism, keeping the same notation, by (with $\mathrm{i} = \sqrt{-1}$):

$$\Phi(x + \mathrm{i}y) = \Phi(x) + \mathrm{i}\Phi(y) \quad \text{for all } x, y \in \mathbb{R}^d.$$

Let us construct the cut-and-project scheme above Λ. Let $x \in \mathbb{R}[\Lambda] \subset \mathbb{R}^n$ be in the $\mathbb{R}$-span of Λ. For $i = 1, 2, \ldots, r$, denote by $p_i(x)$ the orthogonal projection of x onto the line $\mathbb{R}w_i$ and $\widetilde{w}_i := \|w_i\|^{-1} w_i$. Then $p_i(x)$ can be written

$$p_i(x) = \frac{\langle x, w_i \rangle}{\langle w_i, w_i \rangle} w_i = \langle x, \widetilde{w}_i \rangle \widetilde{w}_i,$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean inner product. The point $x \in \mathbb{R}^n$ will be said *K-rational* if the r coefficients $\langle x, \widetilde{w}_i \rangle$ belong to K. In this case, for all $i = 1, \ldots, r$, $\langle x, \widetilde{w}_i \rangle = \sum_{j=1}^d \alpha_{i,j}(x) g_j$ with all coefficients $\alpha_{i,j}(x) \in \mathbb{Q}$. Let us define $\Sigma_i : K \to K_{\mathbb{R}} \widetilde{w}_i$ by $\Sigma_i(\xi) = \Sigma(\xi)\widetilde{w}_i$, for all $i = 1, \ldots, r$, and $\Phi_i = \Sigma_i \circ \Psi$ the $\mathbb{R}$-vector space isomorphism from $\mathbb{R}^d$ onto the $\mathbb{R}$-vector space

$$K_{\mathbb{R}} \widetilde{w}_i := (\mathbb{R}\widetilde{w}_i)^{r_1} \times \{\{z_j \widetilde{w}_i\} \mid z = (z_j) \in \mathbb{C}^{2r_2}, z_{r_2+j} = \bar{z}_j \text{ for all } j = 1, 2, \ldots, r_2\}.$$

The $\mathbb{R}$-subspace $\Sigma_i(\mathcal{O}_K)$ of $K_{\mathbb{R}}\widetilde{w}_i$ is a lattice. For any set A, denote by $\mathrm{pr}_k : A^d \to A$ the k-th projection, so that $\mathrm{pr}_k(K_{\mathbb{R}}\widetilde{w}_i) = \sigma_k(K)\widetilde{w}_i$ for all $i = 1, \ldots, r$ and $k = 1, \ldots, d$. Since the mapping

$$\mathbb{R}[\Lambda] \to \coprod_{i=1}^r \mathrm{pr}_1(K_{\mathbb{R}}\widetilde{w}_i), \quad x \to (\langle x, \widetilde{w}_i \rangle)_i \tag{3.13}$$

is injective, as $\mathbb{R}$-morphism of vector spaces, the $\mathbb{R}$-span $\mathbb{R}[\Lambda]$ of Λ is identified by (3.13) with an s-dimensional $\mathbb{R}$-subspace of the first component $\coprod_{i=1}^r K\widetilde{w}_i =$

$\coprod_{i=1}^{r} \mathrm{pr}_1(K_{\mathbb{R}}\widetilde{w}_i)$. Denote by R_K the subspace of $\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i$ which is the closure of

$$\{(\Sigma_i(\langle x, \widetilde{w}_i\rangle))_i \mid x \in \mathbb{R}[\Lambda] \text{ is } K\text{-rational}\}.$$

It is a product of r_1 copies of $\mathbb{R}[\Lambda]$ and r_2 copies of $\mathbb{C}[\Lambda]$, with $\mathrm{pr}_1(R_K) = \mathbb{R}[\Lambda]$.

The space $\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i$ is a K-vector space, the external law being given by

$$\begin{aligned} K \times \coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i &\to \coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i, \\ (\mu, u) &\mapsto \Sigma(\mu)\cdot u \end{aligned} \tag{3.14}$$

with componentwise multiplication, where $u = (u_k)_{k=1,\dots,d}$, so that the external law, on the k-th component, is given by

$$\begin{aligned} K \times \coprod_{i=1}^{r} \mathrm{pr}_k(K_{\mathbb{R}}\widetilde{w}_i) &\to \coprod_{i=1}^{r} \mathrm{pr}_k(K_{\mathbb{R}}\widetilde{w}_i), \\ (\mu, u_k) &\mapsto \sigma_k(\mu)\cdot u_k. \end{aligned} \tag{3.15}$$

The actions (3.14) and (3.15) are extended from K to $\mathbb{R}$ by continuity. Thus, by (3.15) and since the conjugate fields $\sigma_j(K)$, $\sigma_{r_2+j}(K)$ are not subfields of $\mathbb{R}$ for $j = 1, 2, \dots, r_2$ (if $r_2 \neq 0$), the usual scalar product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^n$ should be considered as the restriction to $\mathbb{R}^n$ of the standard hermitian form on $\mathbb{C}^n$; in particular, it is anti-linear for the second variable.

We now construct a real positive definite symmetric bilinear form on $\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i$. Since

$$\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i = \prod_{k=1}^{d} \Big(\coprod_{i=1}^{r} \mathrm{pr}_k(K_{\mathbb{R}})\widetilde{w}_i \Big)$$

it suffices to construct it on the k-th component $\coprod_{i=1}^{r} \sigma_k(K)\widetilde{w}_i$. Let us define $q_k \colon \coprod_{i=1}^{r} \sigma_k(K)\widetilde{w}_i \times \coprod_{i=1}^{r} \sigma_k(K)\widetilde{w}_i \to \mathbb{R}$ by

$$q_k(U, V) := \begin{cases} \sum_{i=1}^{r} u_i v_i & k = 1, 2, \dots, r_1, \\ \sum_{i=1}^{r} (\bar{u}_i v_i + u_i \bar{v}_i) & k = r_1 + 1, \dots, r_1 + r_2, \end{cases} \tag{3.16}$$

where $U = (u_i\widetilde{w}_i)_i$ and $V = (v_i\widetilde{w}_i)_i$. Then we define

$$q \colon \coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i \times \coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i \to \mathbb{R}, \quad q(U, V) := \sum_{k=1}^{r_1+r_2} q_k(\mathrm{pr}_k(U), \mathrm{pr}_k(V)). \tag{3.17}$$

Let

$$\mathcal{G} := \Big\{ \varphi \in \mathbb{R}^{\mathcal{C}} \mid \frac{\varphi(w)}{\|w\|} \in K \text{ for all } w \in \mathcal{C},\ \sum_{i=1}^{r} \|w_i\|^{-1} \varphi(w_i)\widetilde{w}_i = 0 \Big\}.$$

We call $\mathcal{G}$ the *space of relations over K between the generators $w_1, \dots, w_r$*. The space $\mathcal{G}$ can be identified with a subspace of $K^{\mathcal{C}} \simeq K^r$ and thus with a subspace of

$\coprod_{i=1}^{r} K\widetilde{w}_i$. Denote by G its image by $\coprod_{i=1}^{r}\Sigma_i$ in $\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i$ and by $\overline{G}$ resp. $\overline{\mathcal{G}}$, the closure of G resp. $\mathcal{G}$. For all $x \in \mathbb{R}[\Lambda]$ and all $\varphi \in \mathcal{G}$,

$$\Big\langle x, \sum_{i=1}^{r} \|w_i\|^{-1}\varphi(w_i)\widetilde{w}_i\Big\rangle = \sum_{i=1}^{r} \|w_i\|^{-1}\varphi(w_i)\langle x, \widetilde{w}_i\rangle = 0. \tag{3.18}$$

(3.18) implies that $q_1((\|w_i\|^{-1}\varphi(w_i)\widetilde{w}_i)_i, (\langle x, \widetilde{w}_i\rangle\widetilde{w}_i)_i) = 0$. Then the two subspaces $\mathbb{R}[\lambda]$ and $\overline{\mathcal{G}}$ are orthogonal and complementary in $\coprod_{i=1}^{r}\mathbb{R}\widetilde{w}_i \simeq \mathbb{R}^r$ of dimensions s and $r-s$, respectively.

Let us prove that R_K and $\overline{G}$ are orthogonal and complementary of respective dimension sd and $(r-s)d$ in $\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i$. It suffices to prove $q_k(U, V) = 0$ with $U = (\sigma_k(\|w_i\|^{-1}\varphi(w_i))\widetilde{w}_i)_i$ and $V = (\sigma_k(\langle x, \widetilde{w}_i\rangle)\widetilde{w}_i)_i$, for all $k = 2, 3, \ldots, d$, $x \in \mathbb{R}[\Lambda]$ K-rational and $\varphi \in \mathcal{G}$. We have

$$\begin{aligned}\frac{1}{2}q_k(U, V) &= \mathrm{Re}\Big[\sum_{i=1}^{r}\overline{\sigma_k(\|w_i\|^{-1}\varphi(w_i))}\sigma_k(\langle x, \widetilde{w}_i\rangle)\Big]\\ &= \mathrm{Re}\Big[\sigma_k\Big(\sum_{i=1}^{r}\overline{\|w_i\|^{-1}\varphi(w_i)}\langle x, \widetilde{w}_i\rangle\Big)\Big]\\ &= \mathrm{Re}\Big[\sigma_k\Big(\langle x, \sum_{i=1}^{r}\|w_i\|^{-1}\varphi(w_i)\widetilde{w}_i\rangle\Big)\Big] = 0.\end{aligned}$$

We deduce the claim.

The cut-and-project scheme above Λ we have constructed is the following:

$$\Big(\coprod_{i=1}^{r} K_{\mathbb{R}}\widetilde{w}_i \simeq \overline{G}\times R_K \simeq H \times \mathbb{R}[\Lambda], L, \pi, \mathrm{pr}_1\Big)$$

where $L = \coprod_{i=1}^{r}\Sigma_i(\mathcal{O}_K)$ is a lattice in $\overline{G}\times R_K$ and pr_1 such that $\mathrm{pr}_1(R_K) = \mathbb{R}[\Lambda]$, $\mathrm{pr}_1(\overline{G}) = 0$. Because of the structure of the $\mathbb{Z}$-module $\mathbb{Z}[\Lambda-\Lambda]$ given by (ii), it suffices to take $L = \coprod_{i=1}^{r}\Sigma_i(\mathbb{Z}[\lambda])$. The projection mapping π is $\mathrm{Id} - \mathrm{pr}_1$. The internal space, say H, is $\overline{G}\times(R_K \setminus \mathbb{R}[\Lambda])$. By construction, $\pi(L)$ is dense in H and pr_1 is one-to-one on L, onto $\mathrm{pr}_1(L) = \mathbb{Z}[\lambda][w_1, w_2, \ldots, w_r] \supset \mathbb{Z}[\Lambda-\Lambda]$. Note that $\mathbb{Z}[\Lambda-\Lambda]$ is not necessarily a free $\mathbb{Z}[\lambda]$-module, but it is of finite index in $\mathrm{pr}_1(L) = \mathbb{Z}[\lambda][w_1, w_2, \ldots, w_r]$. The Euclidean structure on the cut-and-project scheme given by q is such that $\mathbb{R}[\Lambda]$ and H are orthogonal. The component $R_K\setminus\mathbb{R}[\Lambda]$ of the internal space H is called the *shadow space* in [70].

If the vectors $w_1, w_2, \ldots, w_r$ are K-linearly independent (case (iv-1)) then $\overline{G}$ is trivial and the internal space H is $R_K \setminus \mathbb{R}[\Lambda]$.

This cut-and-project scheme lies above Λ since, for all $\nu \in \Lambda$, $\Lambda - \nu \subset \Lambda - \Lambda \subset \mathbb{Z}[\Lambda-\Lambda] \subset \mathrm{pr}_1(L)$. We deduce the claim.

4 Proof of Corollary 1.2 – ideal lattices

The objectives of this section are the following: (i) to recall some definitions concerning ideal lattices, referring to [13], [14], [15], [16], [17]; (ii) to show that the sublattice (L', q) of (L, q) such that $\mathrm{pr}_1(L') = \mathbb{Z}[\Lambda - \Lambda]$ in the cut-and-project scheme above the SFU-set Λ given by Theorem 1.1 (iv) is a sublattice of an ideal lattice.

It will suffice to show that the canonical bilinear form q defined by (3.17) has suitable properties.

The *canonical involution* (or complex conjugation) of the algebraic number field K generated by the self-similarity λ is the involution $^-\colon K_{\mathbb{R}} \to K_{\mathbb{R}}$ that is the identity on $\mathbb{R}^{r_1}$ and complex conjugation on $\mathbb{C}^{r_2}$. Let $\mathcal{P} := \{\alpha \in K_{\mathbb{R}} \mid \bar{\alpha} = \alpha \text{ and all components of } \alpha \text{ are } > 0\}$. Let us denote by $\mathrm{Tr}\colon K_{\mathbb{R}} \to \mathbb{R}$ the trace map, i.e. $\mathrm{Tr}(x_1, x_2, \ldots, x_d) = x_1 + \cdots + x_d$.

A *generalized ideal* will be by definition a sub $\mathcal{O}_K$-module of K-rank one of $K_{\mathbb{R}}$. As examples, fractional ideals of K are generalized ideals; ideals of the type uI where I is an $\mathcal{O}_K$-ideal and $u \in K_{\mathbb{R}}$ are also generalized ideals.

Proposition 4.1. *Let $b\colon K_{\mathbb{R}} \times K_{\mathbb{R}} \to \mathbb{R}$ be a symmetric bilinear form. The following statements are equivalent*:

(i) *There exists $\alpha \in K_{\mathbb{R}}$ with $\alpha = \bar{\alpha}$ such that*

$$b(x, y) = \mathrm{Tr}(\alpha x \bar{y})$$

for all $x, y \in K_{\mathbb{R}}$.

(ii) *The identity*

$$b(\mu x, y) = b(x, \bar{\mu} y)$$

holds for all $x, y, \mu \in K_{\mathbb{R}}$.

Proof. [13] Proposition 2.1, [14] Proposition 1. □

An *ideal lattice* is a lattice (I, b) where I is a generalized ideal and $b\colon K_{\mathbb{R}} \times K_{\mathbb{R}} \to \mathbb{R}$ which satisfies the equivalent conditions of Proposition 4.1 with $\alpha \in \mathcal{P}$. Ideal lattices with respect to the canonical involution correspond bijectively to Arakelov divisors of the number field K [13], [14].

It is easy to check that Proposition 4.1 is satisfied by the real symmetric bilinear form q defined by (3.16) and (3.17) in Section 3, where $\alpha = (\alpha_i)_{1 \le i \le d}$ with $\alpha_i = 1$ for all i. We assume $r = 1$, i.e. that the degree of the self-similarity λ is equal to the rank $\mathrm{rk}\,\Lambda$. The lattice L' given by Theorem 1.1 (ii) such that $\mathrm{pr}_1(L') = \mathbb{Z}[\Lambda - \Lambda]$ is of finite index in the $\mathcal{O}_K$-module $\Sigma(\mathcal{O}_K)\widetilde{w}_1$ and of K-rank one in $K_{\mathbb{R}}\widetilde{w}_1$. We have

$$q(\mu U, V) = q(U, \mu V) \quad \text{for all } U, V \in K_{\mathbb{R}}\widetilde{w}_1 \text{ and all } \mu \in K_{\mathbb{R}}.$$

Then, by Proposition 4.1, the bilinear form q has the following expression:

$$q(U, V) = \mathrm{Tr}(\alpha\, U \overline{V}) \quad \text{for all } U, V \in K_{\mathbb{R}}\widetilde{w}_1. \tag{4.1}$$

Therefore (L', q) is a sublattice of finite index of an Arakelov divisor of K in bijection with $\mathbb{Z}[\Lambda - \Lambda]$ by the projection mapping $\mathrm{pr}_{1|_L}$. We deduce Corollary 1.2.

The construction of the real bilinear form q in Theorem 1.1 (iv) which provides the Euclidean structure to the cut-and-project scheme is obtained with $\alpha = (1)_{1\le i\le d}$ in Proposition 4.1. Other choices of α are possible, and the parametrization of the set of possible constants $\alpha = \alpha(\widetilde{w}_1)$ in q of (4.1) is studied for instance in Schoof ([99] and related works); see also Neukirch [88].

5 Lower bounds of densities and pseudo-Delone constants

5.1 Pseudo-Delone sphere packings

The following definition is inspired by the "empty sphere" method of Delone [37]. If $A \subset \mathbb{R}^n$ is any nonempty subset of $\mathbb{R}^n$ and Λ is a uniformly discrete set of constant $r > 0$, we define the density of $\mathcal{B}(\Lambda)$ in A by

$$\delta_A(\mathcal{B}(\Lambda)) := \limsup_{t\to+\infty} \frac{\mathrm{vol}\left(\bigcup_{z_i\in\Lambda, \|z_i\|\le t} B(z_i, r/2)\cap A\right)}{\mathrm{vol}(B(0,t)\cap A)}. \tag{5.1}$$

We omit the subscript "$\mathbb{R}^n$" when $A = \mathbb{R}^n$.

Definition 5.1. A uniformly discrete set Λ of $\mathbb{R}^n$, $n \ge 1$, of constant $r > 0$ is *pseudo-Delone* of constant $R_\xi > 0$ when there exists a sequence $\xi := (x_i, T_i)_i$, where $(x_i)_i$ is a sequence of points of $\mathbb{R}^n$ and $(T_i)_i$ a sequence of real numbers, such that, with the notation $A_\xi := \mathbb{R}^n \setminus \bigcup_i \mathring{B}(x_i, T_i)$, the following holds:

(i) for all i, $T_{i+1} \ge T_i$ with $T_i \ge r/2$;
(i) for all i, $\mathring{B}(x_i, T_i)\cap\mathcal{B}(\Lambda) = \emptyset$;
(i) for all $x \in A_\xi$ there exists $\lambda \in \Lambda$ such that $\|x - \lambda\| \le R_\xi$;
(iv) $\lim_{T\to+\infty} \frac{\mathrm{vol}(A_\xi\cap B(0,T))}{\mathrm{vol}(B(0,T))} = 1$.

If it is finite the infimum $\inf\{R_\xi\}$ over all possible point sets $(x_i)_i$ in $\mathbb{R}^n$ and all collections of radii $(T_i)_i$ such that (i) to (iv) is satisfied is called the *pseudo-Delone constant* of Λ and is denoted by $R(\Lambda)$. Let us call *optimal* a collection ξ such that R_ξ is equal to $R(\Lambda)$.

By Zorn's Lemma, optimal collections exist. The portion of space $\bigcup_i \mathring{B}(x_i, T_i)$ defined by an optimal collection is an invariant, independent of the optimal collection used for defining it. Definition 5.1 means that we can remove the portion of ambient space which does not intervene at infinity for the computation of the density of Λ (in $\mathbb{R}^n$). Note that for an SFU-set of $\mathbb{R}^n$, this portion of space does not contribute to the determination of the generators w_i in Theorem 1.1 since $\bigcup_i \mathring{B}(x_i, T_i)$ contains no point of Λ, and it is legitimate to remove it.

5.2 Proof of Theorem 1.3

Let $R_c := \inf\{R(\Lambda) \mid \Lambda \text{ is uniformly discrete of } \mathbb{R}^n \text{ of constant } 1\}$ be the infimum of possible Delone constants over sphere packings of common radius $1/2$. R_c is a function of n only. Then, for all $r > 0$, rR_c is the infimum of Delone constants of uniformly discrete sets of constant r.

Let $r > 0$ and ω_n be the volume of the unit ball of $\mathbb{R}^n$. Let $R \geq rR_c$ and $T > R$ be a real number. Let Λ be a uniformly discrete set of $\mathbb{R}^n$ of constant $r > 0$ which is pseudo-Delone of pseudo-Delone constant R. Let $\xi := (x_i, T_i)_i$ be an optimal sequence and $A_\xi := \mathbb{R}^n \setminus \cup_i \mathring{B}(x_i, T_i)_i$. For all $\varepsilon > 0$, the pseudo-Delone constant of Λ in A_ξ is smaller than $R + \varepsilon$. Then $(B(0, R+\varepsilon) + \Lambda) \cap B(0, T)$ covers the set $B(0, T - R - \varepsilon) \cap A_\xi$. The number of elements of $\Lambda \cap B(0, T)$ is equal to the number of elements of $\Lambda \cap B(0, T) \cap A_\xi$. This number is at least

$$\frac{\omega_n(T-R-\varepsilon)^n - \mathrm{vol}((\mathbb{R}^n \setminus A_\xi) \cap B(0, T-R-\varepsilon))}{\omega_n(R+\varepsilon)^n}$$
$$= \frac{(T-R-\varepsilon)^n}{(R+\varepsilon)^n}\left(1 - \frac{\mathrm{vol}((\mathbb{R}^n \setminus A_\xi) \cap B(0, T-R-\varepsilon))}{\omega_n(T-R-\varepsilon)^n}\right).$$

On the other hand, since all the balls of radius $r/2$ centered at the elements of $\Lambda \cap B(0, T)$ lie within $B(0, T + r/2)$ and also within A_ξ, the proportion of space they occupy in $B(0, T + r/2) \cap A_\xi$ is at least

$$\left(\frac{T-R-\varepsilon}{R+\varepsilon}\right)^n \left(1 - \frac{\mathrm{vol}((\mathbb{R}^n \setminus A_\xi) \cap B(0, T-R-\varepsilon))}{\omega_n(T-R-\varepsilon)^n}\right) \frac{\mathrm{vol}(B(0, r/2))}{\mathrm{vol}(B(0, T+r/2) \cap A_\xi)}. \tag{5.2}$$

But, for all $\varepsilon > 0$,

$$\lim_{T \to +\infty} \frac{\mathrm{vol}((\mathbb{R}^n \setminus A_\xi) \cap B(0, T-R-\varepsilon))}{\omega_n(T-R-\varepsilon)^n} = 0$$

and

$$\lim_{T \to +\infty} \frac{\mathrm{vol}(B(0, T+r/2))}{\mathrm{vol}(B(0, T+r/2) \cap A_\xi)} = 1.$$

Hence, if T is large enough, the quantity (5.2) is greater than

$$\left(\frac{r(T-R-\varepsilon)}{2(R+\varepsilon)(T+r/2)}\right)^n.$$

When T tends to infinity, this quantity tends to $(2(R+\varepsilon)/r)^{-n}$ for all $\varepsilon > 0$, which is a lower bound of $\delta(\mathcal{B}(\Lambda))$. We deduce the claim.

6 Lower bounds of the Delone constant of an SFU-set

The field $K = \mathbb{Q}(\lambda)$ generated by the self-similarity λ of the SFU-set Λ in Theorem 1.1 has its own Euclidean spectrum [26], [27] which leads to specific geometric properties of the Voronoi cell of the lattice L' [33] of the cut-and-project scheme above Λ, where L' such that $\mathrm{pr}_1(L') = \mathbb{Z}[\Lambda - \Lambda]$. By projection by pr_1, in this cut-and-project scheme above Λ, the Delone (or pseudo-Delone) constant of Λ, whatever the occupation of the elements of $\Lambda - \Lambda$ in $\mathbb{Z}[\Lambda - \Lambda]$, reflects the arithmetical features of K (Euclidean minimum, Euclidean spectrum,… [26], [27]) as well as the geometrical characteristics of the central cluster $\{w_1, \ldots, w_r\}$. This applies, in particular, to the case where Λ is a model set (see Proposition 2.10) or a Meyer set (see Definition 2.15 (iii)), since a window in the internal space controls the thickness of the band around the $\mathbb{R}$-span of Λ which is used for selecting the points of the lattice L'. Recall that the Delone (or pseudo-Delone) constant $R(\Lambda)$ of the SFU-set Λ, if finite, "measures" the maximal size of (spherical) holes in Λ with respect to the portion of space where the density is computed (see §2.1.1, §5 and [86]) In the sequel, we recall these notions and refer to [13], [16], [17], [26], [27], [33], [74].

In the following we will only consider the case $r = 1$, i.e. the case where the degree of the self-similarity λ is equal to the rank of $\mathbb{Z}[\Lambda - \Lambda]$, leaving aside the case $r > 1$. Let us give the name of "spanning self-similarity" to a self-similarity for which $r = 1$. Observe that Proposition 2.18 gives answers in the case of a lattice and a spanning self-similarity.

Theorem 6.1, resp. Theorem 6.2, Corollary 6.3 and Theorem 6.4, is a reformulation in the present context of Theorem 3, resp. Theorem 5, Corollary 6 and Theorem 4 (Remark 2), obtained by Cerri [26]. Recall that, for all $\xi \in K$, $\Sigma_1(\xi) = \Sigma(\xi)\widetilde{w}_1$, with $\widetilde{w}_1 = \|w_1\|^{-1} w_1$ the unit vector.

6.1 Euclidean and inhomogeneous spectra of the number field generated by the self-similarity

Let $N_{K/\mathbb{Q}}$ be the norm defined on K by

$$N_{K/\mathbb{Q}}(\xi) = \prod_{i=1}^{d} \sigma_i(\xi) = \prod_{i=1}^{r_1} \sigma_i(\xi) \prod_{i=r_1+1}^{r_1+r_2} |\sigma_i(\xi)|^2 \quad \text{for all } \xi \in K. \tag{6.1}$$

The field K is said to be *norm-Euclidean* if

$$\text{for all } \xi \in K \text{ exists } y \in \mathcal{O}_K \text{ such that } |N_{K/\mathbb{Q}}(\xi - y)| < 1.$$

Following the notations of Section 3 and [26], [27] we extend $N_{K/\mathbb{Q}} \circ \Psi$ from $\mathbb{Q}^d$ to $\mathbb{R}^d$ by the map denoted by $\mathcal{N}$ as follows:

$$\mathcal{N}(x) = \prod_{i=1}^{d} \Big(\sum_{j=1}^{d} x_j \sigma_i(g_j) \Big) \quad \text{for all } x \in \mathbb{R}^d. \tag{6.2}$$

Let $\xi \in K$. The *Euclidean minimum of* ξ (relatively to the norm $N_{K/\mathbb{Q}}$) is the real number $m_K(\xi) := \inf\{|N_{K/\mathbb{Q}}(\xi - y)| \mid y \in \mathcal{O}_K\}$. The *Euclidean minimum of* K (for the norm $N_{K/\mathbb{Q}}$) is denoted by $M(K)$ and is by definition

$$M(K) := \sup_{\xi \in K} m_K(\xi). \tag{6.3}$$

The mapping $m_K \circ \Psi$ defined on $\mathbb{Q}^d$ is extended to $\mathbb{R}^d$ and is denoted by m

$$m(z) := \inf\{|\mathcal{N}(z - l)| \mid l \in \mathbb{Z}^d\} \quad \text{for} z \in \mathbb{R}^d.$$

The *inhomogeneous minimum of* K is denoted by $M(\overline{K})$ and is defined by

$$M(\overline{K}) := \sup_{x \in \mathbb{R}^d} m(x). \tag{6.4}$$

The mapping $m_K \circ \Sigma_1^{-1}$ is extended to $K_{\mathbb{R}}\widetilde{w}_1$ and is denoted by $m_{\overline{K}}$:

$$m_{\overline{K}}(U) := \inf\left\{|\prod_{i=1}^d (U_i - Z_i)| \mid Z = (Z_i)_i \widetilde{w}_1 \in \Sigma(\mathcal{O}_k)\widetilde{w}_1\right\}$$

for $U = (U_i)_i \widetilde{w}_1 \in K_{\mathbb{R}}\widetilde{w}_1$. The set of values of m_K resp. of m is called the *Euclidean spectrum* resp. the *inhomogeneous spectrum of* K. Successive minima are enumerated: the *second inhomogeneous minimum of* K is defined by

$$M_2(\overline{K}) := \sup_{\substack{x \in \mathbb{R}^d \\ m(x) < M(\overline{K})}} m(x)$$

and the *second Euclidean minimum of* K by

$$M_2(K) := \sup_{\substack{\xi \in K \\ m_K(\xi) < M(K)}} m_K(\xi).$$

Iteratively we define ($p \geq 2$)

$$M_{p+1}(\overline{K}) := \sup_{\substack{x \in \mathbb{R}^d \\ m(x) < M_p(\overline{K})}} m(x),$$

and

$$M_{p+1}(K) := \sup_{\substack{\xi \in K \\ m_K(\xi) < M_p(K)}} m_K(\xi).$$

The inhomogeneous minimum $M(\overline{K})$ of K is said to be *isolated* if

$$M_2(\overline{K}) < M(\overline{K}).$$

This isolation phenomenon has been conjectured for $d = 2$ and K totally real by Barnes and Swinnerton-Dyer. Corollary 6.3 below shows that it occurs frequently.

If the inhomogeneous minimum $M(\overline{K})$ of K satisfies the following property,

$$\text{for all } x \in \mathbb{R}^d \text{ exists } l \in \mathbb{Z}^d \text{ such that } |\mathcal{N}(x - l)| \leq M(\overline{K}), \tag{6.5}$$

we will say that $M(\overline{K})$ is *attained*. Note that (6.5) is not verified for the quadratic field $K = \mathbb{Q}(\sqrt{13})$ [74].

Theorem 6.1. *Assume that the degree d of the field K generated by the self-similarity λ of the SFU-sphere packing Λ is ≥ 3 and is equal to the rank of $\mathbb{Z}[\Lambda - \Lambda]$. If the unit rank $r_1 + r_2 - 1$ of K is > 1, in particular if K is totally real, then*

(i) *there exists $\xi \in K$ such that $M(\overline{K}) = m_{\overline{K}}(\Sigma_1(\xi))$,*

(ii) $M(K) = M(\overline{K}) \in \mathbb{Q}$.

Proof. Theorem 3 in [26]. □

The question whether ξ is unique under some assumptions is not clear [26] [27].

Theorem 6.2. *Assume that the degree d of the field K generated by the self-similarity λ of the SFU-sphere packing Λ is ≥ 3 and is equal to the rank of $\mathbb{Z}[\Lambda - \Lambda]$. If the unit rank $r_1 + r_2 - 1$ of K is > 1 and if K is not a CM-field, in particular if K is totally real, there exists a strictly decreasing sequence $(y_p)_{p\geq 1}$ of positive rational integers, which satisfies*

(i) $\lim_{p\to+\infty} y_p = 0$,

(ii) $m(\mathbb{R}^d) = \bigcup_{p\geq 1} \{y_p\}$,

(iii) *for each $p \geq 1$, the set $\{x + \mathbb{Z}^d \mid m(x) = y_p\}$ of classes modulo the lattice $\mathbb{Z}^d$ is finite and lifts up to points of $\mathbb{Q}^d$, i.e. $m(x) = 0$ for all $x \notin \mathbb{Q}^d$.*

Proof. Theorem 5 in [26]. □

From the definitions, the inequality $M(K) \leq M(\overline{K})$ holds for an arbitrary number field, with equality if $d = 2$ (Barnes and Swinnerton-Dyer [74]). Recently Cerri [26] (Corollary 3 of Theorem 3) proved that the equality $M(K) = M(\overline{K})$ does hold true for every number field.

Corollary 6.3. *Under the same hypotheses $M(\overline{K})$ is attained and*

(i) $M_p(K) = M_p(\overline{K})$ *for all* $p > 1$,

(ii) $M_2(\overline{K}) < M(\overline{K})$ ($M(\overline{K})$*is isolated*),

(iii) $M_{p+1}(\overline{K}) < M_p(\overline{K})$ *and* $\lim_{p\to+\infty} M_p(\overline{K}) = 0$ *for all* $p > 1$.

What are the possible fundamental regions of the sublattices L' of $L = \Sigma_1(\mathbb{Z}[\lambda])$ in $K_{\mathbb{R}}\widetilde{w}_1$?

Theorem 6.4. *Denote for all $t > 0$*

$$\mathcal{A}_t := \left\{U = (U_i)\widetilde{w}_1 \in K_{\mathbb{R}}\widetilde{w}_1 \mid |\prod_{i=1}^{d} U_i| \leq t\right\}.$$

If (unit rank) $r_1 + r_2 - 1 > 1$, then

K is norm-Euclidean $\iff$ there existst $\in (0, 1)$ such that $\Sigma_1(\mathcal{O}_K) + \mathcal{A}_t = K_{\mathbb{R}}\widetilde{w}_1$.

Proof. Remark 2 after Theorem 4 in [26]. □

Then if the unit rank $r_1 + r_2 - 1$ of K is strictly greater than 1 and K is norm-Euclidean, there exists $t \in (0, 1)$ such that the number of copies of $\mathscr{A}_t$ to be considered for obtaining the fundamental region of L' is equal to the index of L' in $\Sigma_1(\mathcal{O}_K)\widetilde{w}_1$, an integer multiple of $(\mathcal{O}_K : \mathbb{Z}[\lambda])$. Recall that ([26] Proposition 4):

(i) $M(K) < 1 \Longrightarrow K$ is norm-Euclidean,
(ii) $M(K) > 1 \Longrightarrow K$ is not norm-Euclidean.

If $M(K) = 1$, the conclusion does not hold except if there exists $\xi \in K$ such that $M(K) = m_K(\xi)$; in this case K is not norm-Euclidean. See [27] for computations of $M_p(K)$, $p \geq 1$, with the conventions $M(K) = M_1(K)$, $M(\overline{K}) = M_1(\overline{K})$.

6.2 Proof of Theorem 1.4

Let us now deduce lower bounds of the Delone constant of the SFU-set $\Lambda \subset \mathbb{R}^n$, $n \geq 1$. We will assume that Λ is a Meyer set (i.e. a Delone set that is a subset of a model set), defined by a window Ω (chosen minimal) in the internal space $\mathbb{R}_K \setminus \mathbb{R}[\Lambda]$ of the cut-and-project scheme

$$(K_{\mathbb{R}}\widetilde{w}_1 = (\mathbb{R}_K \setminus \mathbb{R}[\Lambda]) \times \mathbb{R}[\Lambda], L', \pi, \mathrm{pr}_1) \tag{6.6}$$

given by Theorem 1.1 (iv) with L' such that $\mathrm{pr}_1(L') = \mathbb{Z}[\Lambda - \Lambda]$ and the window Ω nonempty, open and relatively compact such that

$$\Lambda \subset \nu + \{\mathrm{pr}_1(U) \mid U \in L',\ \pi(U) \in \Omega\} \subset \nu + \mathbb{R}\widetilde{w}_1 \quad \text{with } \nu \in \Lambda. \tag{6.7}$$

Λ is a Delone set. The central cluster is $\{w_1\}$. Denote by $\mathbb{R}[\Lambda] + \Omega$ the band $\{P = (U, V) \in K_{\mathbb{R}}\widetilde{w}_1 \mid U = \mathrm{pr}_1(P) \in \mathbb{R}[\Lambda],\ V = \pi(P) \in \Omega\}$ parallel to the one-dimensional $\mathbb{R}$-span $\mathbb{R}[\Lambda]$ of Λ. Note that $\mathbb{R}[\Lambda] = \nu + \mathbb{R}\widetilde{w}_1$ in $\mathbb{R}^n$.

In the sequel we will use the notations of Subsection 6.1 and the assumptions of Theorem 6.2.

Definition 6.5. Let $k \geq 2$ be an integer. The self-similar finitely generated Delone set Λ defined by (6.6) and (6.7) is called

(i) *thin* if the following condition on Ω and L'

$$0 < \|\pi\big(\Sigma_1(x - t)\big)\| < dM(K)^{2/d} \tag{6.8}$$

holds for all $t \in \mathcal{O}_K$ such that $\Sigma_1(t) \in L' \cap (\mathbb{R}[\Lambda] + \Omega)$ and all $x \in m_K^{-1}(M(K))$,

(ii) *k-thin* if we have

$$d(M_{k+1}(K))^{2/d} \leq \|\pi(\Sigma_1(x - t))\| < d(M_k(K))^{2/d} \tag{6.9}$$

for all $t \in \mathcal{O}_K$ such that $\Sigma_1(t) \in L' \cap (\mathbb{R}[\Lambda] + \Omega)$ and $x \in \bigcup_{p=1}^{k-1} m_K^{-1}(M_p(K))$.

This definition is consistent with the following facts:

(a) The values of m constitute a strictly decreasing sequence of positive rational integers (Theorem 6.2 (ii)), each integer being reached a finite number of times modulo L' (Theorem 6.2 (iii)).

(b) The infinite (band) cylinder $\mathbb{R}[\Lambda] + \Omega$ parallel to the one-dimensional space $\mathbb{R}[\Lambda]$ can be made sufficiently narrow in order to avoid the set of points z of $\mathbb{Q}^d$ such that $m_K \circ \Psi(z) \in \bigcup_{p=1}^{k-1} M_p(K)$, for all $k \geq 2$, which is a finite union of translates of L'.

(c) Assertion (b) is possible since *free planes*, a fortiori free lines, do exist in any lattice (equal) sphere packings in $\mathbb{R}^d$ once d is large enough. Henk [61] proved the existence of a $\frac{d}{\log_2(d)}$-dimensional affine plane (called free plane) which does not meet any of the spheres in their interiors. Hence, provided that d is large enough, free lines exist in the lattice sphere packing $\mathcal{B}(L')$ corresponding to L' in $K_{\mathbb{R}}\widetilde{w_1}$. Equivalently narrow bands, with section a nonempty open set, about free lines exist in $K_{\mathbb{R}}\widetilde{w_1}$ which do not intersect L'. By continuity, there exist narrow bands, with nonempty open cross-sections, about free lines, which do not intersect any finite union of translates of L'.

For all $x \in K, t \in \mathcal{O}_K$, by the geometric mean inequality, we have

$$\begin{aligned} |N_{K/\mathbb{Q}}(x)|^2 = |N_{K/\mathbb{Q}}(x-t)|^2 &= \prod_{i=1}^{r_1} |\sigma_i(x-t)|^2 \prod_{i=r_1+1}^{r_1+r_2} |\sigma_i(x-t)|^4 \\ &\leq \left(\frac{1}{d}\sum_{i=1}^{r_1} |\sigma_i(x-t)|^2 + \frac{2}{d}\sum_{i=r_1+1}^{r_1+r_2} |\sigma_i(x-t)|^2\right)^d \qquad (6.10) \\ &= \left(\frac{1}{d}\, q\big(\Sigma_1(x-t), \Sigma_1(x-t)\big)\right)^d \end{aligned}$$

where $q(\Sigma_1(x-t), \Sigma_1(x-t)) = \|\operatorname{pr}_1(\Sigma_1(x-t))\|^2 + \|\pi(\Sigma_1(x-t))\|^2$ with the notations of Section 3. We have

$$\|\operatorname{pr}_1(\Sigma_1(x-t))\|^2 = \|(x-t)\widetilde{w}_1\|^2 = \|(v + x\widetilde{w}_1) - (v + t\widetilde{w}_1)\|^2.$$

The set $m_K^{-1}(M(K)) = \{x \in K \mid m_K(x) = M(K) = M(\overline{K})\}$ is such that

$$\Sigma_1(m_K^{-1}(M(K))) = \{\Sigma_1(x) \mid m_K(x) = M(K) = M(\overline{K})\}$$

is finite modulo L' by Theorem 6.1 and Theorem 6.2.

Let us take x in $m_K^{-1}(M(K))$. Then, from (6.10), for all $t \in \mathcal{O}_K$,

$$M(K)^2 \leq \left(\frac{1}{d}(\|(v + x\widetilde{w}_1) - (v + t\widetilde{w}_1)\|^2 + \|\pi(\Sigma_1(x-t))\|^2)\right)^d.$$

Hence, for all $t \in \mathcal{O}_K$ such that $\Sigma_1(t) \in L' \cap \big(\mathbb{R}[\Lambda] + \Omega\big)$,

$$\|(v + x\widetilde{w}_1) - (v + t\widetilde{w}_1)\|^2 \geq dM(K)^{2/d} - \|\pi(\Sigma_1(x-t))\|^2.$$

Since by hypothesis $\nu + x\widetilde{w}_1$ does not belong to Λ and since we consider the elements $t \in \mathcal{O}_k$ for which $\nu + t\widetilde{w_1}$ belongs to Λ (with $\Sigma_1(t) \in L' \cap (\mathbb{R}[\Lambda] + \Omega)$) we have

$$R(\Lambda) \geq \|(\nu + x\widetilde{w}_1) - (\nu + t\widetilde{w}_1)\|.$$

Hence

$$R(\Lambda)^2 \geq dM(K)^{2/d} - \sup \|\pi(\Sigma_1(x - t))\|^2 > 0 \tag{6.11}$$

where the supremum is taken over all $x \in m_K^{-1}(M(K))$ and all $t \in \mathcal{O}_k$ for which $\nu + t\widetilde{w_1}$ belongs to Λ, with $\Sigma_1(t) \in L' \cap (\mathbb{R}[\Lambda] + \Omega)$.

Let us assume that Λ is k-thin and take x in $\bigcup_{p=1}^{k-1} m_K^{-1}(M_p(K))$. Then the supremum in (6.11) is bounded from above by $dM_k(K)^{\frac{2}{d}}$, which allows to deduce the claim.

Using Theorem 1.3 or [86] we deduce a lower bound of the density of the SFU-set Λ.

The assumption "d large enough" from fact (c) and [61] can be replaced by "$d \geq 3$" (after Theorem 6.2) since λ is a spanning self-similarity and since it is easy to check that $\mathbb{Z}^3 + B(0, 1/2)$, a fortiori $\mathbb{Z}^n + B(0, 1/2)$, already contains a free line.

The isolation phenomenon that frequently occurs (Corollary 6.3 (ii) of Theorem 6.2) in higher dimension is likely to occur in $\mathbb{R}^n$ as well by projection for Λ.

Appendix. Crystallography of aperiodic crystals and Delone sets

New aperiodic states of matter call for mathematical idealizations of packings of atoms. A deep understanding of the mathematics which lies behind the usual experimental techniques, such as diffraction (X-rays, electrons, neutrons, synchrotron radiation, etc.) and inverse problems (crystal reconstruction with local atom clustering, long-range order and self-similarities, etc.) adapted to this new context, is required. Indeed, the situation is well known for (periodic) crystals [34], [57], [58], [68], [102] but fairly unknown or at least badly understood for nonperiodic crystals. Quasicrystals and modulated crystals constitute exceptions since the use of cut-and-project sets allows periodization in higher dimension [3], [4], [55], [67] [69], [102]. The parts of mathematics related to the crystallography of aperiodic crystals are:

- Geometry of Numbers and Discrete Geometry [25], [56], [113];
- N-dimensional crystallography when periodization in higher dimension plays a role [83], [90], [101];
- Spectral Theory, Ergodic Theory and Fourier Transform of Delone sets in connection with diffraction [5], [53], [63], [64], [109];
- Harmonic Analysis as far as density is concerned (as an asymptotic measure).

Atoms are viewed as *hard spheres* and aperiodic crystals as Delone sets (sphere centers). Implicitly this means that atoms behave like spheres, that is, to have a spherical

potential. This is far from covering the large variety of possibilities of chemical boundings between atomic species (see [112] for quasicrystalline models of pure Boron for instance). This provides first-order crystalline models from which the computation of the electron density is made possible. Then comparison with experimental data (densities, physical properties, …) leads to refinement of the models.

Looking for a detailed hierarchy of Delone sets from the mathematical side, either from arithmetics [40], [50] [70] or from tiling theory [6], [10], [85], [108], leads to interesting and new questions concerning crystals, without knowing whether these crystals will exist or not. To conclude let us recall the new definition of a crystal (in $\mathbb{R}^3$), which was recently chosen by the International Union of Crystallography [66], and the former one [101]:

Definition (former definition). A crystal is any solid for which the set of atom positions is a finite union of orbits under the action of a crystallographic group.

Definition (new definition). A crystal is any solid having an essentially discrete diffraction diagram.

By the Poisson formula, the new definition covers all cases of solids that are dealt with in the old definition; see [72] for a proof.

References

[1] P. Arnoux, V. Berthé, H. Ei and S. Ito, Tilings, quasicrystals, discrete planes, generalized substitutions and multidimensional continued fractions, in *Discrete models: combinatorics, computation, and geometry* (Paris, 2001), Discrete Math. Theor. Comput. Sci. Proc., AA, Maison Inform. Math. Discrèt., Paris 2001, 59–78.

[2] P. Arnoux and S. Ito, Pisot substitutions and Rauzy fractals, *Bull. Belg. Math. Soc. Simon Stevin* 8 (2001), 181–207.

[3] F. Axel, F. Dénoyer and J.-P. Gazeau (Eds.), *From quasicrystals to more complex systems* (Les Houches School, 1998) , EDP Sciences, Les Ulis , Springer-Verlag, Berlin 2000.

[4] F. Axel and D. Gratias (Eds.), *Beyond quasicrystals*, EDP Sciences, Les Ulis, Springer-Verlag, Berlin 1995.

[5] M. Baake and D. Lenz, Dynamical systems on translation bounded measures: pure point dynamical and diffraction spectra, *Ergod. Theory Dynam. Systems* 24 (2004), 1867–1893.

[6] M. Baake and R. V. Moody (Eds.), *Directions in Mathematical Quasicrystals*, CRM Monograph Series 13, Amer. Math. Soc., Providence, RI, 2000.

[7] M. Baake, R.V. Moody and P. Pleasants, Diffraction from visible lattice points and k-th power free integers, *Discrete Math.* 221 (2000), 3–42.

[8] M. Baake and B. Sing, Kolakoski-(3, 1) is a (deformed) model set, *Canad. Math. Bull.* 47 (2004), 168–190.

[9] C. Bachoc and B. Venkov, Modular forms, lattices and spherical designs, in *Réseaux euclidiens, designs sphériques et formes modulaires*, ed. by Jacques Martinet, Monogr. Enseign. Math. 37, Enseignement Math., Geneva, 2001, 87–111.

[10] C. Bandt, Self-similar tilings and patterns described by mappings, in [85], 45–83.

[11] E. Bannai, Positive definite unimodular lattice with trivial automorphism groups, *Mem. Amer. Math. Soc.* 85 (1990), n^o 429.

[12] G. Barat, V. Berthé, P. Liardet and J. Thuswaldner, Dynamical directions in numeration, *Ann. Inst. Fourier*, to appear.

[13] E. Bayer-Fluckiger, Upper bounds for Euclidean minima of algebraic number fields, *J. Number Theory*, to appear.

[14] E. Bayer-Fluckiger, Ideal lattices in *A panorama of number theory or the view from Baker's garden* (Zürich 1999), ed. by G. Wüstholz , Cambridge University Press, Cambridge 2002, 168–184.

[15] E. Bayer-Fluckiger, Lattices and number fields, in *Algebraic geometry: Hirzebruch 70* (Warsaw, 1998), Contemp. Math. 241, Amer. Math. Soc., Providence, RI, 1999, 69–84.

[16] E. Bayer-Fluckiger and G. Nebe, Euclidean minima for some real number fields, *J. Théor. Nombres Bordeaux*, to appear.

[17] E. Bayer-Fluckiger and I. Suarez, Ideal lattices over totally real number fields and Euclidean minima, *Arch. Math.* 86 (2006), 217–225.

[18] J. Bernat, Arithmetics in β-numeration, preprint 2005.

[19] G. Bernuau and M. Duneau, Fourier analysis of deformed model sets, in [6], 43–60.

[20] K. Bezdek, Compact packings in the Euclidean space, *Beiträge Algebra Geom.* 25 (1987), 79–84.

[21] F. Blanchard, β-expansions and symbolic dynamics, *Theoret. Comput. Sci.* 65 (1989), 131–141.

[22] K. Böröczky, Jr. *Finite packing and covering*, Cambridge Tracts in Math. 154, Cambridge University Press, Cambridge 2004.

[23] D. W. Boyd, The beta expansion for Salem numbers, in *Organic mathematics* (Burnaby, BC, 1995), CMS Conf. Proc. 20, Amer. Math. Soc., Providence, RI, 1997, 117–131.

[24] C. Burdik, C. Frougny, J.-P. Gazeau and R. Krejcar, beta-integers as natural counting systems for quasicrystals, *J. Phys. A* 31 (1998), 6449–6472.

[25] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin 1959.

[26] J.-P. Cerri, Inhomogeneous and Euclidean spectra of number fields with unit rank greater than 1, preprint.

[27] J.-P. Cerri, Euclidean minima of totally real number fields. Algorithmic determination, preprint.

[28] G. Collinet, Homologie rationelle du groupe $\mathrm{O}_n(\mathbb{Z}[\frac{1}{2}])$ et automorphismes des réseaux unimodulaires, *C. R. Math. Acad. Sci. Paris* 335 (2002), no. 2, 127–132.

[29] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, Springer-Verlag, Berlin 1988.

[30] N. Cotfas, G-model mets and their self-similarities, *J. Phys. A* 32 (1999), 8079–8093.

[31] N. Cotfas, On the self-similarities of a model set, *J. Phys. A* 32 (1999), L165–L168.

[32] N. Cotfas and J.-L. Verger-Gaugry, A mathematical construction of n-dimensional quasicrystals starting from g-clusters, *J. Phys. A* 30 (1997), 4283–4291.

[33] R. Coulangeon, Voronoi theory over algebraic number fields, in *Réseaux euclidiens, designs sphériques et formes modulaires*, ed. by Jacques Martinet, Monogr. Enseign. Math. 37, Enseignement Math., Geneva, 2001, 147–162.

[34] J. M. Cowley, *Diffraction physics*, 2nd ed., North-Holland, Amsterdam 1986.

[35] F. M. Dekking, Recurrent sets, *Adv. Math.* 44 (1982), 78–104.

[36] B. N. Delaunay[†], Neue Darstellung der geometrischen Kristallographie, I , *Z. Kristallogr.* A84 (1932), 109–149.

[37] B. N. Delaunay, Sur la sphère vide, *Bull. Acad. Sci. URSS*, VII Ser. 1934, N^o 6, (1934), 793–800.

[38] D. Descombes, *Eléments de théorie des nombres*, Presses Universitaires de France, Paris 1986.

[39] M. Duneau and M. Audier, Approximant phases of quasicrystals, in [62], 283–333.

[40] A. Elkharrat, C. Frougny, J.-P. Gazeau and J.-L. Verger-Gaugry, Symmetry groups for beta-lattices, *Theoret. Comput. Sci.* 319 (2004), 281–305.

[41] S. Fabre, Substitutions et β-systèmes de numération, *Theoret. Comput. Sci.* 137 (1995), 219–236.

[42] D. K. Fadeev, N. P. Dolbilin, S. S. Ryshkov and M. I. Shtogrin, Boris Nikolaevich Delone (on his life and creative work), *Proc. Steklov Inst. Math.* 1992, no. 4, 1–9.

[43] G. Fejes-Tóth, Multiple packing and covering of spheres, *Acta Math. Hungar.* 34 (1979), 165–176.

[44] L. Flatto, J. Lagarias and B. Poonen, The zeta function of the beta transformation, *Ergodic Theory Dynam. Systems* 14 (1994), 237–266.

[45] N. Pythéas Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lecture Notes in Math. 1794, Springer-Verlag, Berlin 2002.

[46] J. Frenkel, *Géométrie pour l'élève-professeur*, appendice I, Hermann, Paris 1973.

[47] C. Frougny, Number representation and finite automata, in *Topics in symbolic dynamics and applications* (Temuco, 1997), London Math. Soc. Lect. Notes Ser. 279, Cambridge University Press, Cambridge 2000, 207–228.

[48] C. Frougny, Numeration systems, Chapter 7 in *Algebraic Combinatorics on Words*, ed. by M. Lothaire, Encyclopedia Math. Appl. 90, Cambridge University Press, Cambridge 2002.

[49] F. Gaehler, P. Kramer, H.-R. Trebin and Knut Urban (Eds.), Proc. 7th Internat. Conf. Quasicrystals, *Materials Science and Engineering A*, 294–296 (2000).

[50] J.-P. Gazeau, Pisot-cyclotomic integers for quasilattices, in [85], 175–198.

[51] J.-P. Gazeau, Counting Systems with Irrational Basis for Quasicrystals, in [3], 195–217.

[†]It seems that B. N. Delone wished to have his name be written "Delone" and not "Delaunay", but it is common to find both in literature.

[52] J.-P. Gazeau and J.-L. Verger-Gaugry, Geometric Study of the Beta-integers with β a Perron Number and Mathematical Quasicrystals, *J. Théor. Nombres Bordeaux* 16 (2004), 125–149.

[53] J.-P. Gazeau and J.-L. Verger-Gaugry, Diffraction spectra of weighted Delone sets on β-lattices with β a quadratic unitary Pisot number, *Ann. Inst. Fourier*, to appear.

[54] J. E. Goodman and J. O'Rourke, *Handbook of discrete and computational geometry*, CRC Press, Boca Raton 1997.

[55] D. Gratias and L. Michel (Eds.), International workshop on aperiodic crystals, Les Houches (1986) *J. Physique* 47 (1986), no. 7, Suppl. Colloq. C3.

[56] P. M. Gruber and C. G. Lerkkerkerker, *Geometry of numbers*, North-Holland, Amsterdam 1987.

[57] A. Guinier, *Theory and techniques for X-ray crystallography*, Dunod, Paris 1964.

[58] T. Hahn (Ed.), *International tables for crystallography*, vol. A, Reidel Publ. Co., Dordrecht 1983.

[59] T. C. Hales, Sphere packings I, *Discrete Comput. Geom.* 17 (1997), 1–51.

[60] T. C. Hales, Sphere packings II, *Discrete. Comput. Geom.* 18 (1997), 135–149.

[61] M. Henk, Free planes in lattice sphere packings, *Adv. Geom*, 5 (2005), 137–144.

[62] F. Hippert and D. Gratias (Eds.), *Lectures on quasicrystals*, EDP Sciences, Les Ulis 1994.

[63] A. Hof, On Diffraction by Aperiodic Structures, Commun. Math. Phys. 169 (1995), 25–43.

[64] A. Hof, Diffraction by aperiodic structures, in [85], 239–268.

[65] S. Ito and M. Kimura, On Rauzy fractal, *Japan J. Indust. Appl. Math.* 8 (1991), 461–486.

[66] International union of crystallography, Report of the executive committee for 1991, *Acta Cryst. Sect. A* 48 (1992), 922–946.

[67] A. Janner and T. Janssen, Superspace groups, *Phys. A* 99 (1979), 47–76.

[68] C. Janot, *Quasicrystals, a primer*, 2nd ed., Clarendon Press, Oxford 1997.

[69] T. Janssen, From Quasiperiodic to more complex systems, in [4], 75–140.

[70] J. C. Lagarias, Geometric models for quasicrystals I. Delone sets of finite type, *Discrete Comput. Geom.* 21 (1999), 161–192.

[71] J. C. Lagarias, The impact of aperiodic order on mathematics, *Materials Science and Engineering A* 294–296 (2000), 186–191.

[72] J. C. Lagarias, Mathematical quasicrystals and the problem of diffraction, in [6], 61–93.

[73] S. Lang, *Algebra*, Addison-Wesley, Reading, Mass., 1965.

[74] F. Lemmermeyer, The Euclidean algorithm in algebraic number fields, *Exposition. Math.* 13 (1995), 385–416.

[75] J. M. Luck, C. Godrèche, A. Janner and T. Janssen, The nature of the atomic surfaces of quasiperiodic self-similar structures, *J. Phys. A* 26 (1993), 1951–1999.

[76] M. Waldschmidt, P. Moussa, J. M. Luck and C. Itzykson (Eds.), *From number theory to physics* (Les Houches, 1989), Springer Proc. Phys. 47, Springer-Verlag, Berlin 1992.

[77] J. Martinet, *Les réseaux parfaits des espaces euclidiens*, Masson, Paris 1996.

[78] Z. Masáková, J. Patera and E. Pelantová, Inflation centers of the cut-and-project quasicrystals, *J. Phys. A* 31 (1998), 1443–1453.

[79] A. Messaoudi, Propriétés arithmétiques et dynamiques du fractal de Rauzy, *J. Théor. Nombres Bordeaux* 10 (1998), 135–162.

[80] Y. Meyer, *Nombres de Pisot, nombres de Salem et analyse harmonique*, Lecture Notes in Math. 117, Springer-Verlag, Berlin 1970.

[81] Y. Meyer, *Algebraic numbers and harmonic analysis*, North-Holland, Amsterdam 1972.

[82] Y. Meyer, Quasicrystals, Diophantine approximation and algebraic numbers, in [4], 3–16.

[83] L. Michel and J. Morzrzymas, *N-dimensional crystallography*, IHES Workshop on Mathematical Crystallography, IHES/P/85/47, August 1985.

[84] R. V. Moody, Meyer sets and their duals, in [85], 403–442.

[85] R. V. Moody (Ed.), *The Mathematics of long-range aperiodic order*, NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci. 489, Kluwer Academic Publishers, Dordrecht 1997.

[86] G. Muraz and J.-L. Verger-Gaugry, On lower bounds of the density of delone sets and holes in sequences of sphere packings, *Experiment. Math.* 14 (2005), 47–57.

[87] G. Muraz and J.-L. Verger-Gaugry, On a generalization of the selection theorem of Mahler, *J. Théor. Nombres Bordeaux* 17 (2005), 237–269.

[88] J. Neukirch, *Algebraic number theory*, Springer-Verlag, Berlin 1999.

[89] J. Oesterlé, Empilements de sphères, Séminaire Bourbaki, Vol. 1989/90, *Astérisque* 189–190 (1990), Exp. No. 727, 375–397.

[90] W. Opechowski, *Crystallographic and metacrystallographic groups*, North-Holland, Amsterdam 1986.

[91] W. Parry, On the β-expansions of real numbers, *Acta Math. Hungar.* 11 (1960), 401–416.

[92] J. Patera (Ed.), *Quasicrystals and discrete geometry* (Toronto, 1995), Fields Inst. Monogr. 10, Amer. Math. Soc., Providence, RI, 1998.

[93] Y. Peres and B. Solomyak, Problems on self-similar sets and self-affine sets: an update, in *Fractal geometry and stochastics*, II (Greifswald/Koserow, 1998), Progr. Probab. 46, Birkhäuser Basel 2000 , 95–106.

[94] G. Rauzy, Nombres algébriques et substitutions, *Bull. Soc. Math. France* 110 (1982), 147–178.

[95] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Hungar.* 8 (1957), 477–493.

[96] C. A. Rogers, *Packing and covering*, Cambridge University Press, Cambridge 1964.

[97] S. S. Ryshkov, Density of an (r, R)-system, *Math. Notes* 16 (1975), 855–858.

[98] M. Schlottmann, Generalized model sets and dynamical systems, in [6], 143–159.

[99] R. Schoof, Computing Arakelov class groups, in *Surveys in algorithmic number theory*, Cambridge University Press, to appear.

[100] J.P. Schreiber, Approximations diophantiennes et problèmes additifs dans les groupes abéliens localement compacts, *Bull. Soc. Math. France* 101 (1973), 297–332.

[101] R. L. E. Schwarzenberger, *N-dimensional crystallography*, Pitman Res. Notes. Math. Ser. 41, Pitman, London 1980.

[102] M. Sénéchal, *Quasicrystals and geometry*, Cambridge University Press, Cambridge 1995.

[103] D. Shechtman, I. Blech, D. Gratias and J. Cahn, Metallic phase with long-range orientational order and no translational symmetry, *Phys. Rev. Lett.* 53 (1984), 1951–1953.

[104] A. Siegel, Représentations des systèmes dynamiques substitutifs non unimodulaires, *Ergodic Theory Dynam. Systems* 23 (2003), 1247–1273.

[105] A. Siegel, Spectral theory and geometric representations of substitutions, in [45], 199–252.

[106] V. F. Sirvent, Identifications and dimension of the Rauzy fractal, *Fractals* 5 (1997), 281–294.

[107] V. F. Sirvent and Y. Wang, Self-affine tiling via substitution dynamical systems and Rauzy fractals, *Pacific J. Math.* 206 (2002), 465–485.

[108] B. Solomyak, Conjugates of beta-numbers and the zero-free domain for a class of analytic functions, *Proc. London Math. Soc.* (3) 68 (1993), 477–498.

[109] B. Solomyak, Dynamics of Self-Similar Tilings, *Ergodic Theory Dynam. Systems* 17 (1997), 695–738.

[110] W. Thurston, Groups, tilings and finite state automata, preprint 1989.

[111] J.-L. Verger-Gaugry, On gaps in Rényi β-expansions of unity for $\beta > 1$ an algebraic number, *Ann. Inst. Fourier*, to appear.

[112] D. Weygand and J.-L. Verger-Gaugry, Modèle d'un quasicristal covalent de Bore pur, *C. R. Acad. Sci. Paris Sér. II* 320 (1995), 253–257.

[113] C. Zong, *Sphere packings*, Springer-Verlag, Berlin 1999.

Nested quasicrystalline discretisations of the line

Jean-Pierre Gazeau[*], *Zuzana Masáková, and Edita Pelantová*

APC, Boite 7020, Université Paris 7-Denis Diderot
2 place Jussieu 75251 Paris Cedex 05, France
email: gazeau@ccr.jussieu.fr

Department of Mathematics, FNSPE, Czech Technical University
Trojanova 13, 120 00 Praha 2, Czech Republic
emails: masakova@km1.fjfi.cvut.cz, pelantova@km1.fjfi.cvut.cz

Abstract. One-dimensional cut-and-project point sets obtained from the square lattice in the plane are considered from a unifying point of view and in the perspective of aperiodic wavelet constructions. We successively examine their geometrical aspects, combinatorial properties from the point of view of the theory of languages, and self-similarity with algebraic scaling factor θ. We explain the relation of the cut-and-project sets to non-standard numeration systems based on θ. We finally examine the substitutivity, a weakened version of substitution invariance, which provides us with an algorithm for symbolic generation of cut-and-project sequences.

Contents

1 Introduction

Initially introduced by Y. Meyer [48], [49] in the context of Harmonic Analysis and more specifically of *harmonious sets*, the cut-and-project sets or model sets have become during the two last decades a kind of geometrical paradigm in quasicrystalline studies. Quasicrystals are those alloys whose first sample was discovered in 1982 by

[*] UMR 7164, (CNRS, Université Paris 7, CEA, Observatoire de Paris)

Shechtman, Blech, Gratias, and Cahn [61], namely the alloy $Al_{0.86}Mn_{0.14}$, characterized by

i) a diffraction pattern like a dense constellation of more-or-less bright spots, which is an indication of a long-range order;
ii) a spatial organisation of those Bragg peaks obeying five- or ten-fold symmetries, at least locally, which indicates a sort of icosahedral organisation in real space with five-fold symmetries;
iii) a spatial organisation of those Bragg peaks obeying specific scale invariance, more precisely invariance under dilations by a factor equal to some power of the golden mean $\tau = \frac{1+\sqrt{5}}{2}$ and manifestly consistent with the five-fold symmetry since $\tau = 2\cos\frac{2\pi}{10}$.

One-dimensional examples which are usually presented as toy geometrical models of quasicrystals [40] appertain to the so-called Fibonacci chain family. They are discrete quasiperiodic subsets of the real line and are often presented as an illustration of the cut-and-project method, mainly developed in this context by [23], [33]. Consider a semi-open band $\mathcal{B}$ obtained by translating the unit square through the square lattice $\mathbb{Z}^2$ along the straight line D_1 of slope ε. D_1 is referred to as a "cut" or "parallel" space or "physical space". Then project on D_1 and along a straight line D_2 the lattice points lying in $\mathcal{B}$. Note that the latter points belong to a unique path made of horizontal segments (A) and vertical segments (B). The resulting sequence of points lying in D_1 are the nodes of a specific Fibonacci chain if $\varepsilon = \frac{1}{\tau}$ and $D_2 = D_1^{\perp}$. Let us denote this set of nodes by $\mathcal{F}$. The chain itself is made of the projected paths and reads $\ldots ABAABABABAABAA\ldots$. Note that a short link B is never adjacent to another B whereas two adjacent long links A can occur. D_2 is called the "internal" space, and $\mathcal{B} \cap D_2$ is the "window" or "acceptance zone", or also "atomic surface".

The set of Fibonacci nodes is equivalently obtained through a purely algebraic filtering procedure. Let us first consider the so-called extension ring of the algebraic integer τ:

$$\mathbb{Z}[\tau] = \{x = m + n\tau \mid m, n \in \mathbb{Z}\} = \mathbb{Z} + \mathbb{Z}\tau.$$

It can be obtained as the projection onto D_1 and along D_2 of the whole square lattice $\mathbb{Z}^2$. There exists in this type of a ring an algebraic conjugation, called Galois automorphism, and defined by

$$x = m + n\tau \mapsto x' = m + n\tau',$$

where $\tau' = -\frac{1}{\tau} = \frac{1-\sqrt{5}}{2}$ is the other root of the golden mean equation $x^2 = x + 1$. Then define the point set $\Sigma(\Omega)$ using an internal sieving rule in the ring $\mathbb{Z}[\tau]$ itself [52]:

$$\Sigma(\Omega) = \{x = m + n\tau \in \mathbb{Z}[\tau] \mid x' = m - n\tfrac{1}{\tau} \in \Omega\} = (\mathbb{Z}[\tau] \cap \Omega)'.$$

The Fibonacci point set $\mathcal{F}$ in the above, with link lengths $A = \tau^2$, $B = \tau$, is precisely that set $\Sigma(\Omega)$ with $\Omega = [0, 1)$.

As was previously mentioned, self-similarity plays a fundamental structural role in the existence of quasicrystals. That property is perfectly illustrated by the Fibonacci point set $\mathcal{F} = \Sigma[0, 1)$ since we check from the algebraic definition that $\tau^2\,\Sigma[0, 1) = \Sigma[0, 1/\tau^2) \subset \Sigma[0, 1)$ and so $\tau^2\,\mathcal{F} \subset \mathcal{F}$: this particular Fibonacci point set is self-similar with scaling factor equal to τ^2. Immediately we get the infinite nested sequence

$$\cdots \subset \mathcal{F}/\tau^{2j-2} \subset \mathcal{F}_j := \mathcal{F}/\tau^{2j} \subset \mathcal{F}/\tau^{2j+2} \subset \cdots, \tag{1.1}$$

as increasing aperiodic discretizations of $\mathbb{R}$. Since the distance between two adjacent points of $\mathcal{F}_j$ is equal to $1/\tau^{2j}$ or $1/\tau^{2j+1}$, it is clear that the inductive limit $\mathcal{F}_\infty := \lim_{j\to\infty} \mathcal{F}_j$ densely fills the real line. It is precisely this property which led us to examine in recent works (see [29], [5], [4] and references therein) the problem of constructing wavelets by following discretization schemes of the real line like (1.1). In [4], the construction was based on multiresolution analysis with spline wavelets earlier elaborated by Lemarié-Rieusset [39] and Bernuau [11], [12] (more details will be given below). Our aim was to eventually apply these wavelets to the analysis of aperiodic structures, like diffraction spectra of Fibonacci chain or some other related spectral problem, and to compare our results with more standard wavelet analysis (e.g. dyadic wavelets) [6].

Let us recall the main features of wavelet analysis in the framework of the Hilbert space $L^2(\mathbb{R})$. More complete information can be found in comprehensive textbooks, like [43]. We shall just outline here the essential of that field whose development since the beginning of the eighties amazingly parallels that one of quasicrystals. Under the name *wavelet* is commonly understood a function $\psi(x) \in L^2(\mathbb{R})$ such that the family of functions $\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k)$ for $j, k \in \mathbb{Z}$ forms an orthonormal (in a restrictive sense) or at least a *Riesz basis* for $L^2(\mathbb{R})$. A family of vectors (v_n) in a separable Hilbert space V is a Riesz basis if and only if each $v \in V$ can be expressed uniquely as $v = \sum_n a_n v_n$ and there exist positive constants K_2 and K_2, $0 < K_1 \leq K_2$, such that

$$K_1 \sum_n |a_n|^2 \leq \Big\| \sum_n a_n v_n \Big\|^2 \leq K_2 \sum_n |a_n|^2$$

for all sequence of scalars a_n. We can say that the v_n's are strongly linearly independent and, if $K_1 = 1 = K_2$, then the basis is orthonormal. A function (or a set of functions) generating through dilations and translations an orthonormal basis for $L^2(\mathbb{R})$ can be found using a *multiresolution analysis* of $L^2(\mathbb{R})$ (shortly MRA), a method settled by S. Mallat [44] and precisely based on an increasing sequence of periodic discretizations of $\mathbb{R}$. The genuine MRA ingredients are:

(i) one *scaling* function $\varphi(x) \in L^2(\mathbb{R})$ such that $\{\varphi(x - k)\}_{k\in\mathbb{Z}}$ is an orthonormal system;

(ii) the Hilbert subspace V_0 which is the linear span of $\{\varphi(x-k) \mid k \in \mathbb{Z}\}$ and which corresponds to the "central" element of the sequence of discretizations of $\mathbb{R}$;

(iii) the increasing sequence of nested Hilbert subspaces $\cdots V_{j-1} \subset V_j \subset V_{j+1} \cdots$ which are defined by $f(x) \in V_0 \Leftrightarrow f(2^j x) \in V_j$ and are such that $\bigcap_j V_j = \{0\}$ and $\bigcup_j V_j$ is dense in $L^2(\mathbb{R})$;

(iv) one wavelet, i.e., a function $\psi(x)$ such that $\{\psi(x-k) \mid k \in \mathbb{Z}\}$ spans the orthogonal complement W_0 of V_0 in $V_1 = V_0 \oplus W_0$.

Note that the orthonormality of the basis $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ of the subspace V_0 is a strong constraint. This condition is usually weakened by just imposing that the system $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ be a Riesz basis of V_0.

We thus note that the dilatation factor is genuinely $\theta = 2$. Indeed, the construction of a wavelet basis within the MRA framework relies on the fact that the lattices $2^{-j}\mathbb{Z}$ are increasing for the inclusion. This property is preserved only when θ is an integer. Then, what about choosing another number θ as a scaling factor? Auscher [8] considered the following problem: given a real number $\theta > 1$, does there exist a finite set $\{\psi_1, \psi_2, \ldots, \psi_\ell\}$ of functions in $L^2(\mathbb{R})$ such that the family $\theta^{j/2}\psi_i(\theta^j x - k)$, $j, k \in \mathbb{Z}$, $1 \le i \le \ell$, is an orthonormal basis for $L^2(\mathbb{R})$? Then he proved that a basis of this type exists if θ is a rational number. More precisely, for $\theta = p/q > 1$, p and q being relatively prime integers, there exists a set of $p - q$ wavelet functions satisfying the previous condition. Now, what about an irrational scaling factor? We already mentioned the works [11], [12] and [4], the latter being mainly devoted to the cases in which θ is encountered in quasicrystallography, like $\tau = (1+\sqrt{5})/2$ or $\tau^2 = (3+\sqrt{5})/2$. Other "quasicrystallographic" numbers have been observed: $1+\sqrt{2}$, $2+\sqrt{3}$. All of them belong to the class of quadratic Pisot–Vijayaraghavan units. To such numbers are associated discretization sequences like in (1.1),

$$\cdots \subset \Lambda_{j-1} \subset \Lambda_j := \Lambda/\theta^j \subset \Lambda_{j+1} \subset \cdots, \tag{1.2}$$

where $\Lambda = \Lambda_0$ is a selfsimilar Delone set with scaling factor θ, $\theta\,\Lambda \subset \Lambda$ and is such that the inductive limit $\Lambda_\infty := \lim_{j\to\infty} \Lambda_j$ densely fills the real line. Recall that by Delone set we mean that Λ is *uniformly discrete* (the distances between any pair of points in Λ are greater than a fixed $r > 0$) and *relatively dense* (there exists $R > 0$ such that $\mathbb{R}$ is covered by intervals of length $2R$ centered at points of Λ).

In 1992 Buhmann and Micchelli [18] proposed a construction of a wavelet spline basis corresponding to non-uniform and non-self-similar knot sequences, which are actually nested sequences of Delone sets

$$\cdots \subset \Lambda_{j-1} \subset \Lambda_j \subset \Lambda_{j+1} \subset \cdots. \tag{1.3}$$

They consider only two successive elements, say Λ_0 and $\Lambda_1 \supset \Lambda_0$ for their purpose of proving the existence of what they call prewavelets, with minimal support, which span the orthogonal complement W_0 of V_0 in $V_1 = V_0 \oplus W_0$. Here, V_0 and V_1 are the spaces of linear combinations of *B-splines* on Λ_0 and Λ_1 respectively. They first suppose that the "refining" of the "coarse" knot sequence Λ_0 leading to the "finer" Λ_1 consists in adding a new knot between each two adjacent knots of Λ_0. The case of multiple insertions is eventually examined. Let us now give some insight on these spline functions, so intimately linked to the notion of a Delone set.

Any Delone set Λ determines a space of splines of order s, $s \geq 2$, in the following way.

Definition 1.1. Let $s \geq 2$. Then $V_0^{(s)}(\Lambda)$ is the closed subspace of $L^2(\mathbb{R})$ defined by

$$V_0^{(s)}(\Lambda) := \left\{ f(x) \in L^2(\mathbb{R}) \mid \tfrac{\mathrm{d}^s}{\mathrm{d}x^s} f(x) = \textstyle\sum_{\lambda \in \Lambda} a_\lambda \delta_\lambda \right\}.$$

An equivalent definition is given in terms of the restriction of functions to intervals determined by consecutive elements of Λ. Suppose $\Lambda = \{\lambda_n \mid n \in \mathbb{Z}\}$, where $\lambda_n < \lambda_{n+1}$ for all $n \in \mathbb{Z}$. There results from Definition 1.1 that

$$V_0^{(s)}(\Lambda) = \{ f \in C^{s-2} \cap L^2(\mathbb{R}) \mid f_{|[\lambda_n, \lambda_{n+1}]} \text{ is a polynomial of degree } \leq s-1 \}.$$

Therefore, $V_0^{(s)}(\Lambda)$ is the space of splines of order s with nodes in Λ. Let us now give a classical result about the existence of a Riesz basis for $V_0^{(s)}(\Lambda)$ [60].

Theorem 1.2. *For all Delone sets $\Lambda = \{\lambda_n \mid n \in \mathbb{Z}\} \subset \mathbb{R}$ and for all $s \geq 2$, there exists a Riesz basis $\{B_\lambda^{(s)} \mid \lambda \in \Lambda\}$ of $V_0^{(s)}(\Lambda)$. The function $B_\lambda^{(s)}$ (called B-spline) is the unique function in $V_0^{(s)}(\Lambda)$ satisfying the following conditions*:

(i) supp $B_\lambda^{(s)} = [\lambda, \lambda']$, *where* $\lambda' \in \Lambda$.
(ii) *The interval* (λ, λ') *contains exactly* $s-1$ *points of* Λ.
(iii) $\int_{\mathbb{R}} B_\lambda^{(s)} = \frac{\lambda' - \lambda}{s}$.

See [60] for proof. Note that (i) and (ii) give precise information on the (compact) support of $B_\lambda^{(s)}$ whilst (iii) is a normalization condition.

The construction of $B_\lambda^{(s)}(x)$ can be carried out in various ways, by recurrence, by using the condition of minimal support, or by inverse Fourier transform. In the latter case, one can prove that the Fourier transform of $B_{\lambda_n}^{(s)}(x + \lambda_n)$ depends on the s-tuple $(\lambda_{n+1} - \lambda_n, \ldots, \lambda_{n+s} - \lambda_n)$ only. Now suppose that the Delone set Λ is of *finite local complexity*, which means [38], [36] that, for all $R > 0$, the point set

$$\bigcup_{\lambda \in \Lambda} \{(\Lambda - \lambda) \cap (-R, R)\}$$

is finite, i.e., local environments of points in Λ are not different in infinite fashions. Typically, such sets Λ are mathematical models for one-dimensional structures having a long-range order, like quasicrystals. We can then assert the following.

Proposition 1.3. *Let $\Lambda \subset \mathbb{R}$ be a Delone set of finite local complexity. Then the B-splines of order s based on Λ are of the form $B_\lambda^{(s)}(x) = \phi_\lambda(x - \lambda)$, $\lambda \in \Lambda$, where the set $\{\phi_\lambda(x) \mid \lambda \in \Lambda\}$ is a finite set of functions with compact support.*

Therefore, in the finite local complexity case, it is possible to partition the indexing set $\mathbb{Z}$ for Λ into a finite set of equivalence classes $\bar{0}$, $\bar{1}$, …, $\bar{q}$, where the equivalence

between k and n is given by

$$B^{(s)}_{\lambda_k}(x+\lambda_k) = B^{(s)}_{\lambda_n}(x+\lambda_n) \quad \text{for all } x.$$

Correspondingly, for a given s, the point set Λ is partitioned into $\Lambda = \bigcup_{\bar{n}=\bar{0}}^{\bar{q}} \Lambda_{\bar{n}}$ with $\Lambda_{\bar{n}} = \{\lambda_k \in \Lambda \mid k \in \bar{n}\}$. The equivalence between k and n means that λ_k and λ_n are left-hand ends of identical s-letter words if we identify each interval $(\lambda_k, \lambda_{k+1})$ with a letter of the allowed alphabet. To each class $\bar{n}$ is biunivocally associated the function $\phi_{\bar{n}}(x) \equiv \phi_{\lambda_k}(x) = B^{(s)}_{\lambda_k}(x+\lambda_k)$, $k \in \bar{n}$. In this way, the space $V_0^{(s)}(\Lambda)$ decomposes into the direct sum

$$V_0^{(s)}(\Lambda) = \bigoplus_{\bar{n}=\bar{0}}^{\bar{q}} V_{0,\bar{n}},$$

where $V_{0,\bar{n}}$ is the closure of the linear span of the functions $\phi_{\bar{n}}(x-\lambda_k)$, $k \in \bar{n}$.

Let us now go back to the case in which there is self-similarity (like in (1.2)) and finite local complexity and let us see which issue holds in term of MRA and existence and properties of wavelets. More concretely, let Λ be a Delone set of finite local complexity and self-similar with inflation factor $\theta > 1$, $\theta\Lambda \subset \Lambda$. Changing the scale allows us to define subspaces $V_j^{(s)}(\Lambda)$, $j \in \mathbb{Z}$, as

$$V_j^{(s)}(\Lambda) = \left\{ f(x) \in L^2(\mathbb{R}) \mid \tfrac{\mathrm{d}^s}{\mathrm{d}x^s} f(x) = \textstyle\sum_{\lambda\in\Lambda} a_\lambda \delta_{\theta^{-j}\lambda} \right\}.$$

Therefore, $V_j^{(s)}(\Lambda)$ is the space of splines of order s with nodes in the jth scaled version of Λ.

We now have at our disposal an inductive chain of spaces allowing analysis at any scale. More precisely, with the above notations, we have the following statement.

Proposition 1.4. *The sequence of subspaces $(V_j^{(s)}(\Lambda))_{j\in\mathbb{Z}}$ is a θ-multiresolution analysis of $L^2(\mathbb{R})$, i.e.,*

(i) *for any $j \in \mathbb{Z}$, $V_j^{(s)}(\Lambda)$ is a closed subspace of $L^2(\mathbb{R})$;*
(ii) $\cdots \subset V_{-1}^{(s)}(\Lambda) \subset V_0^{(s)}(\Lambda) \subset V_1^{(s)}(\Lambda) \subset \cdots$;
(iii) $\bigcup_{j\in\mathbb{Z}} V_j^{(s)}(\Lambda)$ *is dense in* $L^2(\mathbb{R})$;
(iv) $\bigcap_{j\in\mathbb{Z}} V_j^{(s)}(\Lambda) = \{0\}$;
(v) $f(x) \in V_j^{(s)}(\Lambda)$ *if and only if* $f(\theta^{-j}x) \in V_0^{(s)}(\Lambda)$;
(vi) *there exists a finite number of functions $\phi_{\bar{n}}(x) \in V_0^{(s)}$, called scaling functions, such that $\{\phi_{\bar{n}}(x-\lambda_k)\}_{k\in\bar{n},\, 0\le n\le q}$ is a Riesz basis in $V_0^{(s)}$.*

The proof is straightforward from definitions and Proposition 1.3.

As a consequence of the above statements, we have the important result obtained by Bernuau [11], [12]:

Theorem 1.5. *Let $\Lambda = \{\lambda_n \mid n \in \mathbb{Z}\} \subset \mathbb{R}$ be a Delone set of finite local complexity, self-similar with factor $\theta > 1$. Let us denote the elements of $\theta^{-1}\Lambda$ by $\kappa_n = \theta^{-1}\lambda_n$, $n \in \mathbb{Z}$. Then for all $s > 1$ there exists a Riesz basis of $L^2(\mathbb{R})$ of the form*

$$\{\theta^{j/2}\psi_{\kappa_n}^{(s)}(\theta^j x - \kappa_n),\ \kappa_n \in \theta^{-1}\Lambda,\ \kappa_{n+1} \notin \Lambda,\ j \in \mathbb{Z}\},$$

where $\{\psi_{\kappa_n}^{(s)}\}$ is a finite set of compactly supported functions of order C^{s-2}.

Such a result offers the possibility of constructing explicit (spline) wavelet basis for a large class of self-similar Delone set of finite local complexity, like in particular those ones obtained through the cut-and-project method. It was done for some very specific cases in [4]. Nevertheless, we think that a systematic study of the properties of (not necessarily self-similar) cut-and-project sets is still lacking.

The aim of the present paper is to fill this gap. Its content is devoted to the study of one-dimensional cut-and-project sets built in a rather generic way, without supposing a priori any algebraic nature for the respective slopes of D_1 and D_2 and by using the freedom of making the window Ω vary in a continuous way. Hence, our results can be viewed as paving the way to further investigations concerning explicit wavelet constructions for arbitrary cut-and-project sets. Indeed, what we learn from previous works is the importance of knowing in a precise way the environment of each point in the Delone set, i.e. its complexity, in order to build the scaling functions and the associated wavelets. Moreover, we should not underestimate the structural importance of the nested sequence (1.3) in five of its features, namely

(i) the way in which the new points (i.e. the *details*) intertwine the old ones at each step $\Lambda_j \subset \Lambda_{j+1}$ of the increasing sequence (1.3), or, equivalently, the characteristics of the detail sets $\Lambda_{j+1} \setminus \Lambda_j$;
(ii) the way in which the sequence behaves at the limit $j \to \infty$, i.e. the structure of its inductive limit which should be dense in $\mathbb{R}$;
(iii) the specific advantages brought by the self-similarity hypothesis;
(iv) related to (i) and (iii), the specific advantages brought by the *substitivity* hypothesis, a notion which is introduced in (6.2);
(v) the relevance of a specific choice of a nested sequence with regard to the domain of application of the corresponding wavelet analysis.

Note the crucial importance of the first point in connection with the so-called *scaling* or *refinement* equations which couple with the inclusion $V_0 \subset V_1 = V_0 \oplus W_0$. Suppose there exist spline bases $(\varphi_{0,k})_{k\in\mathcal{I}_0\subset\mathbb{Z}}$, $(\varphi_{1,k})_{k\in\mathcal{I}_1\subset\mathbb{Z}}$ for subspaces V_0 and V_1 respectively, and a wavelet basis $(\psi_{0,k})_{k\in\mathcal{J}_0\subset\mathbb{Z}}$ for W_0. Then the following hilbertian decomposition should hold:

$$\varphi_{1,j}(x) = \sum_{k\in\mathcal{I}_0} c_{1,jk}\varphi_{0,k}(x) + \sum_{k\in\mathcal{J}_0} d_{1,jk}\psi_{0,k}(x). \tag{1.4}$$

The way the *tendency* coefficients $c_{1,jk}$ and the *detail* coefficients $d_{1,jk}$ behave for large k is a crucial question in wavelet analysis, and so the way this question depends on the set inclusion $\Lambda_0 \subset \Lambda_1$ deserves special attention.

In consequence, we have organized the paper by following the hierarchy of questions i)–v), and furthermore including in the scheme other aspects of possible interest, like some considerations on numeration systems related to cut-and-project scheme. Section 2 is devoted to the geometrical aspects of those point sets in the line issued from the square lattice in the plane through "cut" and "projection": definition, study of distances between adjacent points, properties of invariance or covariance under the group $\mathrm{SL}(2,\mathbb{Z}) \times \{-1, 1\}$ acting on the square lattice. The material presented there is not specifically new. However it represents an original overview, in which one focuses on the universality of many features of these cut-and-project sets, independently of specific algebraic or substitutional characteristics.

In Section 3 we examine the combinatorial aspects of the cut-and-project sequences from the point of view of the theory of languages: subword complexity, Rauzy graphs, and occurrence of specific classes of finite words (or *factors*) in the bidirectional word biunivocally associated to cut-and-project sequences. One can find there original results (Propositions 3.8, 3.11, and 3.16). The important case of sturmian words is also considered and we establish three properties, 3.18, 3.19, 3.20, describing their factors.

Self-similar cut-and-project sets are the object of Section 4 in which we solve (Theorem 4.1) the precise relation between self-similarity and algebraic properties of the irrational scaling factor(s) on one hand and the irrational numbers involved in the cut-and-project scheme on the other hand.

In Section 5 we revisit the important notion of β-integers, i.e. those real numbers which do not have "β-fractional" part when expanded in "basis" $\beta > 1$, in the light of their possible or not relation with cut and projection, and we give a necessary and sufficient condition for β (Proposition 5.2) under which the positive β-integers coincide with the positive part of a cut-and-project set.

Another original part of the paper is found in Section 6. It is well known that infinite words associated to many cut-and-project sets present the so-called substitution invariance, and this property can be crucial for understanding or even for creating the relation $\Lambda_j \to \Lambda_{j+1}$ in a multiresolution sequence of sets and the scaling equations (1.4) issued from the companion inclusion $V_j \subset V_{j+1}$. Now, even though the substitution invariance is absent for a given bidirectional infinite word $u = \ldots u_{-2}u_{-1}|u_0u_1u_2\ldots$, where u_k belongs to some alphabet $\mathcal{A}$, there exist cases in which it could be "hidden" behind the weaker notion of *substitutivity*, which means that there exists another infinite word $v = \ldots v_{-2}v_{-1}|v_0v_1v_2\ldots$ over an alphabet $\mathcal{B}$ which has substitution invariance and a letter projection $\psi\colon \mathcal{B} \to \mathcal{A}$ such that $u = \ldots u_{-2}u_{-1}|u_0u_1u_2\ldots = \ldots \psi(v_{-2})\psi(v_{-1})|\psi(v_0)\psi(v_1)\psi(v_2)\ldots$.

With regards to this property, we shall give in Section 6 an algorithm (Theorem 6.6) allowing to "pull back", a given word u pertaining to the algebraic cut-and-project scheme to the word v mentioned in the above. In order to illustrate this result, the algorithm is carried out on an example of cut-and-project set defined by the algebraic (Sturm) number $1/\sqrt{2}$.

Eventually, we shall give in the conclusion some hints about possible applications of our results, mainly in direction of wavelet constructions, of mathematical diffraction, and of design of aperiodic pseudo-random number generators.

2 Cut-and-project sequences

In this section we define cut-and-project sequences arising by a projection of a 2-dimensional lattice. We also describe their basic properties, including the invariance under certain transformations. We further show that cut-and-project sequences are geometric representations of a three or two interval exchange. Codings of two interval exchanges are in one-to-one correspondence with mechanical words $\underline{s}_{\alpha,\beta}$, $\overline{s}_{\alpha,\beta}$, (see definition by (2.12) and (2.13)), which are in fact sturmian words (Definition 3.2).

2.1 Definition and properties

The construction of a cut-and-project sequence starts with a choice of a 2-dimensional lattice L and two straight lines D_1, D_2. One of the lines plays the role of the space onto which the lattice L is projected, the other line determines the direction of the projection. If A is an arbitrary non-singular linear map on $\mathbb{R}^2$, then the cut-and-project sequence constructed using a lattice AL and straight lines AD_1, AD_2 is the same as the cut-and-project sequence constructed using a lattice L and straight lines D_1, D_2. Therefore it is not necessary to consider general L and D_1, D_2. Some authors allow arbitrary lattice L and for D_1, D_2 take mutually orthogonal straight lines. Others, including us, prefer to fix the lattice $\mathbb{Z}^2$ and consider arbitrary straight lines D_1, D_2.

Let us take two distinct irrational numbers ε, η and let us consider straight lines $D_1 \colon y = \varepsilon x$, $D_2 \colon y = \eta x$. If we choose vectors

$$\vec{x}_1 = \frac{1}{\varepsilon - \eta}(1, \varepsilon) \quad \text{and} \quad \vec{x}_2 = \frac{1}{\eta - \varepsilon}(1, \eta)$$

in the subspaces D_1, D_2 of $\mathbb{R}^2$, then for every lattice point $(a, b) \in \mathbb{Z}^2$ we have

$$(a, b) = (b - a\eta)\vec{x}_1 + (b - a\varepsilon)\vec{x}_2.$$

Obviously, the projection of $\mathbb{Z}^2$ on D_1 along D_2 is the set

$$\mathbb{Z}[\eta]\vec{x}_1,$$

where $\mathbb{Z}[\eta]$ is the abelian group

$$\mathbb{Z}[\eta] := \{a + b\varepsilon \mid a, b \in \mathbb{Z}\}.$$

Similarly, the projection of $\mathbb{Z}^2$ on D_2 along D_1 is the set $\mathbb{Z}[\varepsilon]\vec{x}_2$. Since the numbers ε, η are irrational, the mappings $(a, b) \mapsto a + b\eta$, $(a, b) \mapsto a + b\varepsilon$ are bijections

between $\mathbb{Z}^2$ and $\mathbb{Z}[\eta]$, resp. $\mathbb{Z}^2$ and $\mathbb{Z}[\varepsilon]$. Therefore there exists also a bijection

$$\star\colon \mathbb{Z}[\eta] \to \mathbb{Z}[\varepsilon]$$

defined by the prescription

$$x = a + b\eta \mapsto x^\star = a + b\varepsilon,$$

which is called the star map. Directly from its definition we obtain

$$(x + y)^\star = x^\star + y^\star \quad \text{for every } x, y \in \mathbb{Z}[\eta].$$

Let us now introduce the definition of cut-and-project sets. It is easy to observe that, as a consequence of irrationality of ε and η, the sets $\mathbb{Z}[\varepsilon]$, $\mathbb{Z}[\eta]$ are dense in $\mathbb{R}$. However, if instead of all the lattice $\mathbb{Z}^2$, we project only those points in $\mathbb{Z}^2$ that belong to a chosen strip parallel to D_1, the resulting set in D_1 has no limit points. The width and position of the projected strip is determined by an interval in D_2. Formally, we have the following definition.

Definition 2.1. Let ε, η be distinct irrational numbers and let Ω be a bounded interval. The set

$$\Sigma_{\varepsilon,\eta}(\Omega) = \{a + b\eta \mid a, b \in \mathbb{Z},\ a + b\varepsilon \in \Omega\} = \{x \in \mathbb{Z}[\eta] \mid x^\star \in \Omega\}$$

is called a cut-and-project set, or C&P set. The interval Ω is called the acceptance window of $\Sigma_{\varepsilon,\eta}(\Omega)$.

The above definition is a special case of the very general 'model sets'. Important contributions to the study of model sets as mathematical models of quasicrystals are due to [36], [38], [49], [50], [52], [51].

Let us mention some of the properties of cut-and-project sequences that follow directly from the definition or were derived by cited authors.

Remark 2.2. (1) Trivially from the definition we have

$$\Sigma_{\varepsilon,\eta}(\Omega_1) \subset \Sigma_{\varepsilon,\eta}(\Omega_2) \quad \text{for } \Omega_1 \subset \Omega_2.$$

More generally, if $(\Omega_i)_{i\in\mathbb{Z}}$ is a sequence of nested bounded intervals such that

$$\cdots \subset \Omega_{i-1} \subset \Omega_i \subset \Omega_{i+1} \subset \cdots \quad \text{and} \quad \bigcup_{i\in\mathbb{Z}} \Omega_i = \mathbb{R},$$

then

$$\cdots \subset \Sigma_{\varepsilon,\eta}(\Omega_{i-1}) \subset \Sigma_{\varepsilon,\eta}(\Omega_i) \subset \Sigma_{\varepsilon,\eta}(\Omega_{i+1}) \subset \cdots$$

and

$$\bigcup_{i\in\mathbb{Z}} \Sigma_{\varepsilon,\eta}(\Omega_i) = \mathbb{Z}[\eta].$$

(2) Since $\mathbb{Z}[\eta]$ and $\mathbb{Z}[\varepsilon]$ are additive groups, the C&P sequence satisfies

$$x + \Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{\varepsilon,\eta}(\Omega + x^\star) \quad \text{for every } x \in \mathbb{Z}[\eta].$$

This property further implies that $\Sigma_{\varepsilon,\eta}(\Omega)$ is not invariant under any translation, i.e. is aperiodic.

(3) Any model set is Delone, see [50]. In our one-dimensional case it implies that there exist a sequence $(x_n)_{n\in\mathbb{Z}}$ and two positive numbers r_1, r_2 such that $r_1 < x_{n+1} - x_n < r_2$ for all $n \in \mathbb{Z}$, and $\Sigma_{\varepsilon,\eta}(\Omega) = \{x_n \mid n \in \mathbb{Z}\}$.

(4) The density of points of $\Sigma_{\varepsilon,\eta}(\Omega)$, defined as

$$\varrho(\Sigma_{\varepsilon,\eta}(\Omega)) := \lim_{N\to+\infty} \frac{\#\left([-N, N] \cap \Sigma_{\varepsilon,\eta}(\Omega)\right)}{2N+1},$$

is proportional to the length of the interval Ω; see [51].

(5) Since Ω is an interval, it is easy to see that there exists a finite set $F^\star \subset \mathbb{Z}[\varepsilon]$ such that $\Omega - \Omega \subset \Omega + F^\star$. Hence there also exists a finite set F such that

$$\Sigma_{\varepsilon,\eta}(\Omega) - \Sigma_{\varepsilon,\eta}(\Omega) \subset \Sigma_{\varepsilon,\eta}(\Omega) + F.$$

Thus $\Sigma_{\varepsilon,\eta}(\Omega)$ satisfies the so-called Meyer property. In fact, every model set $\Lambda \subset \mathbb{R}^n$ is a Meyer set, i.e. is Delone and satisfies $\Lambda - \Lambda \subset \Lambda + F$ for a finite set F, cf. [50]. Note that in this sense a cut-and-project set is a generalization of a lattice, because a lattice satisfies the above property with $F = \{0\}$.

(6) Since $\Sigma_{\varepsilon,\eta}(\Omega)$ is a Meyer set, it is of finite local complexity, i.e., it has only a finite number of local configurations of a fixed size [36]. More precisely, for $\varrho > 0$ we define the ϱ-neighbourhood of a point $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ as

$$N_\varrho(x) = \{y \in \Sigma_{\varepsilon,\eta}(\Omega) \mid |x - y| < \varrho\}.$$

The family of ϱ-neighbourhoods $\{N_\varrho(x) - x \mid x \in \Sigma_{\varepsilon,\eta}(\Omega)\}$ is finite for any positive ϱ. In particular, there is only finitely many distances between adjacent points of any C&P sequence, i.e., the set $\{x_{n+1} - x_n \mid n \in \mathbb{Z}\}$ is finite.

(7) The boundary of the acceptance interval Ω influences the structure of the C&P sequence only trivially. The sets $\Sigma_{\varepsilon,\eta}[c, c+\ell)$, $\Sigma_{\varepsilon,\eta}[c, c+\ell]$, $\Sigma_{\varepsilon,\eta}(c, c+\ell)$, and $\Sigma_{\varepsilon,\eta}(c, c+\ell]$ differ at most in two points. If $c, c+\ell \notin \mathbb{Z}[\varepsilon]$, then all these sets coincide.

(8) If the acceptance window Ω is chosen to be a semi-closed interval, then the number and shape of ϱ-neighbourhoods of a C&P sequence does not depend on the position Ω, but only on its length $|\Omega|$.

(9) The set $\Sigma_{\varepsilon,\eta}(\Omega)$, where $\Omega = [c, c+\ell)$ or $\Omega = (c, c+\ell]$ contains every finite configuration infinitely many times. More precisely, for every $\varrho > 0$ and every $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ there exist infinitely many points $y \in \Sigma_{\varepsilon,\eta}(\Omega)$ such that

$$N_\varrho(x) - x = N_\varrho(y) - y.$$

We say that such C&P sequences are repetitive.

2.2 Distances

As it was mentioned in (6) of Remark 2.2, any cut-and-project sequence has only a finite number of distances between adjacent points. It turns out that the number of distances does not exceed 3. We quote the result of [30], which is a generalization of the famous 3-distance theorem, and provide algorithms for determining the distances for any particular acceptance interval.

Theorem 2.3. *Let Ω be a semi-closed interval. For every $\Sigma_{\varepsilon,\eta}(\Omega)$ there exist positive numbers Δ_1, $\Delta_2 \in \mathbb{Z}[\eta]$ such that the distances between adjacent points in $\Sigma_{\varepsilon,\eta}(\Omega)$ take values in $\{\Delta_1, \Delta_2, \Delta_1+\Delta_2\}$. The numbers Δ_1, Δ_2 depend only on the parameters ε, η and on the length $|\Omega|$ of the interval Ω. They are linearly independent over $\mathbb{Q}$ and satisfy $\Delta_1^\star > 0$, $\Delta_2^\star < 0$, and $\Delta_1^\star - \Delta_2^\star \geq |\Omega|$.*

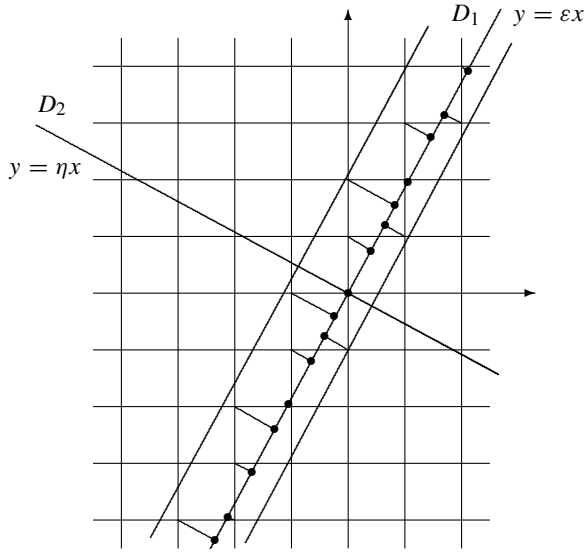


Figure 1. Construction of a cut-and-project sequence and assignment of the infinite word which codes the order of distances in the cut-and-project sequence.

More precisely, every C&P sequence $\Sigma_{\varepsilon,\eta}(\Omega) = \{x_n \mid n \in \mathbb{Z}\}$ has always two or three type of distances between adjacent points, namely

$$\{x_{n+1} - x_n \mid n \in \mathbb{Z}\} = \begin{cases} \{\Delta_1, \Delta_2, \Delta_1 + \Delta_2\} & \text{if } \Delta_1^\star - \Delta_2^\star > |\Omega|, \\ \{\Delta_1, \Delta_2\} & \text{if } \Delta_1^\star - \Delta_2^\star = |\Omega|. \end{cases} \tag{2.1}$$

Therefore one can naturally assign to it a binary or ternary bidirectional infinite word $u_{\varepsilon,\eta}(\Omega) = (u_n)_{n\in\mathbb{Z}}$, for example in the alphabet $\{A, B, C\}$, by

$$u_n = \begin{cases} A & \text{if } x_{n+1} - x_n = \Delta_1, \\ B & \text{if } x_{n+1} - x_n = \Delta_1 + \Delta_2, \\ C & \text{if } x_{n+1} - x_n = \Delta_2. \end{cases} \tag{2.2}$$

An example of construction of a cut-and-project sequence together with the assignment of the infinite word is shown in Figure 1.

The successor of a point x in the C&P sequence is determined using its star-map image $x^\star$. If $\Omega = [c, c + \ell)$, then the inequality $\Delta_1^\star - \Delta_2^\star \geq \ell$ from Theorem 2.5 ensures that the nearest right neighbour of the point x in $\Sigma_{\varepsilon,\eta}(\Omega)$ is equal to $x + \Delta_1$ if $x^\star \in [c, c + \ell - \Delta_1^\star)$, to $x + \Delta_2$ if $x^\star \in [c - \Delta_2^\star, c + \ell)$, or to $x + \Delta_1 + \Delta_2$ if $x^\star \in [c+\ell-\Delta_1^\star, c-\Delta_2^\star)$. Thus we can define a piecewise linear map $f : [c, c+\ell) \to [c, c + \ell)$ which satisfies $f(x^\star) = y^\star$ if y is the nearest right neighbour of x. This mapping plays an important role in our considerations.

Definition 2.4. Let $c \in \mathbb{R}$, $\ell > 0$. The stepping function of the interval $[c, c + \ell)$ is a mapping $f : [c, c + \ell) \to [c, c + \ell)$ defined by

$$f(y) = \begin{cases} y + \Delta_1^\star & \text{if } y \in [c, c + \ell - \Delta_1^\star), \\ y + \Delta_1^\star + \Delta_2^\star & \text{if } y \in [c + \ell - \Delta_1^\star, c - \Delta_2^\star), \\ y + \Delta_2^\star & \text{if } y \in [c - \Delta_2^\star, c + \ell). \end{cases}$$

The graph of the map f is illustrated on Figure 2.

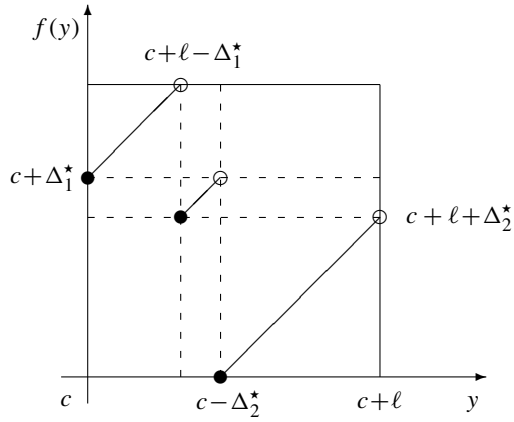
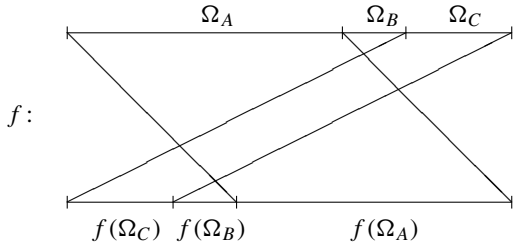


Figure 2. Stepping function f associated to $\Sigma_{\varepsilon,\eta}[c, c+\ell)$. If the distances between neighbours in the C&P set $\Sigma_{\varepsilon,\eta}[c, c+\ell)$ take only two values Δ_1, Δ_2, i.e. $\ell = \Delta_1^\star - \Delta_2^\star$, then the discontinuity points $c + \ell - \Delta_1^\star$ and $c - \Delta_2^\star$ coincide.

The stepping function f of Figure 2 has been studied in the field of dynamical systems under the name of three interval exchange (in case that f has two discontinuity

points) or two interval exchange (if f has only one discontinuity point). In the former situation, the two discontinuity points divide the acceptance interval Ω into three disjoint intervals, say Ω_A, Ω_B, Ω_C, from left to right. The image $f(\Omega) = \Omega$ is again divided into three disjoint intervals $f(\Omega_C)$, $f(\Omega_B)$, $f(\Omega_A)$, in the order from left to right. Therefore one sometimes uses the graphical notation shown in the following scheme.



The relation of three interval exchange to simultaneous approximation of a pair of irrational numbers is treated in [1], [26], [57]. In the theory of symbolic dynamical systems one studies the orbit of a point $x \in \Omega$, under the mapping f, i.e. the sequence $(f^{(n)})_{n\in\mathbb{N}_0}$.[1] Therefore C&P sequences can be viewed as geometric representations of three interval exchange transformations.

Changing continuously the length ℓ of the acceptance interval $\Omega = [c, c + \ell)$ causes discrete changes of the triplet of distances $(\Delta_1, \Delta_2, \Delta_1 + \Delta_2)$. Recall that the triplet does not depend on c, therefore we consider $c = 0$. Let $\Sigma_{\varepsilon,\eta}[0, \ell)$ be a C&P sequence with three distances between its neighbours, and let $\Delta_1^\star > 0$, $\Delta_2^\star < 0$, $\Delta_1^\star + \Delta_2^\star$, be the star map images of these distances. According to (2.1), we must have $\ell < \Delta_1^\star - \Delta_2^\star$. Growing ℓ up to the value $\Delta_1^\star - \Delta_2^\star$ causes appearance of new points in the C&P sequence, which split the large distance $\Delta_1 + \Delta_2$ into two distances Δ_1 and Δ_2. When ℓ reaches the value $\Delta_1^\star - \Delta_2^\star$, the large distance $\Delta_1 + \Delta_2$ disappears completely.

On the other hand, diminishing the length ℓ of the acceptance interval causes that the frequency of the distance $\Delta_1 + \Delta_2$ grows to the detriment of occurrences of the distances Δ_1 and Δ_2. This happens until ℓ reaches a certain limit value for which one of the distances Δ_1 or Δ_2 disappears.

Starting from a given initial value ℓ_0, for which the set $\Sigma_{\varepsilon,\eta}[0, \ell_0)$ has two distances between adjacent points, we can determine by recurrence the increasing sequence of lengths ℓ_n, $n \in \mathbb{Z}$, of the acceptance windows for which $\Sigma_{\varepsilon,\eta}[0, \ell_n)$ has only two distances. The initial value ℓ_0 is determined below (Remark 2.6).

Let $\Delta_{n1}^\star > 0$ and $\Delta_{n2}^\star < 0$ be the star images of distances occurring in the sequence $\Sigma_{\varepsilon,\eta}[0, \ell_n)$, i.e., according to (2.1), $\ell_n = \Delta_{n1}^\star - \Delta_{n2}^\star$.

[1] Note that we use the notation $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$.

If $\Delta^\star_{n1} + \Delta^\star_{n2} > 0$ then

$$\ell_{n-1} := \Delta^\star_{n1}, \qquad \Delta^\star_{(n-1)1} := \Delta^\star_{n1} + \Delta^\star_{n2}, \quad \Delta^\star_{(n-1)2} := \Delta^\star_{n2}. \tag{2.3}$$

If $\Delta^\star_{n1} + \Delta^\star_{n2} < 0$ then

$$\ell_{n-1} := -\Delta^\star_{n2}, \quad \Delta^\star_{(n-1)1} := \Delta^\star_{n1}, \qquad \Delta^\star_{(n-1)2} := \Delta^\star_{n1} + \Delta^\star_{n2}.$$

Similarly, the algorithm which determines the triple ℓ_{n+1}, $\Delta_{(n+1)1}$, $\Delta_{(n+1)2}$ from the triple ℓ_n, Δ_{n1}, Δ_{n2} has the inverse form:

If $\Delta_{n1} > \Delta_{n2}$ then

$$\ell_{n+1} := \Delta^\star_{n1} - 2\Delta^\star_{n2}, \quad \Delta^\star_{(n+1)1} := \Delta^\star_{n1} - \Delta^\star_{n2}, \quad \Delta^\star_{(n+1)2} := \Delta^\star_{n2}. \tag{2.4}$$

If $\Delta_{n1} < \Delta_{n2}$ then

$$\ell_{n+1} := 2\Delta^\star_{n1} - \Delta^\star_{n2}, \quad \Delta^\star_{(n+1)1} := \Delta^\star_{n1}, \qquad \Delta^\star_{(n+1)2} := \Delta^\star_{n2} - \Delta^\star_{n1}.$$

From this algorithm it can be seen that the C&P sequences $\Sigma_{\varepsilon,\eta}[0, \ell_n)$ and $\Sigma_{\varepsilon,\eta}[0, \ell_{n-1})$ have exactly one type of distances in common. It is the shorter one among $\Delta_{(n-1)1}$, $\Delta_{(n-1)2}$. Moreover, the distances in $\Sigma_{\varepsilon,\eta}[0, \ell)$, for $\ell_{n-1} < \ell < \ell_n$, are three, and they are given by the union of the sets of distances for $\Sigma_{\varepsilon,\eta}[0, \ell_n)$ and $\Sigma_{\varepsilon,\eta}[0, \ell_{n-1})$.

2.3 Transformations

We would like to identify those parameters ε, η, Ω which provide essentially the same cut-and-project sequences. For example, we have

$$a + b\eta + \Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{\varepsilon,\eta}(\Omega + a + b\varepsilon) \quad \text{for } a, b \in \mathbb{Z}. \tag{2.5}$$

Such translation of the C&P sequence corresponds to a translation of the lattice $\mathbb{Z}^2$. The group of all linear transformations of the lattice $\mathbb{Z}^2$ onto itself is

$$G = \{\mathbb{A} \in M_2(\mathbb{Z}) \mid \det \mathbb{A} = \pm 1\}.$$

Consider the matrix $\mathbb{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. For arbitrary irrational numbers ε, η and arbitrary interval Ω it holds that

$$\begin{aligned}
\Sigma_{\varepsilon,\eta}(\Omega) &= \{p + q\eta \mid p, q \in \mathbb{Z},\ pq\varepsilon \in \Omega\} \\
&= \{(1, \eta)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \mid p, q \in \mathbb{Z},\ (1, \varepsilon)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \Omega\} \\
&= \{(1, \eta)\mathbb{A}\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \mid p, q \in \mathbb{Z},\ (1, \varepsilon)\mathbb{A}\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \Omega\} \\
&= \{(a + c\eta, b + d\eta)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \mid p, q \in \mathbb{Z},\ (a + c\varepsilon, b + d\varepsilon)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \Omega\}
\end{aligned}$$

$$= (a + c\eta)\big\{\big(1, \tfrac{b+d\eta}{a+c\eta}\big)\binom{p}{q} \mid p, q \in \mathbb{Z},\ \big(1, \tfrac{b+d\varepsilon}{a+c\varepsilon}\big)\binom{p}{q} \in \tfrac{1}{a+c\varepsilon}\Omega\big\}$$
$$= (a + c\eta)\,\Sigma_{\frac{b+d\varepsilon}{a+c\varepsilon}, \frac{b+d\eta}{a+c\eta}}\big(\tfrac{1}{a+c\varepsilon}\Omega\big).$$

Let us study the consequences of the above relation if we choose for the matrix $\mathbb{A}$ one of the three generators $\mathbb{A}_1 = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$, $\mathbb{A}_2 = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$, $\mathbb{A}_3 = \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$ of the group G,

$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{1+\varepsilon,\ 1+\eta}(\Omega), \tag{2.6}$$
$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{-\varepsilon,-\eta}(-\Omega), \tag{2.7}$$
$$\Sigma_{\varepsilon,\eta}(\Omega) = \eta\ \Sigma_{\frac{1}{\varepsilon}, \frac{1}{\eta}}(\tfrac{1}{\varepsilon}\Omega). \tag{2.8}$$

The mentioned transformations were used in [30] for the proof of the following theorem.

Theorem 2.5. *For every irrational numbers ε, η, $\varepsilon \neq \eta$ and every bounded interval Ω, there exist $\tilde{\varepsilon} \in (-1, 0)$, $\tilde{\eta} > 0$ and an interval $\tilde{\Omega}$, satisfying $\max(1+\tilde{\varepsilon}, -\tilde{\varepsilon}) < |\tilde{\Omega}| \leq 1$, such that*

$$\Sigma_{\varepsilon,\eta}(\Omega) = s\,\Sigma_{\tilde{\varepsilon},\tilde{\eta}}(\tilde{\Omega}) \quad \textit{for some}\ \ s \in \mathbb{R}.$$

Moreover, if $|\tilde{\Omega}| \neq 1$, then the distances between adjacent points in $\Sigma_{\tilde{\varepsilon},\tilde{\eta}}(\tilde{\Omega})$ are $\tilde{\eta}$, $1 + \tilde{\eta}$, and $1 + 2\tilde{\eta}$. The distances take only two values, $\tilde{\eta}$ and $1 + \tilde{\eta}$, if $|\tilde{\Omega}| = 1$.

According to the above theorem, every C&P sequence is geometrically similar to another C&P sequence whose parameters satisfy certain restricted conditions. In particular, without loss of generality we can consider $\varepsilon \in (-1, 0)$, $\eta > 0$ and the length of the acceptance interval Ω in the range $(\max(1 + \varepsilon, -\varepsilon), 1]$. If moreover we are interested only in the ordering of the distances in the C&P sequence and not in their actual lengths, i.e., we consider only the infinite word $u_{\varepsilon,\eta}(\Omega)$, we can choose any fixed $\eta > 0$. The words $u_{\varepsilon,\eta_1}(\Omega)$, $u_{\varepsilon,\eta_2}(\Omega)$ coincide for $\eta_1 \neq \eta_2$. Therefore choosing $\eta = -\frac{1}{\varepsilon}$, which corresponds to a cut-and-project scheme with orthogonal projection, causes no loss of generality when studying only combinatorial properties of C&P sequences. The choice of η however influences the geometry of the sequences, such as existence of self-similarity factor, etc. (cf. Section 4).

Remark 2.6. Note that according to Theorem 2.5, the length $|\Omega|$ of the interval Ω being equal to 1 is the only case among $(\max(1 + \tilde{\varepsilon}, -\tilde{\varepsilon}), 1]$ for which the C&P set has only two distances between neighbours. We shall thus take it as the initial case for the algorithm given in (2.3), (2.4). We have

$$\ell_0 = 1, \quad \Delta^{\star}_{01} = 1 + \varepsilon, \quad \Delta^{\star}_{02} = \varepsilon. \tag{2.9}$$

Example 2.7. As an example, let us study the case $|\Omega| = 1$, which gives a C&P sequence with two distances between adjacent points. Set $\alpha = -\varepsilon \in (0, 1)$ and put $\Omega = (\beta - 1, \beta]$ for some $\beta \in \mathbb{R}$ as the acceptance window. Since the condition $a + b\varepsilon \in \Omega$ rewrites as $\beta - 1 < a - b\alpha \leq \beta$, we obtain $a = \lfloor b\alpha + \beta \rfloor$ and the C&P

sequence is of the form

$$\Sigma_{-\alpha,\eta}(\beta-1,\beta] = \{\lfloor b\alpha+\beta\rfloor + b\eta \mid b\in\mathbb{Z}\}. \tag{2.10}$$

Since $\alpha, \eta > 0$, the sequence $x_n := \lfloor n\alpha+\beta\rfloor + n\eta$ is strictly increasing and thus the distances between adjacent points of the C&P set $\Sigma_{-\alpha,\eta}(\beta-1,\beta]$ are of the form

$$x_{n+1}-x_n = \eta + \lfloor (n+1)\alpha+\beta\rfloor - \lfloor n\alpha+\beta\rfloor = \begin{cases}\eta+1,\\ \eta.\end{cases} \tag{2.11}$$

From this expression it can be seen that the distances in the C&P sequence are arranged in the same order as 0's and 1's in the so-called lower and upper mechanical word. Recall that the lower mechanical word $\underline{s}_{\alpha,\beta}\colon \mathbb{Z}\to\{0,1\}$ is defined by the prescription

$$\underline{s}_{\alpha,\beta}(n) = \lfloor (n+1)\alpha+\beta\rfloor - \lfloor n\alpha+\beta\rfloor, \tag{2.12}$$

where α is called the slope and β the intercept of the word $\underline{s}_{\alpha,\beta}$. Similarly, upper mechanical word $\bar{s}_{\alpha,\beta}\colon \mathbb{Z}\to\{0,1\}$ is defined by the prescription

$$\bar{s}_{\alpha,\beta}(n) = \lceil (n+1)\alpha+\beta\rceil - \lceil n\alpha+\beta\rceil. \tag{2.13}$$

The infinite word $u_{\varepsilon,\eta}(\beta-1,\beta]$ with parameters $\eta>0$, $\varepsilon=-\alpha$ is in fact the lower mechanical word $\underline{s}_{\alpha,\beta}$. Similarly, the choice $[\beta,\beta+1)$ for the acceptance window provides the upper mechanical word $\bar{s}_{\alpha,\beta}$. The mechanical words are in fact related to the well-known sturmian words; see Definition 3.2 and Remark 3.3.

3 Combinatorial properties of C&P sequences

Ordering of the distances in the C&P sequence $\Sigma_{\varepsilon,\eta}(\Omega)$ on the real line defines naturally an infinite binary or ternary word $u_{\varepsilon,\eta}(\Omega)$ (cf. equation (2.2)). In this section we describe some combinatorial properties of these infinite words. Some of the results derived here can be found in [27]. Nevertheless, the geometric approach to three interval exchange makes the proof simpler.

Obviously, geometrically similar C&P sequences correspond to the same infinite words. Therefore according to Theorem 2.5 we can consider only

$$\varepsilon\in(-1,0),\ \eta>0 \quad\text{and}\quad \Omega=[c,c+\ell),\ \text{where } \max(1+\varepsilon,-\varepsilon)<\ell\le 1. \tag{3.1}$$

In this case the stepping function has the form

$$f(y) = \begin{cases} y+1+\varepsilon & \text{if } y\in[c,c+\ell-1-\varepsilon) =: \Omega_A,\\ y+1+2\varepsilon & \text{if } y\in[c+\ell-1-\varepsilon,c-\varepsilon) =: \Omega_B,\\ y+\varepsilon & \text{if } y\in[c-\varepsilon,c+\ell) =: \Omega_C.\end{cases} \tag{3.2}$$

For simplicity, we denote the discontinuity points of the stepping function

$$\delta_1 := c+\ell-1-\varepsilon, \quad \delta_2 := c-\varepsilon.$$

As was already mentioned, the infinite word $u_{\varepsilon,\eta}(\Omega)$ is defined over a binary alphabet if and only if the length of the acceptance window is $\ell = 1$, because in that case $\delta_1 = \delta_2$ and thus $\Omega_B = \emptyset$. Otherwise the alphabet of $u_{\varepsilon,\eta}(\Omega)$ has three letters.

Let us recall some basic notions of combinatorics on words. An alphabet $\mathcal{A}$ is a finite set of symbols or letters. A finite concatenation w of letters is called a finite word. The set of all finite words (including the empty word ϵ) over the alphabet $\mathcal{A}$ is denoted by $\mathcal{A}^*$. The concatenation of n letters a is denoted by a^n. The length of a word w is the number of letters concatenated in w, it is denoted by $|w|$. One considers also one-directional infinite words

$$u = u_0u_1u_2u_3\dots$$

and bidirectional infinite words

$$u = \dots u_{-2}u_{-1}u_0u_1u_2\dots.$$

In relation to C&P sequences, mainly bidirectional infinite words are important. We denote the set of such words by $\mathcal{A}^{\mathbb{Z}}$. A word $w = w_0w_1\dots w_{k-1}$ is called a *factor* of a word $u \in \mathcal{A}^{\mathbb{Z}}$ if $w = u_iu_{i+1}\dots u_{i+k-1}$ for some i. Note that such i is called the *occurrence* of w in u. The set of factors of a word $u \in \mathcal{A}^{\mathbb{Z}}$ with the length n is denoted by

$$\mathcal{L}_n = \{u_iu_{i+1}\dots u_{i+n-1} \mid i \in \mathbb{Z}\}.$$

The set of all factors of the word u (the *language* of u) is denoted by

$$\mathcal{L} = \bigcup_{n\in\mathbb{N}} \mathcal{L}_n.$$

The number of different n-tuples that appear in the infinite word is given by the so-called complexity function; see for example [3].

Definition 3.1. The complexity of a word $u \in \mathcal{A}^{\mathbb{Z}}$ is a mapping $\mathcal{C}\colon \mathbb{N} \to \mathbb{N}$ such that

$$\mathcal{C}(n) = \#\{u_iu_{i+1}\cdots u_{i+n-1} \mid i \in \mathbb{Z}\} = \#\mathcal{L}_n.$$

Obviously, if u is an infinite word over a k-letter alphabet, then its complexity satisfies

$$1 \le \mathcal{C}(n) \le k^n \quad \text{for every } n \in \mathbb{N}.$$

It is known [53] that if there exists an $n \in \mathbb{N}$ such that $\mathcal{C}(n) \le n$, then the word u is periodic, i.e. of the form $u = \dots www\dots$ for a finite word w. An aperiodic word of minimal complexity thus satisfies $\mathcal{C}(n) = n + 1$ for all $n \in \mathbb{N}$. An example of such a word is the word $\dots 0001000\dots$ in the alphabet $\{0, 1\}$. The structure of such words is little interesting, since the occurrence of the letter 1 is singular. Obviously, they cannot be obtained by a cut-and-project scheme. In order to avoid such strange phenomena, we consider only those words which have reasonable density of their letters. The density of a letter a in the infinite word $u = \dots u_{-2}u_{-1}u_0u_1u_2\dots$ is

defined by

$$\varrho_a := \lim_{k\to\infty} \frac{\#\{i \in \mathbb{Z} \cap [-k,k] \mid u_i = a\}}{2k+1}$$

if the limit exists.

Definition 3.2. An infinite word $u \in \mathcal{A}^{\mathbb{Z}}$ is called sturmian if $\mathcal{C}(n) = n+1$ for all $n \in \mathbb{N}$ and the densities of its letters are irrational.

Such words have been extensively studied. We shall focus on them later in this section. Let us mention that the condition $\mathcal{C}(n) = n+1$ for all $n \in \mathbb{N}$ in the case of one-directional infinite words already implies irrationality of the densities of letters. Our notion of sturmian words follows [54], however, sturmian words are often considered only as one-directional. A survey of properties of one-directional sturmian words can be found in [41].

3.1 Complexity

For the determination of the factors in the infinite bidirectional word $u_{\varepsilon,\eta}(\Omega)$ associated with the C&P sequence $\Sigma_{\varepsilon,\eta}(\Omega)$ it is essential to study the stepping function f. Its properties imply that the word $w = w_0 w_1 \dots w_{k-1}$ in the alphabet $\mathcal{A} = \{A, B, C\}$ is a factor of $u_{\varepsilon,\eta}(\Omega)$ if and only if there exists an $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ such that

$$x^\star \in \Omega_{w_0},\ f(x^\star) \in \Omega_{w_1},\ \dots,\ f^{k-1}(x^\star) \in \Omega_{w_{k-1}},$$

where $\Omega_A, \Omega_B, \Omega_C$ are defined in (3.2). This means that

$$w = w_0 w_1 \dots w_{k-1} \in \mathcal{L}_k \iff \Omega_{w_0} \cap f^{-1}(\Omega_{w_1}) \cap \dots \cap f^{-(k-1)}(\Omega_{w_{k-1}}) \neq \emptyset.$$

In case that $w = w_0 w_1 \dots w_{k-1} \in \mathcal{L}_k$, we denote

$$\Omega_w := \Omega_{w_0} \cap f^{-1}(\Omega_{w_1}) \cap \dots \cap f^{-(k-1)}(\Omega_{w_{k-1}}).$$

Properties of the stepping function f imply that Ω_w is an interval, closed from the left, open from the right. Obviously, we have

$$\Omega = \bigcup_{w\in\mathcal{L}_k} \Omega_w,$$

where the union is disjoint. In order that points $x^\star, y^\star \in \Omega$ belong to different intervals $x^\star \in \Omega_{w^{(1)}}$, $y^\star \in \Omega_{w^{(2)}}$, where $w^{(1)} \neq w^{(2)}$, $w^{(1)}, w^{(2)} \in \mathcal{L}_k$, there must exist $i = 0, 1, \dots, k-1$ such that at least one discontinuity point of the function f lies between $f^i(x^\star)$ and $f^i(y^\star)$. Thus boundaries between intervals Ω_w for $w \in \mathcal{L}_k$ are all points z such that $f^i(z)$ is a discontinuity point of the function f, i.e. $f^i(z) \in \{\delta_1, \delta_2\}$. This implies that the number of different factors of the word $u_{\varepsilon,\eta}(\Omega)$ of length k is equal to the number of elements,

$$\#\mathcal{L}_k = \#\{c, \delta_1, f^{-1}(\delta_1), \dots, f^{-k+1}(\delta_1), \delta_2, f^{-1}(\delta_2), \dots, f^{-k+1}(\delta_2)\}. \qquad (3.3)$$

For determination of the cardinality of the set $\mathcal{L}_k$, i.e. complexity of the infinite word, we need to use two properties of the stepping function f, which follow from the irrationality of ε.

$$\begin{aligned} &1)\ \text{Let } x \in \Omega. \text{ Then } f^i(x) \neq x \text{ for all } i \in \mathbb{Z},\ i \neq 0. \\ &2)\ \text{Let } x, y \in \Omega. \text{ Then there is } i \in \mathbb{Z} \text{ such that } f^i(x) = y \text{ iff } x - y \in \mathbb{Z}[\varepsilon]. \end{aligned} \tag{3.4}$$

Since $c = f(\delta_2)$ the property 1) implies that $c, \delta_2, f^{-1}(\delta_2), \ldots, f^{-k+1}(\delta_2)$ are distinct. Thus

$$\mathcal{C}(k) = \#\mathcal{L}_k \geq k + 1 \quad \text{for all } \ k \in \mathbb{N}.$$

Equality holds only in the case that $\delta_1 = \delta_2$. Therefore for C&P sequences $u_{\varepsilon,\eta}(\Omega)$ with parameters ε, η, Ω satisfying (3.1) it holds that

$$\mathcal{C}(k) = k + 1 \iff |\Omega| = 1.$$

We have thus derived the following well-known fact.

Remark 3.3. Every mechanical word (2.12) or (2.13) is a sturmian word. The opposite is also true [22, 54].

The results about the complexity function of all C&P sequences are summarized in the following theorem.

Theorem 3.4 ([30]). *Let $\mathcal{C}$ be the complexity function of the infinite word $u_{\varepsilon,\eta}(\Omega)$ with $\Omega = [c, c + \ell)$, and let f be the corresponding stepping function.*

- *If $\ell \notin \mathbb{Z}[\varepsilon]$ then*
$$\mathcal{C}(n) = 2n + 1 \quad \textit{for } \ n \in \mathbb{N}.$$
- *If $\ell \in \mathbb{Z}[\varepsilon]$ then there exists a unique $n_0 \in \mathbb{N}_0$ such that*
$$\mathcal{C}(n) = \begin{cases} 2n + 1 & \textit{for } \ n \leq n_0, \\ n + n_0 + 1 & \textit{for } \ n > n_0. \end{cases}$$

Obviously, generic cut-and-project sequences have complexity $2n + 1$. In case that the length of the acceptance window is in $\mathbb{Z}[\varepsilon]$, the cut-and-project sequence has a specific property which is explained in the following remark.

Remark 3.5. Theorem 3.4 says that infinite words $u_{\varepsilon,\eta}(\Omega)$ with $|\Omega| \in \mathbb{Z}[\varepsilon]$ have complexity $\mathcal{C}(n) = n + const.$ for sufficiently large n. One-directional words with such complexity are called quasisturmian by Cassaigne in [21]. This author shows that such words have a sturmian structure, i.e., up to a finite prefix they are images under a morphism of a one-direction sturmian word. Following the same ideas, one can show that the bidirectional infinite word $u_{\varepsilon,\eta}(\Omega)$ with $|\Omega| \in \mathbb{Z}[\varepsilon]$ corresponding to a C&P sequence satisfies the following: there exists a sturmian word $v = \ldots v_{-2}v_{-1}|v_0v_1v_2\ldots \in \{0, 1\}^{\mathbb{Z}}$ and finite words $W_0, W_1 \in \{A, B, C\}^*$ such that

$$u_{\varepsilon,\eta}(\Omega) = \ldots W_{v_{-2}} W_{v_{-1}} | W_{v_0} W_{v_1} W_{v_2} \ldots,$$

i.e., $u_{\varepsilon,\eta}(\Omega)$ can be obtained by concatenation of words W_0, W_1 in the order of 0's and 1's in the sturmian word v.

Example 3.6. Consider $\varepsilon = -\frac{1}{\tau}$ and $\eta = \tau$, where $\tau = \frac{1}{2}(1+\sqrt{5})$; see Introduction. For the acceptance window choose $\Omega_1 = \big[-\frac{7}{\tau}+4, -\frac{17}{\tau}+11\big)$, $\Omega_2 = \Omega_1 + 2 - \frac{3}{\tau}$. Since the length of the acceptance windows $\ell = |\Omega_1| = |\Omega_2|$ satisfies $\ell = 7 - \frac{10}{\tau} \in (\max(1+\varepsilon, \varepsilon), 1]$, according to Theorem 2.5, the distances between adjacent points in the C&P sequences $\Sigma_{\varepsilon,\eta}(\Omega_1)$, $\Sigma_{\varepsilon,\eta}(\Omega_2)$ are: $1+\tau$, coded by the letter A; $1+2\tau$, coded by the letter B; and τ, coded by the letter C.

Using (2.5) the sequences $\Sigma_{\varepsilon,\eta}(\Omega_1)$, $\Sigma_{\varepsilon,\eta}(\Omega_2)$ are the same, up to a shift by $2+3\tau$. Both of them contain 0, since $0 \in \Omega_1, 0 \in \Omega_2$. Figure 3 shows a segment of the infinite words coding these sequences, where we mark the point 0 in both of them.

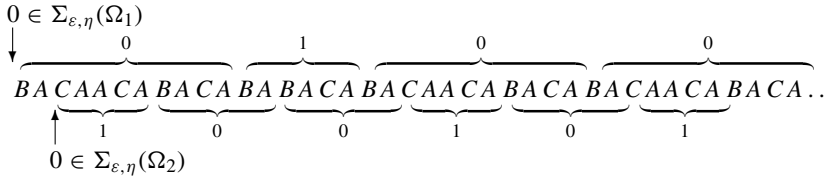


Figure 3. Block structure of quasisturmian words.

It can be shown that the word $u_{\varepsilon,\eta}(\Omega_1)$ can be obtained from the upper mechanical sequence $\bar{s}_{\frac{1}{\tau},-\frac{1}{\tau^2}}$ by substituting for 0 the word $w_0 = BACAACA$ and for 1 the word $w_1 = BABACA$. Similarly, the infinite word $u_{\varepsilon,\eta}(\Omega_2)$ can be obtained from the upper mechanical sequence $\bar{s}_{\frac{1}{\tau},-\frac{1}{\tau}}$ by substituting for 0 the word $w_0 = BACABA$ and for 1 the word $w_1 = CAACA$.

3.2 Properties of the language

Let us study the language of the infinite word $u_{\varepsilon,\eta}(\Omega)$. Since the stepping function f_Ω corresponding to the C&P sequence $\Sigma_{\varepsilon,\eta}(\Omega)$ and the stepping function $f_{\Omega+z}$ corresponding to the C&P sequence $\Sigma_{\varepsilon,\eta}(\Omega+z)$ satisfy the relation

$$f_{\Omega+z}(x) = z + f_\Omega(x-z),$$

the language $\mathcal{L}$ and the complexity $\mathcal{C}$ of a C&P sequence depend only on the length of the acceptance interval Ω and not on its position.

First we determine the density of a given factor. Recall that the density of a particular factor w in the infinite bidirectional word $u = \ldots u_{-2}u_{-1}u_0u_1u_2\ldots$ is defined by

$$\varrho_w := \lim_{k\to\infty} \frac{\#\{i \in \mathbb{Z} \cap [-k,k] \mid u_i u_{i+1}\ldots u_{i+n-1} = w\}}{2k+1} \tag{3.5}$$

if the limit exists. Elements $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ such that the word corresponding to the k-tuple of the right neighbours of x is $w = w_0 w_1 \dots w_{k-1} \in \mathcal{L}_k$ satisfy $x^\star \in \Omega_w$. Therefore the occurrences of the factor w in the infinite word $u_{\varepsilon,\eta}(\Omega)$ are given by the set

$$\{x \in \mathbb{Z}[\eta] \mid x^\star \in \Omega_w\}.$$

Since Ω_w is a semi-closed interval $\Omega_w \subset \Omega$, it is a C&P set, and its density is proportional to the length of the acceptance window (see fact 4 of Remark 2.2). This implies that the density of a factor $w \in \mathcal{L}$ in $u_{\varepsilon,\eta}(\Omega)$ is given by

$$\varrho_w = \frac{|\Omega_w|}{|\Omega|}.$$

Another important property of the language of the infinite word $u_{\varepsilon,\eta}(\Omega)$ is given by the following proposition.

Proposition 3.7. *The language $\mathcal{L}$ of the infinite word $u_{\varepsilon,\eta}[c, c+\ell)$ is stable under mirror image, i.e.,*

$$w = w_0 w_1 \dots w_{n-1} \in \mathcal{L} \implies \overline{w} = w_{n-1} w_{n-2} \dots w_0 \in \mathcal{L}.$$

Moreover, the densities of the factors w and $\overline{w}$ coincide: $\varrho_w = \varrho_{\overline{w}}$.

Proof. Since the language of the infinite word $u_{\varepsilon,\eta}[c, c+\ell)$ depends only on the length ℓ of the acceptance interval and not on its position, it suffices to show the statement for the infinite word $u_{\varepsilon,\eta}\big[-\frac{\ell}{2}, \frac{\ell}{2}\big)$, which codes the C&P sequence $\Sigma_{\varepsilon,\eta}\big[-\frac{\ell}{2}, \frac{\ell}{2}\big)$. If $-\frac{\ell}{2} \notin \mathbb{Z}[\varepsilon]$ then

$$\Sigma_{\varepsilon,\eta}\big[-\tfrac{\ell}{2}, \tfrac{\ell}{2}\big) = \Sigma_{\varepsilon,\eta}\big(-\tfrac{\ell}{2}, \tfrac{\ell}{2}\big),$$

and thus it is a centrally symmetric set and the statement of the proposition is obvious. If $-\frac{\ell}{2} = a + b\varepsilon \in \mathbb{Z}[\varepsilon]$, then for the proof it suffices to realize that the central symmetry of the set $\Sigma_{\varepsilon,\eta}\big[-\frac{\ell}{2}, \frac{\ell}{2}\big)$ is broken by a unique point, namely $a + b\eta$. Since every factor $w \in \mathcal{L}$ occurs in $u_{\varepsilon,\eta}\big[-\frac{\ell}{2}, \frac{\ell}{2}\big)$ infinitely many times, we can still use the same argument to justify the proposition. □

For sturmian words the above property is well known, its proof can be found in [41].

For different lengths $\ell_1, \ell_2 \in (\max(-\varepsilon, 1+\varepsilon), 1]$ the languages of the infinite words $u_{\varepsilon,\eta}[c, c+\ell_1)$, $u_{\varepsilon,\eta}[c, c+\ell_2)$ are different. However, if we are interested only in factors of a given length n, the sets $\mathcal{L}_n$ can coincide even for infinite words corresponding to acceptance intervals of different lengths. Let $u_{\varepsilon,\eta}(\Omega)$ be an infinite word with the length of the acceptance window $|\Omega| = \ell$. We denote $\mathcal{L}_n(\ell)$ its set of factors of length n. For example $\mathcal{L}_1(\ell)$ is equal to the alphabet $\{A, B, C\}$ for every length $\max(-\varepsilon, 1+\varepsilon) < \ell < 1$. Let us now see how much we can change the length ℓ of the acceptance interval Ω without changing the set $\mathcal{L}_n(\ell)$.

Proposition 3.8. *Let $n \in \mathbb{N}$ be fixed. Denote by $\mathcal{C}_\ell$ the complexity function of the infinite word $u_{\varepsilon,\eta}(\Omega)$ with $\Omega = [c, c+\ell)$. Define*

$$\mathcal{D}_n = \{\ell \mid \max(-\varepsilon, 1+\varepsilon) < \ell \leq 1,\ \mathcal{C}_\ell(n) < 2n+1\}.$$

Then the elements of $\mathcal{D}_n$ divide the interval $(\max(-\varepsilon, 1+\varepsilon), 1]$ *into a finite disjoint union of sub-intervals such that $\mathcal{L}_n(\ell)$ is constant on the interior of each of these intervals.*

The proof for the special case $\varepsilon = -\frac{1}{\tau}$, $\eta = \tau$, can be found in [47]. The demonstration of the general statement follows analogous ideas.

Example 3.9. Consider again the parameters $\varepsilon = -\frac{1}{\tau}$, $\eta = \tau$. For the sake of illustration of the previous proposition, let us choose $n = 4$ and find the division of the interval $(\max(-\varepsilon, 1+\varepsilon), 1] = \left(\frac{1}{\tau}, 1\right]$ into intervals such that the set $\mathcal{L}_4(\ell)$ is constant on the interior of these intervals. For that, we need to find ℓ so that $\mathcal{C}_\ell(n) < 2n+1$ for $n = 4$. Using (3.3) this happens if

$$f^{(k)}(\delta_1) = \delta_2 \quad \text{or} \quad f^{(k)}(\delta_2) = \delta_1 \quad \text{for some } k = 0, 1, 2, 3.$$

For solving these equations, one has to realize that not only the discontinuity points $\delta_1 = c + \ell - 1 - \varepsilon$, $\delta_2 = c - \varepsilon$ depend on ℓ, but also the prescription for the function f depends on it. However, since every iteration of f is piecewise linear, the above equations can be easily solved. We find that

$$\mathcal{D}_4 = \{4 - 2\tau, -4 + 3\tau, 1\}.$$

The division of the interval $\left(\frac{1}{\tau}, 1\right]$ by the elements of $\mathcal{D}_4$ is illustrated in Figure 4.

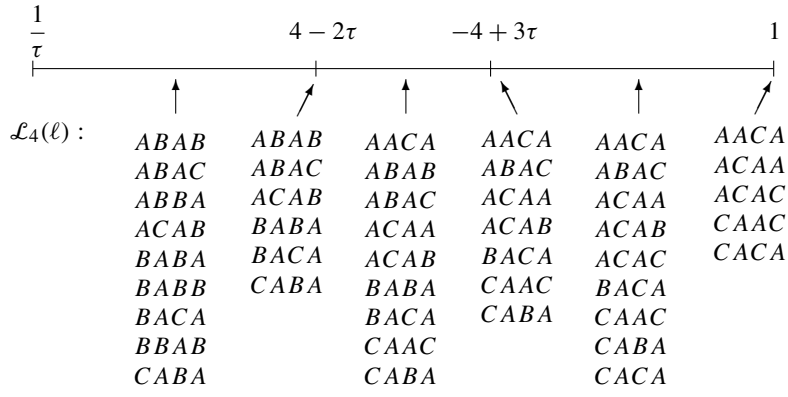


Figure 4. The appearance of factors of length 4 in $\Sigma_{-\frac{1}{\tau},\tau}(-\Omega)$ in function of the length of the acceptance window.

The figure also shows the set of factors of length 4 for each of the subintervals and for the singular lengths $\ell \in \mathcal{D}_4$. Note that the set of factors corresponding to the interior of a subinterval is a union of sets of factors corresponding to the boundary points of the subinterval, for example

$$\mathcal{L}_4(\ell) = \mathcal{L}_4(4-2\tau) \cup \mathcal{L}_4(-4+3\tau) \quad \text{for all } \ell \in (4-2\tau, -4+3\tau).$$

3.3 Special factors

Let us introduce some important notions which help us understand the structure of factors in the language $\mathcal{L}$. The notions have been introduced in [20]. Consider arbitrary bidirectional infinite word v in an alphabet $\mathcal{A}$,

$$v = \ldots v_{-2}v_{-1}v_0v_1v_2\ldots$$

For every factor $w \in \mathcal{L}$ of v there exists at least one letter $a \in \mathcal{A}$ such that $aw \in \mathcal{L}$. Such letter a is called a left extension of the factor w. The set of left extensions of the factor w is denoted by $\mathrm{Lext}(w) \subset \mathcal{A}$.

Remark 3.10. If $\tilde{w}$ is a prefix of the factor w, then

$$\mathrm{Lext}(\tilde{w}) \supseteq \mathrm{Lext}(w).$$

Since for every factor $\tilde{w} \in \mathcal{L}_{n+1}$ we have $\tilde{w} = aw$ for some $w \in \mathcal{L}_n$ and a letter $a \in \mathrm{Lext}(w)$, the increment of the complexity function can be computed as

$$\Delta\mathcal{C}(n) = \mathcal{C}(n+1) - \mathcal{C}(n) = \#\mathcal{L}_{n+1} - \#\mathcal{L}_n = \textstyle\sum_{w\in\mathcal{L}_n}(\#\,\mathrm{Lext}(w) - 1). \quad (3.6)$$

Similarly one can define the notion of right extension of a factor and obtain analogical relation

$$\Delta\mathcal{C}(n) = \textstyle\sum_{w\in\mathcal{L}_n}(\#\,\mathrm{Rext}(w) - 1). \quad (3.7)$$

Obviously, for determining the increment of complexity, only such factors w are interesting that have $\#\,\mathrm{Lext}(w) \geq 2$ or $\#\,\mathrm{Rext}(w) \geq 2$. Such factors are called left (resp. right) special factor.

Let us study these notions for infinite words corresponding to C&P sequences. Proposition 3.7 implies:

$$w \text{ is a left special factor of } u_{\varepsilon,\eta}(\Omega)$$
$$\Updownarrow$$
$$\overline{w} \text{ is a right special factor of } u_{\varepsilon,\eta}(\Omega).$$

Therefore we can limit our considerations to the study of left special factors. Theorem 3.4 implies that

$$1 \leq \Delta\mathcal{C}(n) \leq 2.$$

Thus for every $n \in \mathbb{N}$ there exists at least one and at most two left special factors of length n. Let us explain how one can decide whether a given factor $w \in \mathcal{L}_n$ is a

left special factor or not. Recall from the beginning of Section 3.1 that every factor $w \in \mathcal{L}_n$ is linked with an interval $\Omega_w \subset \Omega$ of the form $\Omega_w = [a, b)$, where

$$a, b \in \{c, \delta_1, f^{-1}(\delta_1), \dots, f^{-k+1}(\delta_1), \delta_2, f^{-1}(\delta_2), \dots, f^{-k+1}(\delta_2)\},$$

and δ_1, δ_2 are the discontinuity points of the stepping function f.

For every $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ such that $x^\star \in \Omega_w$ the n-tuple of distances in the right neighbourhood of the point x corresponds to the word w. The nearest left neighbour of the point x is determined by $f^{-1}(x^\star)$. In order that the word w is a left special factor, the interior Ω_w° must contain at least one discontinuity point of f^{-1}. If the discontinuity point of f^{-1} lies only on the boundary of Ω_w, then w has only one left extension and thus is not a left special factor. The discontinuity points of the function f^{-1} are

$$c + \ell + \varepsilon = f(\delta_1) \quad \text{and} \quad c + 1 + \varepsilon = f(c) = f^2(\delta_2).$$

Properties (3.4) of the stepping function imply that if $\ell \notin \mathbb{Z}[\varepsilon]$ then the discontinuity points of f^{-1} do not belong to the set

$$\{c, \delta_1, f^{-1}(\delta_1), \dots, f^{-n+1}(\delta_1), \delta_2, f^{-1}(\delta_2), \dots, f^{-n+1}(\delta_2)\}$$

for any $n \in \mathbb{N}$, and, therefore, if a discontinuity point of f^{-1} lies in Ω_w then it lies in its interior. We can therefore conclude with the following proposition, which is proved in a different way in [27].

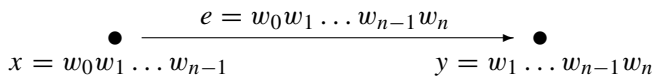
Proposition 3.11. *Let $\ell \notin \mathbb{Z}[\varepsilon]$. Consider the one-directional infinite word $u^{(i)} = u_0^{(i)} u_1^{(i)} u_2^{(i)} u_3^{(i)} \dots$, $i = 1, 2$, coding the orbits $\{f^n(c + \ell + \varepsilon) \mid n \in \mathbb{N}_0\}$ and $\{f^n(c + 1 + \varepsilon) \mid n \in \mathbb{N}_0\}$. Then a finite word w is a left special factor of $u_{\varepsilon,\eta}(\Omega)$ if and only if it is a prefix of $u^{(1)}$ or $u^{(2)}$.*

Remark 3.12. Since $c + \ell + \varepsilon$ is the image of δ_1, which is on the boundary between intervals Ω_B, Ω_C, then every prefix of the infinite word $u^{(1)}$ has $\{B, C\}$ as its left extension. Similarly, every prefix of $u^{(2)}$ has in its left extension letters A, B. It can happen that a word w is a prefix of both $u^{(1)}$ and $u^{(2)}$. Then $\mathrm{Lext}(w) = \{A, B, C\}$. However, since the infinite words $u^{(1)}, u^{(2)}$ are different, starting from a certain length of the factor w we have $\#\,\mathrm{Lext}(w) = 2$.

3.4 Rauzy graphs

Another important tool for the study of combinatorial properties of infinite words are the so-called Rauzy graphs [56], [7].

Definition 3.13. Let u be an infinite word in the alphabet $\mathcal{A}$ and let $\mathcal{L}_n$ be the set of its factors of length n, $n \in \mathbb{N}$. Rauzy graph Γ_n is a directed graph whose set of vertices is $\mathcal{L}_n$ and set of directed edges is $\mathcal{L}_{n+1}$. The edge $e \in \mathcal{L}_{n+1}$ starts at a vertex $x \in \mathcal{L}_n$ and ends at a vertex $y \in \mathcal{L}_n$ if x is a prefix of e and y is its suffix, i.e.:



The number of edges starting at a vertex x is called the outdegree of x and denoted by $d^+(x)$, the number of edges ending at x is called the indegree of x and denoted by $d^-(x)$.

From the definition of a Rauzy graph, we have

$$d^+(w) = \#\,\mathrm{Rext}(w) \quad \text{and} \quad d^-(w) = \#\,\mathrm{Lext}(w). \tag{3.8}$$

Remark 3.10 implies that if Γ_{n+1} contains a vertex with outdegree K, then the graph Γ_n contains a vertex with outdegree $\geq K$. A similar statement holds for indegrees. Therefore

$$\max_{w\in\mathcal{L}_n} d^-(w) \geq \max_{w\in\mathcal{L}_{n+1}} d^-(w) \quad \text{and} \quad \max_{w\in\mathcal{L}_n} d^+(w) \geq \max_{w\in\mathcal{L}_{n+1}} d^+(w). \tag{3.9}$$

Example 3.14. Let us consider the lower mechanical word

$$u_n = \left\lfloor \tfrac{n+1}{\tau} \right\rfloor - \left\lfloor \tfrac{n}{\tau} \right\rfloor,\ n \in \mathbb{Z},\ \text{where } \tau = \tfrac{1+\sqrt{5}}{2}.$$

Using Example 2.7, this infinite word in the alphabet $\{0, 1\}$ is a coding of the C&P sequence $\Sigma_{-\frac{1}{\tau},\eta}(\beta - 1, \beta]$. According to Remark 3.3, it is a sturmian word, i.e. of complexity $\mathcal{C}(n) = n + 1$. It can be easily computed that

$$\begin{aligned}
\mathcal{L}_3 &= \{010, 011, 101, 110\},\\
\mathcal{L}_4 &= \{0101, 0110, 1010, 1011, 1101\},\\
\mathcal{L}_5 &= \{01011, 01101, 10101, 10110, 11010, 11011\}.
\end{aligned}$$

The Rauzy graphs Γ_3, Γ_4 are illustrated in Figure 5.

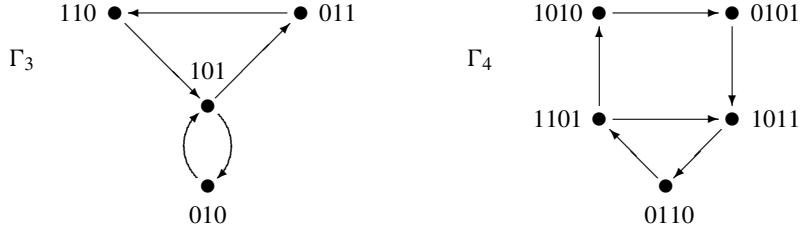


Figure 5. Rauzy graphs for the Fibonacci word.

Let us list some of the properties of the Rauzy graph Γ_n of the infinite word $u_{\varepsilon,\eta}[c, c + \ell)$ corresponding to a C&P sequence.

(1) The graph Γ_n is strongly connected for every $n \in \mathbb{N}$. It means that for every pair of vertices x, y of the graph there exists a directed path starting at x ending at y. This is a consequence of the repetitivity of the infinite word $u_{\varepsilon,\eta}[c, c+\ell)$; see (9) of Remark 2.2.

(2) For the length of the acceptance window $\ell = 1$, the infinite word $u_{\varepsilon,\eta}[c, c+1)$ is sturmian and thus $\Delta\mathcal{C}(n) = 1$ for every $n \in \mathbb{N}$. Using (3.6), (3.7) and (3.8) for every n the graph Γ_n contains exactly one vertex $x \in \mathcal{L}_n$ with outdegree 2 and exactly one vertex $y \in \mathcal{L}_n$ with indegree 2. These vertices may or may not coincide, as we have seen in Example 3.14.
For $\ell \in \mathbb{Z}[\varepsilon]$ the Theorem 3.4 implies $\Delta\mathcal{C}(n) = 1$ for sufficiently large n. Therefore the graphs have the same indegrees and outdegrees as in the sturmian case.

(3) If $\ell \notin \mathbb{Z}[\varepsilon]$, then using Theorem 3.4 we have $\Delta\mathcal{C}(n) = 2$ for all $n \in \mathbb{N}$. Since the language of the infinite word $u_{\varepsilon,\eta}[c, c+\ell)$ is stable under mirror image (Proposition 3.7), relations (3.6) and (3.7) imply that in the graph Γ_n there is either one vertex with indegree 3 and one with outdegree 3, or there are two vertices with indegree 2 and two with outdegree 2. Remark 3.12 states that a vertex with out or indegree 3 can occur only in a graph Γ_n for small n.

(4) Let us denote by $\overline{\Gamma}_n$ the graph created from Γ_n by the change of the orientation of the edges. Then $\overline{\Gamma}_n$ and Γ_n are isomorphic graphs, i.e., there exists a bijection π between the vertices of $\overline{\Gamma}_n$ and Γ_n such that for every two vertices x, y of $\overline{\Gamma}_n$ there is a directed edge from x to y if and only there is a directed edge in the graph Γ_n from $\pi(x)$ to $\pi(y)$. This property follows from Proposition 3.7.

The last mentioned property can be stated in an even stronger version if we consider the densities of factors in $\mathcal{L}_{n+1}$ as labels of the edges in the graph Γ_n.

Definition 3.15. If the densities of all factors of the infinite word u are well defined, every edge e in the Rauzy graph Γ_n can be assigned a non-negative number, namely the density ϱ_e of the factor e. The resulting graph is called a weighted Rauzy graph.

For every vertex x of the weighted Rauzy graph Γ_n we have obviously a 'conservation law',

$$\sum_{\substack{\text{edge } e \\ \text{endingin } x}} \varrho_e = \sum_{\substack{\text{edge } f \\ \text{startingin } x}} \varrho_f. \tag{3.10}$$

With the mentioned properties we can prove that the factors in a C&P word take at most 5 values.

Proposition 3.16. *Let $\mathcal{L}_n$ be the set of factors of length n of the infinite word $u_{\varepsilon,\eta}[c, c+\ell)$, $\ell \notin \mathbb{Z}[\varepsilon]$. The densities of factors in $\mathcal{L}_n$ take at most 5 values, i.e.*

$$\#\{\varrho_w \mid w \in \mathcal{L}_n\} \leq 5.$$

Proof. Consider the weighted Rauzy graph Γ_n of $u_{\varepsilon,\eta}[c, c+\ell)$. If for every vertex x of Γ_n we have $d^+(x) = d^-(x) = 1$, then the relation (3.10) implies that the density of the edge e ending at x and of the edge f starting at x coincide. Since the graph is strongly connected, these edges are different, $e \neq f$. Denote by y the starting vertex of the edge e and by z the ending vertex of the edge f. From the graph Γ_n we remove the vertex x and edges e, f and replace it by a new edge starting at y and ending at z. We assign the new edge with the weight $\varrho_e = \varrho_f$. This reduction of the graph is illustrated in Figure 6.

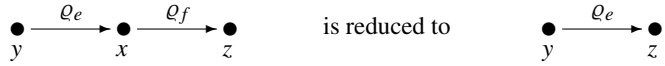


Figure 6. Reduction of the weighted Rauzy graph.

The reduction of the Rauzy graph Γ_n is repeated until there are no vertices with both outdegree and indegree 1. The resulting graph is called the reduced weighted Rauzy graph $R\Gamma_n$. The construction implies that also $R\Gamma_n$ is a strongly connected graph, the weights of its edges satisfy the conservation law and the set of weights of the graph $R\Gamma_n$ is the same as the set of weight of the graph Γ_n. Moreover, the graph $\overline{R\Gamma}_n$ created by reversing the direction of edges in $R\Gamma_n$ is isomorphic to $R\Gamma_n$.

Using the property 3 of the Rauzy graph Γ_n for sufficiently large n there are two vertices with outdegree 2, the outdegree of the remaining vertices is 1. Similarly, there are two vertices with indegree 2 and the indegree of other vertices is 1. It may happen that a vertex with outdegree 2 coincides with a vertex with indegree 2. This implies that the reduced Rauzy graph has 2, 3 or 4 vertices.

Let us discuss the case that $R\Gamma_n$ has 4 vertices, i.e., the case when none of the vertices has at the same time indegree and outdegree 2. It can be easily derived that the reduced weighted Rauzy graph has one of the forms illustrated in Figure 7. Since all the possible reduced graphs have six edges, the original weighted Rauzy graph has at most six different densities. We can eliminate the sixth value in the graphs G_1, G_2 and G_3 using the conservation law. In the graph G_1 we have $\varrho_1 = \varrho_4 + \varrho_6 = \varrho_3$. Similarly in the graph G_2 we have $\varrho_1 = \varrho_2 + \varrho_5 = \varrho_3$. In the graph G_3 we have $\varrho_1 = \varrho_4 - \varrho_6 = \varrho_3$.

The conservation law is not sufficient for reducing the number of densities in the graph G_4. Here we use the property 4 of the Rauzy graph of a C&P, namely that by changing the direction of the edges in G_4 we obtain an isomorphic graph $\bar{G}_4$. The graphs are illustrated on Figure 8. The only permutation π of the vertices which realizes the isomorphism of the graphs G_4 and $\bar{G}_4$ is the permutation $\pi(x) = v$, $\pi(y) = z$, $\pi(z) = y$, $\pi(v) = x$. The isomorphism preserves the densities, thus $\varrho_2 = \varrho_4$, $\varrho_5 = \varrho_6$.

We have thus solved the case that the reduced weighted graph $R\Gamma_n$ has 4 vertices. If $R\Gamma_n$ has 2 or 3 vertices, then such a graph has at most 5 edges. Therefore there are at most 5 values of densities.

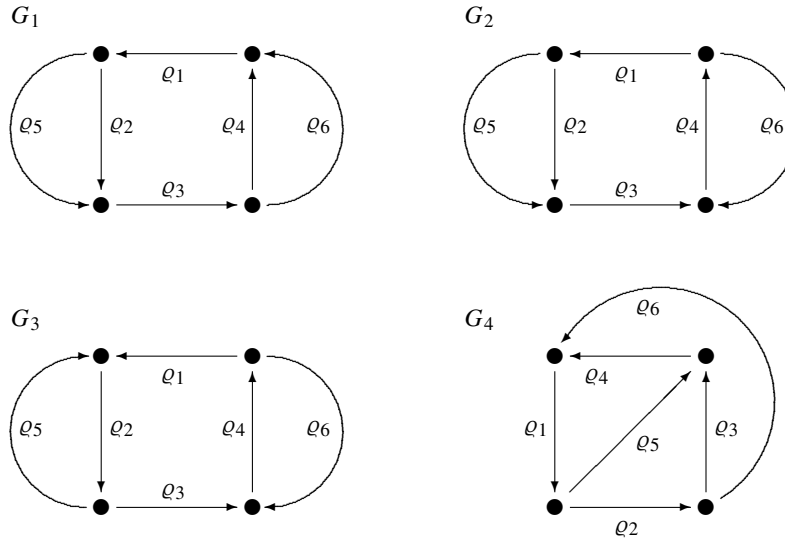


Figure 7. Possible reduces weighted Rauzy graphs.

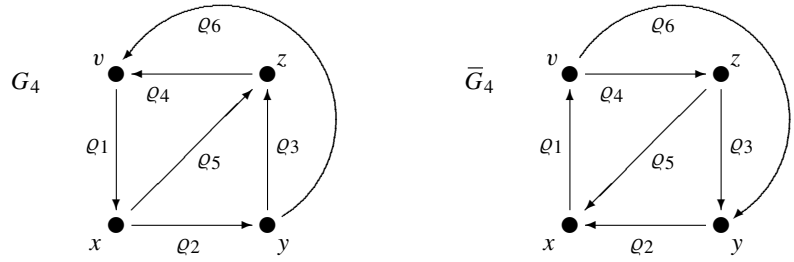


Figure 8. Isomorphic reduced graphs G_4 and $\overline{G}_4$.

For small values of n it can happen that the graph Γ_n has one vertex with outdegree 3 and one vertex with indegree 3, the other vertices having both outdegree and indegree 1. In this case the reduced Rauzy graph $R\Gamma_n$ has 1 or 2 vertices and at most 4 edges, thus the number of different values of densities is less or equal to 4. □

Remark 3.17. In the case that $\ell \in \mathbb{Z}[\varepsilon]$ the densities of factors of length n of the infinite word $u_{\varepsilon,\eta}[c, c+\ell)$ take at most 3 values for sufficiently large n, because the resulting word is either sturmian or quasisturmian and the number of densities can be read from the corresponding reduced Rauzy graph, which has always at most three edges. Let us mention that the fact that the densities of factors in sturmian words take at most three values has been stated in [14]; in fact, it can be deduced already from [62].

3.5 Sturmian words

As we have seen, sturmian words can be defined in several different equivalent ways, namely as

- bidirectional infinite words with complexity $\mathcal{C}(n) = n+1$ and irrational densities of letters;
- mechanical words $\underline{s}_{\alpha,\beta}$, $\overline{s}_{\alpha,\beta}$ with irrational slope α;
- codings of cut-and-project sequences $u_{\varepsilon,\eta}(\Omega)$, where Ω is a semi-closed interval of unit length, and ε, η are irrational numbers satisfying $\varepsilon \in (-1, 0)$, $\eta > 0$.

There exist other equivalent definitions, for example using the number of palindromes of given length or using the so-called return words. For a nice overview of these definitions see [13], [41].

The arithmetical definition of mechanical words allows one to easily derive further combinatorial properties of sturmian words. In Example 2.7 we have shown that the upper mechanical word $\overline{s}_{\alpha,\beta}$ corresponds to a cut-and-project sequence with an acceptance window which is closed from the left and open from the right. The upper mechanical word $\underline{s}_{\alpha,\beta}$ corresponds to a cut-and-project sequence with acceptance interval of opposite type. Since $\Sigma_{\varepsilon,\eta}(-\Omega) = -\Sigma_{\varepsilon,\eta}(\Omega)$ and since the language of a sturmian sequence is closed under reversal, for the study of the properties of the language we can limit our considerations to upper mechanical words $\overline{s}_{\alpha,\beta}$, see (2.13).

Let us now prove three properties which have been used in [5] for the construction of aperiodic wavelets. Note that Property 3.19 can be found already in [21].

Property 3.18. *The number of letters* 1 *in a factor of length* n *of the mechanical word* $\overline{s}_{\alpha,\beta}$ *is equal to* $\lfloor n\alpha \rfloor$ *or* $\lceil n\alpha \rceil$.

Proof. Consider a factor w of length n, $w = \overline{s}_{\alpha,\beta}(i)\overline{s}_{\alpha,\beta}(i+1)\dots\overline{s}_{\alpha,\beta}(i+n-1)$. Since the alphabet of the mechanical word is $\{0, 1\}$, the number of letters 1 in w is equal to

$$\sum_{j=0}^{n-1} \overline{s}_{\alpha,\beta}(i+j) = \lceil (i+n)\alpha + \beta \rceil - \lceil i\alpha + \beta \rceil$$

$$= \Big\lceil n\alpha + \underbrace{i\alpha + \beta - \lceil i\alpha + \beta \rceil}_{\in(-1,0)} \Big\rceil = \begin{cases} \lfloor n\alpha \rfloor, \\ \lceil n\alpha \rceil. \end{cases}$$

□

Property 3.19. *All* $n+1$ *factors of length* n *of the mechanical word* $\overline{s}_{\alpha,\beta}$ *appear in the factor* w *of length* $2n$ *of the mechanical word* $\overline{s}_{\alpha,-\alpha}$, *given by*

$$w = \overline{s}_{\alpha,-\alpha}(-n+1)\overline{s}_{\alpha,-\alpha}(-n+2)\dots\overline{s}_{\alpha,-\alpha}(0)\dots\overline{s}_{\alpha,-\alpha}(n).$$

Proof. Example 2.7 says that $\overline{s}_{\alpha,\beta}$ codes the distances in the cut-and-project sequence $\Sigma_{-\alpha,\eta}[\beta, \beta+1)$ for arbitrary $\eta > 0$. Since the language of a cut-and-project sequence

does not change by translation of the acceptance interval, we can study without loss of generality the language of the mechanical word $\bar{s}_{\alpha,-\alpha}$, i.e. of the cut-and-project sequence $\Sigma_{-\alpha,\eta}[-\alpha, 1-\alpha)$. The stepping function has a unique discontinuity point, namely $\delta_1 = 0$. The same considerations as for determining the complexity in Section 3.1 lead to the fact that the acceptance window $\Omega = [-\alpha, 1-\alpha)$ is divided by n points $\delta_1, f^{-1}(\delta_1), \ldots, f^{-n+1}(\delta_1)$ into $n+1$ disjoint subintervals closed from the left and open from the right, say $\Omega_{w^{(1)}}, \Omega_{w^{(2)}}, \ldots, \Omega_{w^{(n+1)}}$, with the following property: if x, y are elements of $\Sigma_{-\alpha,\eta}[-\alpha, 1-\alpha)$, then the n-tuples of distances starting from x and from y coincide if and only if $x^\star$, $y^\star$ belong to the same interval $\Omega_{w^{(i)}}$ for some $1 \le i \le n+1$. Since the left boundary points of all the intervals belong to $\mathbb{Z}[\alpha]$, these boundary points are star map images of points of $\Sigma_{-\alpha,\eta}[-\alpha, 1-\alpha)$. The boundary points of the intervals $w^{(i)}$ are explicitly given by

$$-\alpha = f(\delta_1),\ 0 = \delta_1,\ f^{-1}(\delta),\ \ldots,\ f^{-n+1}(\delta_1).$$

Therefore it suffices to consider all n-tuples of distances in $\Sigma_{-\alpha,\eta}[-\alpha, 1-\alpha)$, starting at point 0, at its right neighbour, and at its $n-1$ left neighbours. From Example 2.7 we know that every element $x_k \in \Sigma_{-\alpha,\eta}[-\alpha, 1-\alpha)$ has the form $x_k = \lceil k\alpha - \alpha \rceil + k\eta$ for $k \in \mathbb{Z}$. Thus $x_0 = 0$ and we must study the n-tuples of distances between points $x_{-n+1}, x_{-n+2}, \ldots, x_0, x_1, \ldots, x_{n+1}$. Since $\bar{s}_{\alpha,-\alpha}(k)$ codes the distance between x_k and x_{k+1}, the proof is finished. □

Property 3.20. *The number of factors of length n in the mechanical word $\bar{s}_{\alpha,\beta}$ prefixed by* 1 *is equal to* $\lceil n\alpha \rceil$.

Proof. Property 3.19 implies that for the description of the first letter of all $n+1$ different factors of length n it suffices to focus on letters $\bar{s}_{\alpha,-\alpha}(-n+1), \bar{s}_{\alpha,-\alpha}(-n+2), \ldots, \bar{s}_{\alpha,-\alpha}(0), \bar{s}_{\alpha,-\alpha}(1)$. The number of letters 1 among them is

$$\sum_{k=-n+1}^{1} \bar{s}_{\alpha,-\alpha}(k) = \lceil \alpha \rceil - \lceil -n\alpha \rceil = 1 - \lceil -n\alpha \rceil = \lceil n\alpha \rceil. \qquad \square$$

Let us mention an interesting consequence of Property 3.18. If w and w' are factors of $\bar{s}_{\alpha,\beta}$ of the same length, then the numbers of letters 1 in w and in w' differ at most by 1. This obviously implies that also the numbers of letters 0 in w and w' differ at most by 1. Infinite words with this property are called balanced. Sturmian words are balanced. On the other hand, every aperiodic balanced infinite word is sturmian. We have thus obtained another equivalent definition of sturmian words. The above implies other properties:

- Either 00 or 11 is not a factor of a sturmian word.
- If 00 is not a factor of $\bar{s}_{\alpha,\beta}$ or $\underline{s}_{\alpha,\beta}$ and if 01^x0 is a factor, then $x = b$ or $x = b+1$, where $b = \left[\frac{\alpha}{1-\alpha}\right]$.

Other interesting properties of sturmian words concern substitution invariance. This is the topic of the following section.

Remark 3.21. Generic cut-and-project sequences with three distances between adjacent points do not have an explicit formula for determining the n-th letter, which exists for sturmian words. Therefore the study of properties analogous to that mentioned in this subsection is significantly more difficult [27].

4 Selfsimilarity of C&P sequences

We now turn our attention to cut-and-project sequences with self-similarity. We say that a set $\Lambda \subset \mathbb{R}$ is self-similar if there exists a factor $\gamma > 1$ such that

$$\gamma \Lambda \subset \Lambda.$$

An infinite bidirectional word corresponding to a self-similar C&P set may have many interesting properties, namely under some very general condition it is a fixed point of a nontrivial morphism, or it is an image of such a fixed point. These properties are studied in Section 6.

In this section we describe the conditions on the parameters ε, η, and interval Ω, under which the C&P set $\Sigma_{\varepsilon,\eta}(\Omega)$ is self-similar. For that we need to recall some basic number theoretical notions that will be useful also in studying the invariance of C&P sets under morphisms. It turns out that ε and η are different roots of one quadratic equation with integer coefficients. Therefore, we restrict ourselves to notions connected to quadratic numbers.

For an irrational number α we denote by $\mathbb{Q}(\alpha)$ the minimal number field containing $\mathbb{Q}$ and α. If α is a quadratic number, i.e. an irrational solution of a quadratic equation with integer coefficients, then

$$\mathbb{Q}(\alpha) = \{a + b\alpha \mid a, b \in \mathbb{Q}\}.$$

The other root α' of the quadratic equation is the algebraic conjugate of α and obviously we have $\alpha' \in \mathbb{Q}(\alpha)$. On $\mathbb{Q}(\alpha)$ one defines the mapping

$$x = a + b\alpha \in \mathbb{Q}(\alpha) \mapsto x' = a + b\alpha' \in \alpha \in \mathbb{Q}(\alpha),$$

which is (the so-called Galois) automorphism on $\mathbb{Q}(\alpha)$. This means that it satisfies $(x + y)' = x' + y'$ and $(xy)' = x'y'$ for all $x, y \in \mathbb{Q}(\alpha)$.

A root of a monic quadratic polynomial with integer coefficients is called a quadratic integer. A quadratic integer γ is a quadratic Pisot number, if $\gamma > 1$ and its algebraic conjugate γ' satisfies $|\gamma'| < 1$. The following result may be found also in [10].

Theorem 4.1.

(1) *The C&P sequence $\Sigma_{\varepsilon,\eta}(\Omega)$ is self-similar if and only if ε is a quadratic number, $\eta = \varepsilon'$ is its algebraic conjugate, and the closure $\overline{\Omega}$ of the acceptance Ω contains the origin. In that case*

$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{\varepsilon,\varepsilon'}(\Omega) = \Sigma_{\eta',\eta}(\Omega) = \{x \in \mathbb{Z}[\eta] \mid x' \in \Omega\}.$$

(2) *If γ is the self-similarity factor of $\Sigma_{\eta',\eta}(\Omega)$, then γ is a quadratic Pisot number in $\mathbb{Q}[\eta]$.*

Proof. First let us show that if ε and η are mutually conjugated quadratic numbers and $0 \in \overline{\Omega}$ then $\Sigma_{\varepsilon,\eta}(\Omega)$ is self-similar. For that we have to find a self-similarity factor γ, such that $\gamma\,\Sigma_{\eta',\eta}(\Omega) \subset \Sigma_{\eta',\eta}(\Omega)$.

Let ε, η be the roots of the equation $Mx^2 = Kx + L$ for some integers K, L, M. We look for γ in the form $\gamma = a + Mb\eta$ for some integers a, b. Such γ satisfies $\gamma \in \mathbb{Z}[\eta]$ and $\gamma\eta = a\eta + Mb\eta^2 = Lb + \eta(a + Kb) \in \mathbb{Z}[\eta]$. Therefore $\gamma\mathbb{Z}[\eta] \subset \mathbb{Z}[\eta]$.

Since $\gamma \in \mathbb{Z}[\eta] \subset \mathbb{Q}(\eta)$ we determine the image of γ under the Galois automorphism $\gamma' = a + Mb\eta'$. Clearly, $\gamma\gamma' \in \mathbb{Z}$. Since η' is irrational the set $\mathbb{Z}[\eta'] = \mathbb{Z} + \mathbb{Z}\eta'$ is dense in $\mathbb{R}$, and thus there are infinitely many choices of $a, b \in \mathbb{Z}$ so that $\gamma' \in (0, 1)$. Together with the fact $0 \in \overline{\Omega}$ it follows that $\gamma'\Omega \subset \Omega$. We use the above to obtain

$$\begin{aligned}\gamma\,\Sigma_{\eta',\eta}(\Omega) &= \gamma\{x \in \mathbb{Z}[\eta] \mid x' \in \Omega\} = \{\gamma x \in \gamma\mathbb{Z}[\eta] \mid \gamma' x' \in \gamma'\Omega\} \\ &\subset \{\gamma x \in \mathbb{Z}[\eta] \mid \gamma' x' \in \Omega\} \subseteq \{y \in \mathbb{Z}[\eta] \mid y' \in \Omega\} = \Sigma_{\eta',\eta}(\Omega).\end{aligned}$$

Since $\gamma' \in (0, 1)$ and $\gamma\gamma' \in \mathbb{Z}$, we have $|\gamma| > 1$. If $\gamma > 1$ it is the desired self-similarity factor, in the opposite case we choose γ^2 for the self-similarity factor.

Let us prove the necessary condition for the self-similarity of a C&P set. Let $\gamma > 1$ satisfy $\gamma\,\Sigma_{\varepsilon,\eta}(\Omega) \subset \Sigma_{\varepsilon,\eta}(\Omega)$. For a chosen point $x = a + b\eta \in \Sigma_{\varepsilon,\eta}(\Omega)$ we have $\gamma x \in \Sigma_{\varepsilon,\eta}(\Omega) \subset \mathbb{Z}[\eta]$. Therefore there must exist integers $\tilde{a}$, $\tilde{b}$ such that $\gamma x = \gamma(a + b\eta) = \tilde{a} + \tilde{b}\eta$. This implies

$$\gamma = \frac{\tilde{a} + \tilde{b}\eta}{a + b\eta} \quad \text{and} \quad \eta = \frac{-\tilde{a} + a\gamma}{\tilde{b} - b\gamma}.$$

Therefore $\mathbb{Q}(\gamma) = \mathbb{Q}(\eta)$.

Let $(x_n)_{n\in\mathbb{Z}}$ be the strictly increasing sequence such that $\Sigma_{\varepsilon,\eta}(\Omega) = \{x_n \mid n \in \mathbb{Z}\}$. Recall that the distances between neighbouring points of $\Sigma_{\varepsilon,\eta}(\Omega)$ take values $x_{n+1} - x_n \in \{\Delta_1, \Delta_2, \Delta_1 + \Delta_2\}$, where Δ_1, Δ_2 are positive numbers in $\mathbb{Z}[\eta]$ linearly independent over $\mathbb{Q}$. Take an index n such that $x_{n+1} - x_n = \Delta_1$. Since $\Sigma_{\varepsilon,\eta}(\Omega)$ is self-similar with the factor γ, both γx_n and γx_{n+1} belong to $\Sigma_{\varepsilon,\eta}(\Omega)$. Therefore the gap between the two points is filled by distances Δ_1, Δ_2 and $\Delta_1 + \Delta_2$. It follows that the distance $\gamma x_{n+1} - \gamma x_n$ is an integer combination of Δ_1, Δ_2 with positive coefficients,

$$\gamma\Delta_1 = \gamma x_{n+1} - \gamma x_n = k_{11}\Delta_1 + k_{12}\Delta_2$$

for some non-negative integers k_{11}, k_{12}. Analogously we obtain

$$\gamma\Delta_2 = k_{21}\Delta_1 + k_{22}\Delta_2, \quad k_{21}, k_{22} \in \mathbb{N}_0.$$

We denote by $\mathbb{K}$ the 2×2 matrix $\mathbb{K} = (k_{ij})$ and write the above as

$$\mathbb{K}\begin{pmatrix}\Delta_1 \\ \Delta_2\end{pmatrix} = \gamma\begin{pmatrix}\Delta_1 \\ \Delta_2\end{pmatrix}.$$

This means that γ is an eigenvalue of the integer-valued 2×2 matrix $\mathbb{K}$ and as such is a root of a monic quadratic polynomial with integer coefficients. Since $\mathbb{Q}(\eta) = \mathbb{Q}(\gamma)$ and η is irrational, γ is a quadratic integer.

The eigenvector corresponding to γ is $\binom{\Delta_1}{\Delta_2}$. As Δ_1, Δ_2 belong to the quadratic field $\mathbb{Q}(\eta) = \mathbb{Q}(\gamma)$, we can apply the Galois automorphism to obtain the other eigenvector and eigenvalue of the matrix $\mathbb{K}$,

$$\mathbb{K}\begin{pmatrix}\Delta_1'\\ \Delta_2'\end{pmatrix} = \gamma'\begin{pmatrix}\Delta_1'\\ \Delta_2'\end{pmatrix}.$$

The Perron–Frobenius theorem for positive integer matrices implies that $|\gamma'| < \gamma$.

We now explain the relation between the Galois automorphism and the star map in the cut-and-project scheme. Take any pair of points $x, y \in \Sigma_{\varepsilon,\eta}(\Omega)$, $x < y$. Their distance $y - x$ belongs to $\Sigma_{\varepsilon,\eta}(\Omega) - \Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{\varepsilon,\eta}(\Omega - \Omega)$. From the definition of the C&P set, we have $(y - x)^\star \in \Omega - \Omega$. Now let $x = \gamma^m x_n$, $y = \gamma^m x_{n+1}$ for any integer power m and for some n such that $x_{n+1} - x_n = \Delta_1$, or $x_{n+1} - x_n = \Delta_2$ respectively. From the self-similarity of $\Sigma_{\varepsilon,\eta}(\Omega)$, the points x, y belong to $\Sigma_{\varepsilon,\eta}(\Omega)$, and hence $(\gamma^m \Delta_1)^\star$, $(\gamma^m \Delta_2)^\star \in \Omega - \Omega$. Therefore the sequence of vectors

$$\mathbb{K}^m\begin{pmatrix}\Delta_1^\star\\ \Delta_2^\star\end{pmatrix} = \left(\mathbb{K}^m\begin{pmatrix}\Delta_1\\ \Delta_2\end{pmatrix}\right)^\star = \left(\gamma^m\begin{pmatrix}\Delta_1\\ \Delta_2\end{pmatrix}\right)^\star = \begin{pmatrix}(\gamma^m\Delta_1)^\star\\ (\gamma^m\Delta_2)^\star\end{pmatrix} \tag{4.1}$$

is bounded with $m \to \infty$. In the above we have used the property of the star map $(kx)^\star = kx^\star$ for $x \in \mathbb{Z}[\eta]$ and any integer k. Since the eigenvectors of the matrix $\mathbb{K}$ form a basis of $\mathbb{R}^2$, we can write

$$\begin{pmatrix}\Delta_1^\star\\ \Delta_2^\star\end{pmatrix} = \alpha_1\begin{pmatrix}\Delta_1\\ \Delta_2\end{pmatrix} + \alpha_2\begin{pmatrix}\Delta_1'\\ \Delta_2'\end{pmatrix}$$

for some real coefficients α_1, α_2. Substituting into (4.1) we derive that the sequence of vectors

$$\alpha_1\gamma^m\begin{pmatrix}\Delta_1\\ \Delta_2\end{pmatrix} + \alpha_2\gamma'^m\begin{pmatrix}\Delta_1'\\ \Delta_2'\end{pmatrix}$$

is bounded. Since $\gamma > 1$ we have $\alpha_1 = 0$ and $|\gamma'| < 1$. We can conclude that γ is a quadratic Pisot number and $\binom{\Delta_1^\star}{\Delta_2^\star} = \alpha_2\binom{\Delta_1'}{\Delta_2'}$. The lengths Δ_1, Δ_2 belong to $\mathbb{Z}[\eta]$ and hence can be written in the form $\Delta_1 = a_1 + b_1\eta$, $\Delta_2 = a_2 + b_2\eta$ for some integers a_1, a_2, b_1, b_2. We have

$$\alpha_2 = \frac{a_1 + b_1\varepsilon}{a_2 + b_2\varepsilon} = \frac{a_1 + b_1\eta'}{a_2 + b_2\eta'},$$

which implies $(a_1b_2 - a_2b_1)(\varepsilon - \eta') = 0$. Since Δ_1, Δ_2 are linearly independent over $\mathbb{Q}$, we have $a_1b_2 - a_2b_1 \neq 0$ and thus $\varepsilon = \eta'$ as the theorem claims. The star map in such a cut-and-project scheme coincides with the Galois automorphism on the quadratic field $\mathbb{Q}(\eta) = \mathbb{Q}(\varepsilon) = \mathbb{Q}(\gamma)$.

The last to be verified is that $\overline{\Omega}$ contains the origin. Since $\gamma\,\Sigma_{\varepsilon,\eta}(\Omega) \subset \Sigma_{\varepsilon,\eta}(\Omega)$, it follows easily that $\gamma'\Omega \subset \Omega$. This implies $0 \in \overline{\Omega}$. □

From the proof of the above theorem it follows that if ε, η are mutually conjugated quadratic numbers and $0 \in \overline{\Omega}$, (i.e. $\Sigma_{\varepsilon,\eta}(\Omega)$ is self-similar) there exists infinitely many factors γ such that $\gamma\,\Sigma_{\varepsilon,\eta}(\Omega) \subset \Sigma_{\varepsilon,\eta}(\Omega)$. It can be shown that all of these factors are quadratic Pisot numbers in $\mathbb{Z}[\eta]$. The exact description of all self-similarity factors of a given C&P set is straightforward, but rather technical. For a generalisation of self-similarity studied on the most common example $\eta = \tau$ we refer to [45].

Finally, let us mention that first results about self-similar Delone sets with Meyer property (which include C&P sets) have been obtained by Meyer in [49]. He shows that the self-similarity factor of such sets must be a Pisot or Salem number, i.e. an algebraic integer > 1 with all conjugates in the unit disc. In an even more general setting, self-similarity of Delone sets was studied in [34], [35].

5 Non-standard numeration systems and C&P sequences

Another example of self-similar sets are sequences formed by β-integers. We show how they are related to C&P sequences. For the definition of β-integers we introduce the notion of β-expansion, which has been first given by Rényi [58]. The β-expansions are studied from the arithmetical point of view for example in [59], [55], [28], [16].

Let β be a real number greater than 1. For a non-negative $x \in \mathbb{R}$ we find a unique k such that $\beta^k \le x < \beta^{k+1}$ and put

$$x_k := \left[\frac{x}{\beta^k}\right], \quad r_k := x - x_k\beta^k.$$

The coefficients x_i, $i \in \mathbb{Z}$, $i \le k-1$ we define recursively

$$x_i := [\beta r_{i+1}], \quad r_i := \beta r_{i+1} - x_i.$$

The described procedure is called the greedy algorithm. It ensures that

$$x = \sum_{i=-\infty}^{k} x_i\beta^i.$$

The above expression of x using an infinite series is called the β-expansion of x. For $\beta = 2$ or $\beta = 10$ we obtain the usual binary or decimal expansion of x. The real numbers x for which the coefficients $x_{-1}, x_{-2}, x_{-3}, \ldots$ in the β-expansion of $|x|$ vanish, are called β-integers. They form the set denoted by $\mathbb{Z}_\beta$,

$$\mathbb{Z}_\beta = \Big\{ \pm \textstyle\sum_{i=0}^{k} x_i\beta^i \mid \sum_{i=0}^{k} x_i\beta^i \text{ is a } \beta\text{-expansion of an } x \ge 0 \Big\}.$$

The greedy algorithm implies that if $x = \sum_{i=-\infty}^{k} x_i\beta^i$ is the β-expansion of a number x, then $\sum_{i=-\infty}^{k+1} x_{i-1}\beta^i$ is the β-expansion of βx. Therefore we trivially have for β-integers

$$\beta\mathbb{Z}_\beta \subset \mathbb{Z}_\beta.$$

Let us mention that $\mathbb{Z}_\beta$ has also many other self-similarity factors.

If β is an integer greater than 1, the set of β-integers coincides with rational integers, $\mathbb{Z}_\beta = \mathbb{Z}$. Drawn on the real line, the distances between adjacent points of $\mathbb{Z}_\beta$ are all 1, and all integers > 1 are self-similarity factors of $\mathbb{Z}_\beta$. In this case the coefficients (digits) x_i in a β-expansion take values $0, 1, \ldots, \beta - 1$, and every finite sequence formed by these digits is a β-expansion of some number x.

The situation is very different if $\beta \notin \mathbb{Z}$. As a consequence of the greedy algorithm, the digits in a β-expansion take values $0, 1, \ldots, \lceil \beta \rceil - 1$. However, not all strings of these digits correspond to a number x as its β-expansion. Which sequences of digits are permissible in β-expansions and what are the distances between adjacent points in $\mathbb{Z}_\beta$ depends on the so-called Rényi development of 1. We define a mapping

$$T_\beta(x) = \beta x - \lfloor \beta x \rfloor \quad \text{for } x \in [0, 1].$$

Put $t_i := \lfloor \beta T_\beta^{i-1}(1) \rfloor$ for $i = 1, 2, 3, \ldots$. The sequence

$$d_\beta(1) = t_1 t_2 t_3 \ldots$$

is called the Rényi development of 1.

In order to decide whether a series $\sum_{i=0}^{n} x_i \beta^i$ is a β-expansion, we use the condition of Parry [55].

Proposition 5.1. *Let $\beta > 1$. Then $\sum_{i=0}^{n} x_i \beta^i$ is a β-expansion of a number x if and only if the word $x_i x_{i-1} \ldots x_0$ is lexicographically strictly smaller than $t_1 t_2 t_3 \ldots$ for all $0 \le i \le n$.*

In [63] it is shown that the distances between neighbouring points in the set $\mathbb{Z}_\beta$ are of the form

$$\sum_{k=1}^{\infty} \frac{t_{i+k}}{\beta^k} \quad \text{for } i = 0, 1, 2, \ldots.$$

A necessary condition in order that $\mathbb{Z}_\beta$ has only finitely many distances between neighbouring points is that the Rényi development of 1 is eventually periodic. The construction of $d_\beta(1)$ implies that

$$1 = \sum_{i=1}^{\infty} \frac{t_i}{\beta^i}. \tag{5.1}$$

If moreover $d_\beta(1)$ is eventually periodic, β is a root of a monic polynomial with integer coefficients. Such β is an algebraic integer.

Our aim is to describe which parameters have to be chosen in order that the sets $\Sigma_{\varepsilon,\eta}(\Omega)$ and $\mathbb{Z}_\beta$ coincide on the positive half-axis,[2] i.e. when

$$\Sigma_{\varepsilon,\eta}(\Omega) \cap \mathbb{R}_0^+ = \mathbb{Z}_\beta \cap \mathbb{R}_0^+. \tag{5.2}$$

[2] Positive half-axis $[0, +\infty)$ is denoted by $\mathbb{R}_0^+$.

In [19] an example of such a relation is given together with the parameters ε, η, Ω. In particular, the authors study the case of quadratic Pisot units. All quadratic Pisot units can be expressed as the positive roots of a quadratic equation

$$\beta^2 = m\beta + 1 \text{ for } m \geq 1 \quad \text{or} \quad \beta^2 = m\beta - 1 \text{ for } m \geq 3.$$

It is shown that

$$\Sigma_{\beta',\beta}\big[-1, -\tfrac{1}{\beta'}\big) \cap \mathbb{R}_0^+ = \mathbb{Z}_\beta \cap \mathbb{R}_0^+ \quad \text{for } \beta^2 = m\beta + 1,\ m \geq 1, \tag{5.3}$$

$$\Sigma_{\beta',\beta}\big[0, \tfrac{1}{\beta'}\big) \cap \mathbb{R}_0^+ = \mathbb{Z}_\beta \cap \mathbb{R}_0^+ \quad \text{for } \beta^2 = m\beta - 1,\ m \geq 3. \tag{5.4}$$

The above equalities imply that the C&P sequences with given windows have two distances only between adjacent points. Thus from Theorem 2.5 they are geometrically similar to C&P sequences with unit acceptance interval, and therefore the infinite binary words corresponding to β-integers are sturmian words.

In the following proposition we prove that a quadratic Pisot unit β is the only example of a basis for which positive β-integers coincide with the restriction of a cut-and-project set to its positive part.

Proposition 5.2. *Positive part of the set $\mathbb{Z}_\beta$ coincides with the positive part of a cut-and-project set $\Sigma_{\varepsilon,\eta}(\Omega)$ if and only if β is a quadratic Pisot unit.*

Proof. One implication is obvious from (5.3) and (5.4). Let us prove $\Rightarrow$. Since $\mathbb{Z}_\beta$ is a self-similar set, we impose the requirement of self-similarity also on the C&P sets, which implies that ε, η are mutually conjugated quadratic numbers, i.e., $\varepsilon = \eta'$ and $0 \in \overline{\Omega}$. According to Theorem 4.1 the self-similarity factor β is a quadratic Pisot number. Thus it remains to show that β is a unit.

All quadratic Pisot numbers can be expressed as the positive roots of a quadratic equation

$$\beta^2 = m\beta + n \text{ for } 1 \leq n \leq m \quad \text{or} \quad \beta^2 = m\beta - n \text{ for } 1 \leq n \leq m-2.$$

Our considerations can thus be divided into two cases.

1. Let $\beta^2 = m\beta + n$, $m \geq n \geq 1$. Then the conjugated root to β is the number $\beta' \in (-1, 0)$. The Rényi development of 1 has the form

$$d_\beta(1) = mn,$$

and the distances between β-integers are

$$\sum_{i=1}^{\infty} \frac{t_i}{\beta^i} = 1 \quad \text{and} \quad \sum_{i=1}^{\infty} \frac{t_{i+1}}{\beta^i} = \beta - t_1 = \beta - m = \frac{n}{\beta}.$$

In this case a series $\sum_{i=0}^{k} x_i \beta^i$ with non-negative integer coefficients x_i is a β-expansion if $x_i x_{i-1}$ is strictly lexicographically smaller than mn, i.e. $x_i x_{i-1} \prec mn$ for all $1 \leq i \leq k$, which means that $x_i \in \{0, 1, \ldots, m\}$ and every digit $x_i = m$ in the string $x_k x_{k-1} \ldots x_1 x_0$ is followed by a digit $x_{i-1} \leq n - 1$.

Since in a self-similar C&P set the star map and the Galois automorphism coincide, we can find a candidate for the acceptance interval Ω in order that (5.2) be satisfied. In the following estimations we use $\beta' \in (-1, 0)$. For $x = \sum_{i=0}^{k} x_i \beta^i \in \mathbb{Z}_\beta$ we have

$$x' = \sum_{i=0}^{k} x_i \beta'^i < m + m\beta'^2 + m\beta'^4 + \cdots = \frac{m}{1-\beta'^2}.$$

Similarly,

$$x' = \sum_{i=0}^{k} x_i \beta'^i > m\beta' + m\beta'^3 + m\beta'^5 + \cdots = \frac{m\beta'}{1-\beta'^2}.$$

Clearly, the only candidate for Ω is the interval $\left[\frac{m\beta'}{1-\beta'^2}, \frac{m}{1-\beta'^2}\right)$. It is obvious that for such a window one inclusion of (5.2) is verified,

$$\Sigma_{\eta',\eta}(\Omega) \cap \mathbb{R}_0^+ \supseteq \mathbb{Z}_\beta \cap \mathbb{R}_0^+.$$

Since $\mathbb{Z}_\beta$ has only two possible distances between neighbouring elements, namely 1 and $\frac{n}{\beta}$, the equality in the above inclusion is reached according to (8) of Remark 2.2 if $|\Omega| = \Delta'_1 - \Delta'_2$, i.e., if

$$\frac{m}{1-\beta'^2} - \frac{m\beta'}{1-\beta'^2} = \frac{m}{1+\beta'} = 1 - \frac{n}{\beta}.$$

Using the quadratic equation $\beta'^2 = m\beta' + n$ we obtain the condition $(1-n)\beta' = 0$ which implies $n = 1$. Thus β is a unit.

2. Let us study the case $\beta^2 = m\beta - n$, $m - 2 \geq n \geq 1$. Here the conjugated root β' belongs to the interval $(0, 1)$. The Rényi development of 1 has coefficients $t_1 = m - 1$ and $t_i = m - n - 1$ for $i \geq 2$, i.e.

$$d_\beta(1) = (m-1)(m-n-1)(m-n-1)\ldots = (m-1)(m-n-1)^\omega.$$

In this case a β-expansion of a number x has digits in the set $\{0, 1, 2, \ldots, m-1\}$ and forbidden are the strings of digits equal or lexicographically greater than $(m-1)(m-n-1)^s(m-n)$ for arbitrary non-negative integer s. The distances between neighbouring β-integers are

$$\sum_{i=1}^{\infty} \frac{t_i}{\beta^i} = 1 \quad \text{and} \quad \sum_{i=1}^{\infty} \frac{t_{i+1}}{\beta^i} = \beta - t_1 = \beta - (m-1) = 1 - \frac{n}{\beta}.$$

Again using the Galois conjugation of $x = \sum_{i=0}^{k} x_i \beta^i \in \mathbb{Z}_\beta$ we find a candidate on the acceptance interval Ω for (5.2),

$$0 \leq x' = \sum_{i=0}^{k} x_i \beta'^i < m - 1 + (m-2)\beta' + (m-2)\beta'^2 + \cdots = 1 + \frac{m-2}{1-\beta'}.$$

Let us therefore set $\Omega = \left[0,\, 1+\frac{m-2}{1-\beta'}\right)$. In order that a C&P set with such an acceptance window has only two distances between neighbours, we must have

$$|\Omega| = 1 + \frac{m-2}{1-\beta'} = \Delta'_1 - \Delta'_2 = 1 - 1 + \frac{n}{\beta'} = \frac{n}{\beta'}.$$

After manipulations we obtain $(n-1)\beta' = 0$, which implies $n = 1$. This completes the proof. □

6 C&P sequences and substitutions

Construction of an arbitrarily long segment of a C&P sequence directly from the definition of $\Sigma_{\varepsilon,\eta}(\Omega)$ is numerically very demanding, since precise computation with irrational numbers requires a special arithmetics dependent on the form in which the irrational numbers ε, η are given. For a class of self-similar C&P sequences the sequence of distances between adjacent points (i.e. the infinite word $u_{\varepsilon,\eta}(\Omega)$) can be generated effectively using substitution rules. For this purpose we introduce the following notions.

The set $\mathcal{A}^*$ of finite words on an alphabet $\mathcal{A}$ equipped with the empty word ϵ and the operation of concatenation is a free monoid. A morphism on the monoid $\mathcal{A}^*$ is a map $\varphi\colon \mathcal{A}^* \to \mathcal{A}^*$ satisfying $\varphi(wz) = \varphi(w)\varphi(z)$ for any pair of words $w, z \in \mathcal{A}^*$. Clearly, the morphism φ is determined by $\varphi(a)$ for all $a \in \mathcal{A}$. The action of a morphism φ can be easily extended to one-directional infinite words $u = u_0u_1u_2\ldots$ over $\mathcal{A}$ by the prescription

$$\varphi(u) = \varphi(u_0u_1u_2\ldots) = \varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots.$$

Let $u = u_0u_1u_2\ldots$ be an infinite word over an alphabet $\mathcal{A}$ and let φ be a morphism on $\mathcal{A}^*$ satisfying $|\varphi(a)| \geq 1$ for all $a \in \mathcal{A}$ and $|\varphi(u_0)| > 1$. We say that the word u is invariant under the substitution φ if u is its fixed point, i.e. $\varphi(u) = u$.

The φ-invariance of u implies that $\varphi(u_0)$ has the form $\varphi(u_0) = u_0u'$ for some non-empty word $u' \in \mathcal{A}^*$ and that $\varphi^n(u) = u$ for every $n \in \mathbb{N}$. The word $\varphi^n(u_0)$ is a prefix of the fixed point u and its length grows to infinity with n, therefore we can formally write $u = \lim_{n\to\infty} \varphi^n(u_0)$. The substitution under which an infinite word u is invariant allows one to generate u starting from the initial letter u_0 repeating the rewriting rules infinitely many times.

As an example of a substitution invariant C&P sequence let us recall the β-integers, as presented in the previous section. Consider first $\mathbb{Z}_\beta$ for $\beta^2 = m\beta + 1$, where the distances are $\Delta_1 = 1$ and $\Delta_2 = \frac{1}{\beta}$. Associating the letter A to the distance 1 and the letter B to the distance $\frac{1}{\beta}$ we create a one-directional infinite word u in the alphabet $\{A, B\}$. It can be easily seen from the properties of β-expansions and from the Parry condition that, if x, y are neighbours in $\mathbb{Z}_\beta$ such that $y - x = 1$, then, between the points βx and βy, there is m times the distance 1 followed by one distance $\frac{1}{\beta}$.

Similarly, if x, y are neighbours in $\mathbb{Z}_\beta$ such that $y - x = \frac{1}{\beta}$, then the points βx, βy are also neighbours and have distance 1. The above considerations imply that the infinite word u is not changed if every letter A is replaced by the finite word $A^m B$ and every letter B is replaced by A. We say that the word u corresponding to $\mathbb{Z}_\beta$ is invariant under the substitution φ given by

$$\varphi(A) = A^m B, \quad \varphi(B) = A.$$

Similarly we can derive that for β-integers where $\beta^2 = m\beta - 1$, the infinite word u corresponding to $\mathbb{Z}_\beta$ is invariant under the substitution

$$\varphi(A) = A^{m-1} B, \quad \varphi(B) = A^{m-2} B.$$

We have presented the substitutions only for those β-integers that correspond to C&P sequences. However in general, every $\mathbb{Z}_\beta$ which has a finite number of distances between neighbours is invariant under a non-trivial substitution [25].

To every substitution φ in the alphabet $\mathcal{A} = \{a_1, \ldots, a_k\}$ one associates naturally the substitution matrix $M \in M_k(\mathbb{N}_0)$, where

$$M_{ij} = \text{the number of letters } a_j \text{ in the word } \varphi(a_i).$$

If all letters a_i of the alphabet $\mathcal{A}$ have a well-defined density ϱ_i in the infinite word u invariant under the substitution φ, then the vector $(\varrho_1, \varrho_2, \ldots, \varrho_k)$ is a left eigenvector of the matrix M. If the matrix M is primitive, i.e. it has a positive power, then the substitution is called *primitive*. A fixed point of a primitive substitution can be represented geometrically as a self-similar sequence in the following way.

According to the Perron–Frobenius theorem, the matrix M has an up to a scalar factor unique positive right eigenvector $(y_1, y_2, \ldots, y_k)^T$ corresponding to the dominant eigenvalue, say λ. To the infinite word $u = u_0 u_1 u_2 \ldots$ we associate the sequence $(z_n)_{n \in \mathbb{N}_0}$ such that

$$z_0 = 0 \quad \text{and} \quad z_{n+1} - z_n = y_i \quad \text{if } u_n = a_i. \tag{6.1}$$

The sequence $(z_n)_{n \in \mathbb{N}_0}$ is self-similar since we have

$$\lambda\{z_n \mid n \in \mathbb{N}_0\} \subset \{z_n \mid n \in \mathbb{N}_0\}.$$

6.1 Substitution invariance

Infinite words corresponding to C&P sequences are bidirectional. From the property 2 of Remark 2.2 it follows that, without loss of generality, we can consider only those C&P sequences which have $0 \in \Omega$, i.e. $0 \in \Sigma_{\varepsilon,\eta}(\Omega)$. We define a pointed bidirectional infinite word $u_{\varepsilon,\eta}(\Omega) = \ldots u_{-2} u_{-1} | u_0 u_1 u_2 \ldots$ such that $u_0 u_1 u_2 \ldots$ corresponds to the order of distances between adjacent points of $\Sigma_{\varepsilon,\eta}(\Omega)$ on the right of 0, and $\ldots u_{-3} u_{-2} u_{-1}$ corresponds to the order of distances between adjacent points of $\Sigma_{\varepsilon,\eta}(\Omega)$ on the left of 0. This word is ternary or binary.

Let $u = \ldots u_{-2}u_{-1}|u_0u_1u_2\ldots$ be a pointed bidirectional infinite word over an alphabet $\mathcal{A}$. Let φ be a morphism on $\mathcal{A}^*$ such that $|\varphi(u_{-1})| > 1$ and $|\varphi(u_0)| > 1$. We say that the word u is invariant under the substitution φ if it satisfies

$$u = \ldots u_{-2}u_{-1}|u_0u_1u_2\ldots = \ldots\varphi(u_{-2})\varphi(u_{-1})|\varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots = \varphi(u).$$

In this case we formally write

$$u = \lim_{n\to\infty} \varphi^n(u_{-1})|\varphi^n(u_0).$$

Let us mention what is known about the substitution invariance of infinite words associated to C&P sequences. First consider the binary words. As explained in Example 2.7, all such words coincide with lower and upper mechanical words $\underline{s}_{\alpha,\beta}$, $\overline{s}_{\alpha,\beta}$ for irrational $\alpha \in (0, 1)$ and any real $\beta \in [0, 1)$. The question about substitution invariance of mechanical words has been solved independently by different authors [64], [15], [9]. In order to state the result we need to introduce the notion of a Sturm number.

Definition 6.1. A quadratic irrational number $\alpha \in (0, 1)$ whose algebraic conjugate α' satisfies $\alpha' \notin (0, 1)$ is called a Sturm number.

Let us mention that originally Sturm numbers were defined by a special form of their continued fraction. The characterization presented here is due to [2]. The necessary and sufficient condition for substitution invariance of mechanical words is given by the following theorem [9].

Theorem 6.2. *The mechanical word $\underline{s}_{\alpha,\beta}$ resp. $\overline{s}_{\alpha,\beta}$ for irrational $\alpha \in (0, 1)$ and real $\beta \in [0, 1)$ is invariant under a substitution if and only if*

(i) *α is a Sturm number,*
(ii) *$\beta \in \mathbb{Q}(\alpha)$,*
(iii) *$\alpha' \le \beta' \le 1 - \alpha'$ or $1 - \alpha' \le \beta' \le \alpha'$.*

If we represent the substitution invariant lower mechanical word $\underline{s}_{\alpha,\beta}$ geometrically, as described in (6.1), we find that this geometrical representation coincides with the C&P sequence $\Sigma_{\varepsilon,\varepsilon'}(\beta - 1, \beta]$, where $\alpha = -\varepsilon$. Similar statement is valid for the substitution invariant upper mechanical word $\overline{s}_{\alpha,\beta}$. This implies that substitution invariance of a binary word u associated to a C&P sequence forces the existence of a self-similar C&P sequence $\Sigma_{\varepsilon,\varepsilon'}(\Omega)$ such that $u = u_{\varepsilon,\varepsilon'}(\Omega)$.

Substitution invariance of ternary words corresponding to C&P sequences has not yet been solved completely. The authors however conjecture that, even in this case, substitution invariance forces self-similarity of the corresponding C&P sequence.

6.2 Substitutivity

The original aim for studying substitution properties of C&P sequences was the possibility of symbolic generation of $u_{\varepsilon,\eta}(\Omega)$. For our present purpose, it is enough to

consider a property weaker than substitution invariance, namely the substitutivity. We take the formulation of Durand [24].

Definition 6.3. We say that the infinite word u over an alphabet $\mathcal{A}$ is substitutive if there exist an infinite word v over an alphabet $\mathcal{B}$ and a letter projection $\psi : \mathcal{B} \to \mathcal{A}$ such that v is invariant under a substitution φ on $\mathcal{B}^*$ and

$$\ldots \psi(v_{-2})\psi(v_{-2})|\psi(v_0)\psi(v_1)\psi(v_2)\ldots = \ldots u_{-2}u_{-1}|u_0u_1u_2\ldots.$$

If moreover φ is a primitive substitution, then the infinite word u is said to be primitive substitutive.

If an infinite word u is substitutive, it can be constructed in such a way that generating by substitution the word v and using the projection ψ allows us to obtain u.

Using Theorem 2.5 we can without loss of generality consider the ternary words associated to C&P sequences $\Sigma_{\varepsilon,\eta}[c, c+\ell)$, where

$$\varepsilon \in (-1, 0), \quad \eta > 0, \quad c \le 0 < c + \ell \quad \text{and} \quad \max(-\varepsilon, 1+\varepsilon) < \ell < 1. \tag{6.2}$$

The description of infinite words associated to C&P sequences which are substitutive can be derived from the paper of Adamczewski [1].

Theorem 6.4. *Let $\varepsilon, \eta, c, \ell$ satisfy* (6.2)*. The infinite word $u_{\varepsilon,\eta}[c, c+\ell)$ is primitive substitutive if and only if ε is a quadratic irrational number and $c, \ell \in \mathbb{Q}(\varepsilon)$.*

From a practical point of view it is important to know the procedure which, given a substitutive word u over a ternary alphabet $\{A, B, C\}$, allows to determine an alphabet $\mathcal{B}$, a substitution invariant word v over $\mathcal{B}$ and a projection $\psi\colon \mathcal{B} \to \mathcal{A}$ such that $\psi(v) = u$. The bidirectional pointed word $v = \ldots v_{-2}v_{-1}|v_0v_1v_2$ is in fact given by the initial letters $v_{-1}|v_0$ and the substitution φ under which it is invariant, since we have $v = \lim_{n\to\infty} \varphi^n(v_{-1})|\varphi^n(v_0)$.

In the rest of this section we describe the algorithm for solving this problem in the case that the parameters satisfy besides the necessary conditions (6.2), an additional condition that $-\varepsilon$ is a Sturm number, i.e. the algebraic conjugate ε' of ε satisfies $\varepsilon' < -1$ or $\varepsilon' > 0$. Using the transformations (2.6) and (2.7) we have

$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{-1-\varepsilon,-1-\eta}(-\Omega),$$

and thus we can without loss of generality consider only $\varepsilon' > 0$. As for the parameter η, we know that, for fixed ε, c, ℓ satisfying (6.2), the words $u_{\varepsilon,\eta}[c, c+\ell)$ coincide for all choices of $\eta > 0$. In the case that $\varepsilon' > 0$ it is suitable to put $\eta = \varepsilon'$. Then $\Sigma_{\varepsilon,\varepsilon'}[c, c+\ell)$ is a self-similar set. This is a crucial property for proving the correctness of the algorithm presented below.

Remark 6.5. The sequence $\Sigma_{\varepsilon,\varepsilon'}[c, c+\ell)$ with parameters satisfying (6.2) has according to Theorem 2.5 three types of distances between adjacent points, namely

$\varepsilon', 1+\varepsilon', 1+2\varepsilon'$. According to Theorem 4.1 it is a self-similar set. The stepping function f on the acceptance interval $\Omega = [c, c+\ell)$ has in this case the form

$$f(x) = \begin{cases} x+1+\varepsilon & \text{for } x \in [c, c+\ell-1-\varepsilon) =: \Omega_A, \\ x+1+2\varepsilon & \text{for } x \in [c+\ell-1-\varepsilon, c-\varepsilon) =: \Omega_B, \\ x+\varepsilon & \text{for } x \in [c-\varepsilon, c+\ell) =: \Omega_C. \end{cases}$$

This function is a bijection on the acceptance interval Ω, i.e. is invertible.

Algorithm.

Input: quadratic $\varepsilon \in (-1, 0)$ with $\varepsilon' > 0$, $c, \ell \in \mathbb{Q}(\varepsilon)$ such that $c \le 0 < c+\ell$, $\max(-\varepsilon, 1+\varepsilon) < \ell \le 1$.

Output: alphabet $\mathcal{B}$, letters $v_{-1}, v_0 \in \mathcal{B}$, morphism φ on $\mathcal{B}^*$, projection $\psi : \mathcal{B} \to \mathcal{A}$.

Step 1. Find a quadratic unit $\gamma \in (0, 1)$ such that $\gamma\mathbb{Z}[\varepsilon] = \mathbb{Z}[\varepsilon]$ and its conjugate $\gamma' > 1$. It results in solving a Diophantine equation (more precisely the so-called Pell equation) which has always a solution.

Step 2. For $x \in \Omega$ we define

$$g_\gamma(x) = \tfrac{1}{\gamma} f^{-\operatorname{ind}(x)}(x), \text{ where } \operatorname{ind}(x) = \min\{i \in \mathbb{N}_0 \mid f^{-i}(x) \in \gamma\Omega\}. \tag{6.3}$$

Find the minimal set $S \subset \Omega$ such that

$$\{c, c+\ell-1-\varepsilon, c-\varepsilon\} \subseteq S \quad \text{and} \quad g_\gamma(S) \subseteq S.$$

Such a set is finite. Let us denote its elements by $S = \{c_0, c_1, \ldots, c_k\}$, where $c = c_0 < c_1 < \cdots < c_k$, and let $c_{k+1} := c + \ell$. Note that the elements of the set S divide the acceptance window into small subintervals

$$\Omega = \bigcup_{i=0}^{k} [c_i, c_{i+1}).$$

Step 3. Define the alphabet $\mathcal{B} := \{0, 1, \ldots, k\}$ and to every letter $i \in \mathcal{B}$ associate the number

$$j_i = \min\{j \in \mathbb{N} \mid f^j(\gamma c_i) \in \gamma\Omega\}$$

and the word $\varphi(i) = w_0^{(i)} w_1^{(i)} \ldots w_{j_i-1}^{(i)}$ by the prescription

$$w_j^{(i)} := m \in \mathcal{B} \quad \text{if } f^j(\gamma c_i) \in [c_m, c_{m+1}).$$

Step 4. Define the initial letters of the infinite word $v = \ldots v_{-2} v_{-1} | v_0 v_1 v_2 \ldots$ over the alphabet $\mathcal{B}$ as

$$v_0 = m \in \mathcal{B} \quad \text{if } 0 \in [c_m, c_{m+1}),$$
$$v_{-1} = m \in \mathcal{B} \quad \text{if } f^{-1}(0) \in [c_m, c_{m+1}).$$

Step 5. Define the projection $\psi\colon \mathcal{B} \to \mathcal{A} = \{A, B, C\}$ by

$$\psi(i) = \begin{cases} A & \text{if } c_i \in \Omega_A, \\ B & \text{if } c_i \in \Omega_B, \\ C & \text{if } c_i \in \Omega_C. \end{cases}$$

Theorem 6.6. *Let parameters ε, η, c, ℓ satisfy* (6.2). *Let moreover ε be a quadratic irrational number such that $\varepsilon' > 0$, and let $c, \ell \in \mathbb{Q}(\varepsilon)$. Then the alphabet $\mathcal{B}$, the letters $v_{-1}, v_0 \in \mathcal{B}$, the morphism φ on $\mathcal{B}^*$, and the projection $\psi\colon \mathcal{B} \to \mathcal{A}$, defined in the above algorithm, satisfy*

$$u_{\varepsilon,\eta}[c, c+\ell) = \psi(v) \quad \textit{with } v = \lim_{n\to\infty} \varphi^n(v_{-1})|\varphi^n(v_0).$$

The proof of the theorem follows the same ideas as in [46], where the correctness of the algorithm for $\varepsilon = -\frac{1}{\tau}$ is shown. Note that the crucial point in the algorithm is to ensure that the set S of Step 2 is finite.

Remark 6.7. 1. Given the infinite word $u_{\varepsilon,\eta}(\Omega)$, the substitution φ is not unique. Indeed, the ambiguity is found in the choice of the unit γ in Step 1 of the algorithm. Note that if γ has required properties, then so does any power γ^j, $j \in \mathbb{N}$.

2. The cardinality of the alphabet $\mathcal{B}$ is given by the cardinality of the set S, which depends on the choice of γ. Taking a power of γ as the unit factor may reduce the number of letters in the alphabet.

3. If the word $u_{\varepsilon,\eta}(\Omega)$ is not only substitutive but is also a fixed point of a substitution, then a suitable choice of γ (of sufficiently high power of the minimal factor satisfying Step 1) in the algorithm yields the substitution under which $u_{\varepsilon,\eta}(\Omega)$ is invariant.

Let us illustrate the algorithm for finding the substitution in an example.

Example 6.8. Consider the C&P sequence $\Sigma_{\varepsilon,\eta}[c, c+\ell)$ with parameters

$$\varepsilon = -\frac{1}{\sqrt{2}}, \quad \eta = \varepsilon' = \frac{1}{\sqrt{2}}, \quad c = 0, \quad \ell = -2 + 2\sqrt{2} = -2 - 4\varepsilon.$$

Such parameters clearly satisfy the assumptions of the algorithm. The distances between adjacent points of $\Sigma_{\varepsilon,\eta}[c, c+\ell)$ are $\Delta_1 = 1 + \eta = 1 + \frac{1}{\sqrt{2}}$, $\Delta_2 = \eta = \frac{1}{\sqrt{2}}$, and $\Delta_1 + \Delta_2 = 1 + 2\eta = 1 + \sqrt{2}$, and therefore the explicit expression of the stepping function in our case is

$$f(x) = \{ \begin{cases} x + 1 + \varepsilon & \text{for } \in [0, -3 - 5\varepsilon) =: \Omega_A, \\ x + 1 + 2\varepsilon & \text{for } x \in [-3 - 5\varepsilon, -\varepsilon) =: \Omega_B, \\ x + \varepsilon & \text{for } x \in [-\varepsilon, -2 - 4\varepsilon) =: \Omega_C. \end{cases}$$

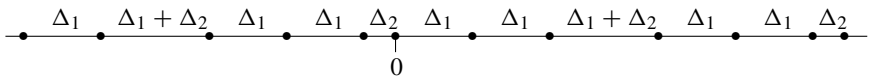
From that we derive the formula for the inverse function

$$f^{-1}(x) = \begin{cases} x - \varepsilon & \text{for } x \in [0, -2 - 3\varepsilon), \\ x - 1 - 2\varepsilon & \text{for } x \in [-2 - 3\varepsilon, 1 + \varepsilon), \\ x - 1 - \varepsilon & \text{for } x \in [-1 - \varepsilon, -2 - 4\varepsilon). \end{cases}$$

We know that $0 \in \Sigma_{\varepsilon,\eta}[c, c + \ell)$. We can generate other elements of the C&P sequences on the right from 0 using the stepping function f. The elements on the left are generated using f^{-1}. We have

$$\Sigma_{\varepsilon,\eta}[c, c + \ell) = \{\ldots, -5 - 8\eta, -4 - 6\eta, -3 - 5\eta, -2 - 3\eta, -1 - 2\eta, -\eta, \\ 0, 1 + \eta, 2 + 2\eta, 3 + 4\eta, 4 + 5\eta, 5 + 6\eta, 5 + 7\eta, \ldots\},$$

and graphically,



The corresponding bidirectional infinite word $u_{\varepsilon,\eta}[c, c+\ell)$ is obtained by replacing Δ_1 with the letter A, Δ_2 with the letter C, and $\Delta_1 + \Delta_2$ with the letter B, i.e.,

$$u_{\varepsilon,\eta}[c, c + \ell) = \ldots BABAAC|AABAAC \ldots$$

Let us now use the algorithm described above to derive the substitution generating the word $u_{\varepsilon,\eta}[c, c + \ell)$. We proceed according to the steps of the algorithm.

Step 1. Put $\gamma = 3 + 4\varepsilon = 3 - 2\sqrt{2}$. It is obvious that γ is a unit in $\mathbb{Z}[\varepsilon] \cap (0, 1)$ and its algebraic conjugate γ' satisfies $\gamma' = \frac{1}{\gamma} = 3 - 4\varepsilon = 3 + 2\sqrt{2} > 1$. We have yet to verify that $\gamma\mathbb{Z}[\varepsilon] = \mathbb{Z}[\varepsilon]$. For this it suffices to show that $\gamma\varepsilon \in \mathbb{Z}[\varepsilon]$ and $\gamma^{-1}\varepsilon \in \mathbb{Z}[\varepsilon]$. We have

$$\gamma\varepsilon = \varepsilon(3 + 4\varepsilon) = 2 + 3\varepsilon \in \mathbb{Z}[\varepsilon],$$
$$\gamma^{-1}\varepsilon = \varepsilon(3 - 4\varepsilon) = -2 + 3\varepsilon \in \mathbb{Z}[\varepsilon],$$

where we have used that $\varepsilon^2 = \frac{1}{2}$.

Step 2. We need to find the minimal set $S \subset \Omega$ closed under the action of g_γ, containing the points

$$c = 0, \quad c + \ell - 1 - \varepsilon = -3 - 5\varepsilon, \quad c - \varepsilon = -\varepsilon.$$

Thus we search for values of all iterations $g_\gamma^j(x)$, $j \in \mathbb{N}$, of the above points. Important for the definition (6.3) of the function g_γ is the index of a point x, i.e. the first exponent $i \in \mathbb{N}_0$ such that $f^{-i}(x)$ belongs to the interval

$$\gamma\Omega = (3 + 4\varepsilon)[0, -2 - 4\varepsilon) = [0, -14 - 20\varepsilon).$$

Let us find the image under g_γ of the point $x = 0$. Clearly, $0 \in \gamma\Omega$, thus $\mathrm{ind}(0) = 0$. So $g_\gamma(0) = \frac{1}{\gamma} f^0(0) = 0$ and $g_\gamma^j(0) = 0$ for all $j \in \mathbb{N}$.

Let us find the image under g_γ of the point $x = -\varepsilon$. We have

$$f^{-2}(-\varepsilon) = f^{-1}(-1-2\varepsilon) = -2-3\varepsilon \in \gamma\Omega,$$

which implies

$$\mathrm{ind}(-\varepsilon) = 2 \quad \text{and} \quad g_\gamma(-\varepsilon) = \frac{1}{\gamma} f^{-2}(-\varepsilon) = (3-4\varepsilon)(-2-3\varepsilon) = -\varepsilon.$$

Thus $g_\gamma^j(-\varepsilon) = -\varepsilon$ for all $j \in \mathbb{N}$.

Let us now find the image under g_γ of the point $x = -3-5\varepsilon$. We have

$$\begin{aligned} f^{-4}(-3-5\varepsilon) &= f^{-3}(-4-6\varepsilon) = f^{-2}(-5-8\varepsilon) \\ &= f^{-1}(-6-9\varepsilon) = -7-10\varepsilon \in \gamma\Omega, \end{aligned}$$

which implies

$$\begin{aligned} \mathrm{ind}(-3-5\varepsilon) = 4 \quad \text{and} \quad g_\gamma(-\varepsilon) &= \frac{1}{\gamma} f^{-4}(-3-5\varepsilon) \\ &= (3-4\varepsilon)(-7-10\varepsilon) = -1-2\varepsilon. \end{aligned}$$

Therefore the set S must contain the point $-1-2\varepsilon$. In order to find further iterations of g_γ on the point $-3-5\varepsilon$, we determine $g_\gamma(-1-2\varepsilon)$. We have seen that $f^{-1}(-1-2\varepsilon) = -2-3\varepsilon \in \gamma\Omega$. Thus

$$\mathrm{ind}(-1-2\varepsilon) = 1 \quad \text{and} \quad g_\gamma(-1-2\varepsilon) = \frac{1}{\gamma} f^{-1}(-1-2\varepsilon) = -\varepsilon.$$

Altogether, we obtain

$$\begin{aligned} g_\gamma^j(0) &= 0 && \text{for all } j \in \mathbb{N}, \\ g_\gamma^j(-\varepsilon) &= -\varepsilon && \text{for all } j \in \mathbb{N}, \\ g_\gamma(-3-5\varepsilon) &= -1-2\varepsilon \\ g_\gamma^j(-3-5\varepsilon) &= -\varepsilon && \text{for all } j \in \mathbb{N},\ j \geq 2. \end{aligned}$$

We can conclude that S contains four elements,

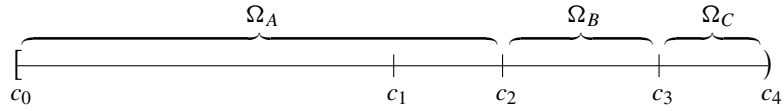
$$c_0 = c = 0, \quad c_1 = -1-2\varepsilon, \quad c_2 = -3-5\varepsilon, \quad c_3 = -\varepsilon,$$

and we put $c_4 = -2-4\varepsilon$.

Note that the elements of the set S divide the acceptance window into small subintervals

$$\begin{aligned} \Omega &= \bigcup_{i=0}^{3} [c_i, c_{i+1}) \\ &= [0, -1-2\varepsilon) \cup [-1-2\varepsilon, -3-5\varepsilon) \cup [-3-5\varepsilon, -\varepsilon) \cup [-\varepsilon, -2-4\varepsilon), \end{aligned}$$

as it is illustrated in the following figure.



Step 3. Since S has four elements, we have the alphabet $\mathcal{B} := \{0, 1, 2, 3\}$ on four letters. In order to define the substitution we compute the iterations $f^j(\gamma c_i)$ and we stop when $f^j(\gamma c_i) \in \gamma\Omega$. First take $i = 0$. We have $\gamma c_0 = 0$ and the iterations

$$f^0(0) = 0 \in [c_0, c_1),$$
$$f^1(0) = 1 + \varepsilon \in [c_0, c_1),$$
$$f^2(0) = 2 + 2\varepsilon \in [c_2, c_3) \text{ (thus } j_0 = 6 \text{ and } \varphi(0) = 002013),$$
$$f^3(0) = 3 + 4\varepsilon \in [c_0, c_1),$$
$$f^4(0) = 4 + 5\varepsilon \in [c_1, c_2),$$
$$f^5(0) = 5 + 6\varepsilon \in [c_3, c_4),$$
$$f^6(0) = 5 + 7\varepsilon \in \gamma\Omega.$$

Note that the word $\varphi(0) = 002013$ is formed by the indices m of the left-end-points of the intervals $[c_m, c_{m+1})$ read in the column. Similarly for $i = 1$, we have $\gamma c_1 = \gamma(-1 - 2\varepsilon) = -7 - 10\varepsilon$. Thus

$$f^0(-7 - 10\varepsilon) = -7 - 10\varepsilon \in [c_0, c_1),$$
$$f^1(-7 - 10\varepsilon) = -6 - 9\varepsilon \in [c_0, c_1),$$
$$f^2(-7 - 10\varepsilon) = -5 - 8\varepsilon \in [c_2, c_3) \text{ (hence } j_1 = 5 \text{ and } \varphi(1) = 00202),$$
$$f^3(-7 - 10\varepsilon) = -4 - 6\varepsilon \in [c_0, c_1),$$
$$f^4(-7 - 10\varepsilon) = -3 - 5\varepsilon \in [c_2, c_3),$$
$$f^5(-7 - 10\varepsilon) = -2 - 3\varepsilon \in \gamma\Omega.$$

For $i = 2$, we have $\gamma c_2 = \gamma(-3 - 5\varepsilon) = -19 - 27\varepsilon$. Therefore

$$f^0(-19 - 27\varepsilon) = -19 - 27\varepsilon \in [c_0, c_1),$$
$$f^1(-19 - 27\varepsilon) = -18 - 26\varepsilon \in [c_0, c_1),$$
$$f^2(-19 - 27\varepsilon) = -17 - 25\varepsilon \in [c_2, c_3),$$
$$f^3(-19 - 27\varepsilon) = -16 - 23\varepsilon \in [c_0, c_1) \text{ (hence } j_2 = 8 \text{ and } \varphi(2) = 00202013),$$
$$f^4(-19 - 27\varepsilon) = -15 - 22\varepsilon \in [c_2, c_3),$$
$$f^5(-19 - 27\varepsilon) = -14 - 20\varepsilon \in [c_0, c_1),$$
$$f^6(-19 - 27\varepsilon) = -13 - 19\varepsilon \in [c_1, c_2),$$
$$f^7(-19 - 27\varepsilon) = -12 - 18\varepsilon \in [c_3, c_4),$$
$$f^8(-19 - 27\varepsilon) = -12 - 17\varepsilon \in \gamma\Omega.$$

Finally for $i = 3$ we have $\gamma c_3 = \gamma(-\varepsilon) = -2 - 3\varepsilon$. Thus

$$f^0(-2-3\varepsilon) = -2-3\varepsilon \in [c_0, c_1),$$
$$f^1(-2-3\varepsilon) = -1-2\varepsilon \in [c_1, c_2) \quad (\text{ hence } j_3 = 3 \text{ and } \varphi(3) = 013),$$
$$f^2(-2-3\varepsilon) = -\varepsilon \in [c_3, c_4),$$
$$f^3(-2-3\varepsilon) = 0 \in \gamma\Omega.$$

Altogether, we have the substitution

$$\varphi(0) = 002013,$$
$$\varphi(1) = 00202,$$
$$\varphi(2) = 00202013,$$
$$\varphi(3) = 013.$$

Step 4. Since $0 \in [c_0, c_1)$ and $f^{-1}(0) = -\varepsilon \in [c_3, c_4)$, we put as the initial letters $v_0 = 0$, $v_{-1} = 3$. Note that 0 is a prefix of $\varphi(0)$ and 3 is a suffix of $\varphi(3)$, thus the word $v = \lim_{n\to\infty} \varphi^n(v_{-1})|\varphi^n(v_0) = \lim_{n\to\infty} \varphi^n(3)|\varphi^n(0)$ is well defined.

Step 5. As $c_0, c_1 \subset \Omega_A, c_2 \subset \Omega_B, c_3 \subset \Omega_C$, we have the projection $\psi : \{0, 1, 2, 3\} \to \mathcal{A} = \{A, B, C\}$ by

$$\psi(0) = \psi(1) = A, \quad \psi(2) = B, \quad \psi(3) = C.$$

Let us write the subsequent iterations of the substitution φ applied to the pair of initial letters $3|0$, i.e. $\varphi^n(3)|\varphi^n(0)$. We have for $n = 0, 1, 2, \ldots$

$$\begin{array}{c} 3|0 \\ 013|002013 \\ 00201300202013|00201300201300202013002013002020130 \\ \vdots \end{array}$$

Since each row is a factor of the next one, in the limit we obtain the infinite word

$$v = \ldots 00201300202013|00201300201300202013002013002020130\ldots.$$

Now we apply the letter projection ψ, which collapses the letters 0 and 1. We have

$$\begin{array}{rl} v = & \ldots 1\,3\,0\,0\,2\,0\,2\,0\,1\,3\,|\,0\,0\,2\,0\,1\,3\,0\,0\,2\,0\,1\,3\,0\,0\,2\,0\,2\,0\,1\,3\,0\,0\ldots \\ & \quad\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\quad\downarrow \\ u_{\varepsilon,\eta}[c, c+\ell) = & \ldots ACAABABAAC|AABAACAABAACAABABAACAA\ldots. \end{array}$$

Let us study the second iteration of the above substitution φ. We obtain

$$\varphi^2(0) = \varphi(002013) = 00201300201300202013002013002020130,$$
$$\varphi^2(1) = \varphi(00202) = 00201300201300202013002013002020130,$$

$\varphi^2(2) = \varphi(00202013) = 00201300201300202013002013002020130020130020201300201300202013,$
$\varphi^2(3) = \varphi(013) = 00201300202013.$

Note that $\varphi^2(0) = \varphi^2(1)$. Therefore we can consider the letters 0 and 1 as identic. It enables us to define a new substitution $\tilde{\varphi}\colon \{A, B, C\}^* \to \{A, B, C\}^*$ by

$$\begin{aligned}
\tilde{\varphi}(A) &= AABAACAABAACAABABAACAABAACAABABAAC,\\
\tilde{\varphi}(B) &= AABAACAABAACAABABAACAABAACAABABAAC\\
&\quad \circ AABAACAABABAAC,\\
\tilde{\varphi}(C) &= AABAACAABABAAC,
\end{aligned}$$

under which the word $u_{\varepsilon,\eta}[c, c+\ell)$ is invariant; namely we have

$$u_{\varepsilon,\eta}[c, c+\ell) = \lim_{n\to\infty} \tilde{\varphi}^n(C)|\tilde{\varphi}^n(A).$$

Note that the symbol $\circ$ in the substitution formula stands for concatenation.

7 Conclusions

In this paper, we attempted to give a unifying view of the one-dimensional cut-and-project point sets obtained from the square lattice in the plane. Part of the work is a review of former results which had to be recalled for the sake of clarity. Let us now indicate some possible continuations or applications of our results.

First, it would be interesting to extend the work [4] to the case of splines of larger regularity for other cut-and-project sets, with quadratic self-similarity, or without self-similarity at all, especially in view the relation between scaling equations and substitution properties of the considered discretizations of $\mathbb{R}$.

Mathematical diffraction of such aperiodic sets obtained by cut and projection should be also envisaged in a systematic way, in relation with the existence of those multiresolution analysis and related wavelets. Concerning diffraction one can find in the literature on quasicrystals many works devoted to this important subject, in which substitutional properties, self-similarity or cut and projection from higher-dimensional lattices play a central role in the elaboration of rigorous results (see for instance [17], [32], [42] and [37] for a recent review on these questions). The analysis of diffraction spectra by using adapted wavelets, i.e. wavelets "living" on the diffracting aperiodic structure, is a project which remains to be really developed.

Another nice application of the results of this paper can be envisaged in the construction of a new type of pseudo-random number generators. A first step in this direction was made in [31], where the authors use sturmian sequences to combine classical periodic pseudo-random sequences to produce an aperiodic pseudo-random sequence. These aperiodic pseudo-random number generators (APRNG) were tested using the DIEHARD test suite and using the Maurer test, and it turned out that sta-

tistical properties of these APRNGs are significantly better than those of the original periodic sequences. Moreover, the authors prove that the APRNG passes the spectral test. It would be very interesting to pursue the study of the APRNGs extending the definition to generic (i.e. ternary) cut-and-project sequences.

Acknowledgement. The authors acknowledge partial support by Czech Science Foundation GA ČR 201/05/0169.

References

[1] B. Adamczewski, Codages de rotations et phénomènes d'autosimilarité, *J. Théor. Nombres Bordeaux* 14 (2002), 351–386.

[2] C. Allauzen, Simple characterization of Sturm numbers, *J. Théor. Nombres Bordeaux* 10 (1998), 237–241.

[3] J.-P. Allouche, Sur la complexité des suites infinies, *Bull. Belg. Math. Soc. Simon Stevin* 1 (1994), 133–143.

[4] M. Andrle, Č.Burdík, J. P. Gazeau, Bernuau spline wavelets and sturmian sequences, *J. Fourier Anal. Appl.* 10 (2004), 269–300.

[5] M. Andrle, Č. Burdík, J. P. Gazeau, R. Krejcar, Wavelet multiresolutions for the Fibonacci chain, *J. Phys A: Math. Gen.* 33 (2000), L47–L51.

[6] M. Andrle, A. El-Kharrat, Wavelet basis for quasicrystal diffraction, in *Group 24: Physical and mathematical aspects of symmetries*, Proceedings of the 24th International Colloquium on Group Theoretical Methods in Physics, Paris, 15-20 July 2002, Institute of Physics Conference Series 173, IOP Publishing, Bristol 2003, 185–188.

[7] P. Arnoux, G. Rauzy, Représentation géométrique de suites de complexité $2n + 1$, *Bull. Soc. Math. France* 119 (1991), 199–215.

[8] P. Auscher, Wavelet bases for $L^2(R)$ with rational dilation factor, in *Wavelets and their applications*, Jones and Barlett Publishers, Boston 1992, 439–452.

[9] P. Baláži, Z. Masáková, E. Pelantová, Complete characterization of substitution invariant sturmian sequences, *Integers: Electronic Journal of Combinatorial Number Theory* 5 (2005), #A14, 23pp.

[10] P. Baláži, E. Pelantová, Cut-and-project sequences, in *Group 24: Physical and mathematical aspects of symmetries*, Proceedings of the 24th International Colloquium on Group Theoretical Methods in Physics, Paris, 15–20 July 2002, Institute of Physics Conference Series 173, IOP Publishing, Bristol 2003, 201–204.

[11] G. Bernuau, *Propriétés spectrales et géométriques des quasicristaux. Ondelettes adaptées aux quasicristaux*, PhD thesis, Ceremade, Université Paris IX Dauphine, France 1998.

[12] G. Bernuau, Wavelet bases adapted to a self-similar quasicrystal, *J. Math. Phys.* 39 (1998), 4213–4225.

[13] J. Berstel, Recent results on extensions of sturmian words, *Internat. J. Algebra Comput.* 12 (2002), 371–385.

[14] V. Berthé, Fréquences des facteurs des suites sturmiennes, *Theoret. Comput. Sci.* 165 (1996), 295–309.

[15] V. Berthé, H. Ei, S. Ito, H. Rao, Invertible substitutions and sturmian words: An application of Rauzy fractals, preprint 2005.

[16] A. Bertrand, Développements en base de Pisot et répartition modulo 1, *C. R. Acad. Sc. Paris, Série A* 285 (1977), 419–421.

[17] E. Bombieri, J. E. Taylor, Which distributions of matter diffract? An initial investigation. *J. Physique* 47 (1986), Suppl. Colloq. C3, C3-19–C3-28.

[18] C. Buhmann, M. Micchelli, Spline prewavelets for nonuniform knots, *Numer. Math.* 61 (1992), 455–474.

[19] Č. Burdík, Ch. Frougny, J. P. Gazeau, R. Krejcar, Beta-integers as natural counting systems for quasicrystals, *J. Phys. A: Math. Gen.* 31 (1998), 6449–6472.

[20] J. Cassaigne, Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc. Simon Stevin* 4 (1997), 67–88.

[21] J. Cassaigne, Sequences with grouped factors, in *Developments in Language Theory III* (*DLT'97*), Aristotle University of Thessaloniki, Thessaloniki 1998, 211–222.

[22] E. M. Coven, G. A. Hedlund, Sequences with minimal block growth, *Math. Systems Theory* 7 (1973), 138–153.

[23] M. Duneau and A. Katz, Quasiperiodic patterns, *Phys. Rev. Lett.* 54 (1985) 2688–2691.

[24] F. Durand, A characterization of substitutive sequences using return words, *Discrete Math.* 179 (1998), 89–101.

[25] S. Fabre, Substitution et β-systèmes de numération, *Theoret. Comput. Sci.* 137 (1995), 219–236.

[26] S. Ferenczi, C. Holton, L. Q. Zamboni, Structure of three interval exchange transformations. I. An arithmetic study, *Ann. Inst. Fourier* 51 (2001), 861–901.

[27] S. Ferenczi, C. Holton, L. Q. Zamboni, Structure of three-interval exchange transformations. II. A combinatorial description of the trajectories, *J. Anal. Math.* 89 (2003), 239–276.

[28] Ch. Frougny, B. Solomyak, Finite beta-expansions, *Ergodic Theory Dynam. Systems* 12 (1992), 713–723.

[29] J. P. Gazeau, J. Patera, Tau wavelets of Haar, *J. Phys. A: Math. Gen.* 29 (1996), 4549–4559.

[30] L. S. Guimond, Z. Masáková, E. Pelantová, Combinatorial properties of infinite words associated with cut-and-project sequences, *J. Théor. Nombres Bordeaux* 15 (2003), 697–725.

[31] L. S. Guimond, Jan Patera, Jiří Patera, Statistical properties and implementation of aperiodic pseudorandom number generators, *Appl. Numer. Math.* 46 (2003), 295–318.

[32] A. Hof, On diffraction by aperiodic structures, *Comm. Math. Phys.* 169 (1995), 25–43.

[33] P. A. Kalugin, A. Y. Kitaev, L. S. Levitov, $Al_{0.86}Mn_{0.14}$: a six-dimensional crystal, *JETP Lett.* 41 (1985), 145–149.

[34] R. Kenyon, The construction of self-similar tilings, *Geom. Funct. Anal.* 6 (1996), 471–488.

[35] R. Kenyon, Inflationary tilings with a similarity structure, *Comment. Math. Helv.* 69 (1994), 169–198.

[36] J. C. Lagarias, Geometric models for quasicrystals I. Delone sets of finite type. *Discrete Comput. Geom.* 21 (1999), 161–191.

[37] J. C. Lagarias, Mathematical quasicrystals and the problem of diffraction, in *Directions in mathematical quasicrystals*, CRM Monogr. Ser. 13, Amer. Math. Soc., Providence, RI, 2000, 61–93.

[38] J. C. Lagarias, P. A. Pleasants, Repetitive Delone sets and quasicrystals, *Ergodic Theory Dynam. Systems* 23 (2003), 831–867.

[39] P. G. Lemarié-Rieusset, Base d'ondelettes sur les groupes de Lie stratifiés, *Bull. Soc. Math. Fr.* 117 (1989), 211–232.

[40] D. Levine, P. J. Steinhardt, Quasicrystals I: Definitions and structure, *Phys. Rev. B* 34 (1986), 596–616.

[41] M. Lothaire, *Algebraic combinatorics on words*, Cambridge University Press, Cambridge 2002.

[42] J. M. Luck, C. Godrèche, A. Janner, T. Janssen, The nature of the atomic surfaces of quasiperiodic self-similar structures, *J. Phys. A* 26 (1993), 1951–1999.

[43] S. Mallat, *A wavelet tour of signal processing*, 2nd ed., Academic Press, San Diego 1999.

[44] S. Mallat, Multiresolution approximation and orthonormal bases of wavelets for $L^2(R)$, *Trans. Amer. Math. Soc.* 315 (1989), 69–87.

[45] Z. Masáková, J. Patera, E. Pelantová, Inflation centers of the cut and project quasicrystals, *J. Phys. A: Math. Gen.* 31 (1998), 1443–1453.

[46] Z. Masáková, J. Patera, E. Pelantová, Substitution rules for aperiodic sequences of the cut and project type, *J. Phys. A: Math. Gen.* 33 (2000), 8867–8886.

[47] Z. Masáková, J. Patera, J. Zich, Classification of Voronoi and Delone tiles in quasicrystals. I. General method. *J. Phys. A* 36 (2003), 1869–1894.

[48] Y. Meyer, *Algebraic numbers and harmonic analysis*, North-Holland Mathematical Library 2, North-Holland, Amsterdam 1972.

[49] Y. Meyer, Quasicrystals, diophantine approximations and algebraic numbers, in *Beyond quasicrystals*, Les Houches 1994, EDP Sciences, Les Ulis, Springer-Verlag, Berlin 1995, 3–16.

[50] R. V. Moody, Meyer sets and their duals, in *The mathematics of long-range aperiodic order*, Waterloo, ON, 1995, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 489, Kluwer Acad. Publ., Dordrecht 1997, 403–441.

[51] R. V. Moody, J. Patera, Densities, minimal distances, and coverings of quasicrystals, *Comm. Math. Phys.* 195 (1998), 613–626.

[52] R. V. Moody, J. Patera, Quasicrystals and icosians, *J. Phys A: Math. Gen.* 26 (1993), 2829–2853.

[53] M. Morse, G. A. Hedlund, Symbolic dynamics I. Sturmian trajectories. *Amer. J. Math.* 60 (1938), 815–866.

[54] M. Morse, G.A. Hedlund, Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* 62 (1940), 1–42.

[55] W. Parry, On the β-expansions of real numbers, *Acta Math. Acad. Sci. Hungar.* 11 (1960), 401–416.

[56] G. Rauzy, Suites à termes dans un alphabet fini, Sémin. Théor. Nombres, Univ. Bordeaux I 1982–1983, Exp. No. 25, 1983, 16p.

[57] G. Rauzy, Une généralisation du développement en fraction continue, Sémin. Delange-Pisot-Poitou, 18e année: 1976/77, *Théorie des nombres*, Fasc. 1, Exp. No. 15, 16p., Secrétariat Math., Paris 1977.

[58] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hung.* 8 (1957), 477–493.

[59] K. Schmidt, On periodic expansions of Pisot numbers and Salem numbers, *Bull. London Math. Soc.* 12 (1980) 269–278.

[60] L. L. Schumaker, *Spline functions: basic theory*, Wiley, New York 1981.

[61] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, Metallic phase with long range orientational order and no translational symmetry, *Phys. Rev. Lett.* 53 (1984), 1951–1953.

[62] V. Sós, On the distribution mod 1 of the sequence $n\alpha$, *Ann. Univ. Sci. Budapest*, Eötvös Sect. Math. 1 (1958), 127–134.

[63] W. P. Thurston, *Groups, tilings, and finite state automata*, Geometry supercomputer project research report GCG1, University of Minnesota 1989.

[64] S. Yasutomi, On sturmian sequences which are invariant under some substitutions, in *Number theory and its applications*, Kyoto 1997, Dev. Math. 2, Kluwer Acad. Publ., Dordrecht, Dordrecht 1999, 347–373.

Hopf algebras in renormalization theory: locality and Dyson–Schwinger equations from Hochschild cohomology

Christoph Bergbauer[*] *and Dirk Kreimer*[†]

Freie Universität Berlin
Institut für Mathematik II
Arnimallee 2-6, 14195 Berlin, Germany
email: bergbau@math.fu-berlin.de

Institut des Hautes Études Scientifiques
35 route de Chartres, 91440 Bures-sur-Yvette, France
email: kreimer@ihes.fr

Abstract. In this review we discuss the relevance of the Hochschild cohomology of renormalization Hopf algebras for local quantum field theories and their equations of motion.

Contents

[*]FU Berlin and IHES. Supported by DFG grant VO 1272/1-1.
[†]CNRS at IHES and Boston University Center for Mathematical Physics.

Introduction and acknowledgments

The relevance of infinite dimensional Hopf and Lie algebras for the understanding of local quantum field theory has been established in the last couple of years. Here, we focus on the role of the 1-cocycles in the Hochschild cohomology of such renormalization Hopf algebras.

After an introductory overview which recapitulates the well-known Hopf algebra of rooted trees we exhibit the crucial connection between 1-cocycles in the Hochschild cohomology of the Hopf algebra, locality and the structure of the quantum equations of motion. For the latter, we introduce combinatorial Dyson–Schwinger equations and show that the perturbation series provides Hopf subalgebras indexed only by the order of the perturbation. We then discuss assorted applications of such equations which focus on the notion of self-similarity and transcendence.

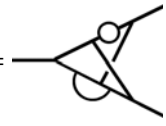
This paper is based on an overview talk given by one of us (D. K.), extended by a more detailed exhibition of some useful mathematical aspects of the Hochschild cohomology of the relevant Hopf algebras. It is a pleasure to thank the organizers of the *75ème Rencontre entre Physiciens Théoriciens et Mathématiciens* for organizing that enjoyable workshop. D. K. thanks Karen Yeats for discussions on the transcendental nature of DSEs. C. B. acknowledges support by the Deutsche Forschungsgemeinschaft under grant VO 1272/1-1. He also thanks Boston University and the IHES for hospitality.

1 Rooted trees, Feynman graphs, Hochschild cohomology and local counterterms

1.1 Motivation

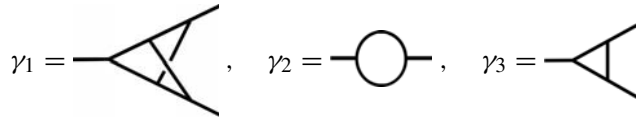
Rooted trees store information about nested and disjoint subdivergences of Feynman graphs in a natural way. This has been used at least implicitly since Hepp's proof of the BPH subtraction formula [23] and Zimmermann's forest formula [38]. However it was only decades later that the algebraic structure of the Bogoliubov recursion was elucidated by showing that it is essentially given by the coproduct and the corresponding antipode of a Hopf algebra on rooted trees [25], [10]. The same result can be formulated more directly in terms of a very similar Hopf algebra on 1PI Feynman graphs [11]. We start with the description in terms of rooted trees which serves as a universal role model for all Hopf algebras of this kind.

For instance, the subdivergences of the ϕ^3 diagram in six spacetime dimensions

$\Gamma =$ 

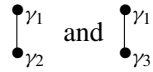
can be represented by the decorated tree

where



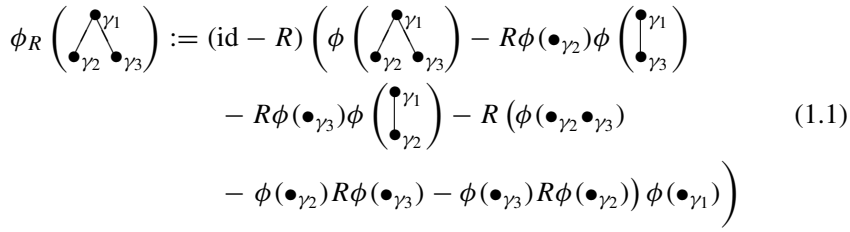
Additional labelling (which we do not care about here) would be needed to keep track of the actual insertion places. However, since one is ultimately interested in the sum of all Feynman graphs of a given order in perturbation theory, for the purpose of the Bogoliubov recursion all possible insertions of γ_2 and γ_3 into γ_1 can be considered at the same time, when due care is given to the resolution of graphs with overlapping divergences into appropriate linear combinations of trees.

In a moment we will need the trees



whose meaning should be clear: They represent the graph γ_1 for which γ_2 or γ_3, respectively, is suitably inserted.

Now how is Γ renormalized? According to the Bogoliubov recursion, the renormalized value is given by



(1.1)

where ϕ denotes the unrenormalized but possibly regularized (if we do not renormalize on the level of the integrand) contribution of the graph which a given tree represents. For example, in dimensional regularization, ϕ is a map into the algebra $V := \mathbb{C}[\epsilon^{-1}, \epsilon]]$ of Laurent series with finite pole part. The map $R\colon V \to V$ is a renormalization scheme. For example, the minimal subtraction scheme is obtained by defining R to be the projector onto the proper pole part, $R(\epsilon^k) = \epsilon^k$ if $k < 0$ and $R(\epsilon^k) = 0$ otherwise. We emphasize though that the use of a regulator can be avoided by defining a suitable renormalization scheme on the level of integrands. Such an approach can then be directly formulated on the level of Dyson–Schwinger equations, where the choice of a renormalization scheme can be non-perturbatively given as the choice of a boundary condition for the accompanying integral equation.

Now consider the polynomial algebra $\mathcal{H}$ generated by all decorated rooted trees of this kind. There is a coproduct on it which disentangles trees into subtrees and thus divergences into subdivergences, as will be discussed in subsection 1.3. Using

this coproduct Δ, the above algebra $\mathcal{H}$ becomes a Hopf algebra. Let S be its antipode. By definition, S satisfies the recursive relation (in Sweedler's notation $\Delta(x) = \mathbb{I} \otimes x + x \otimes \mathbb{I} + \tilde{\sum} x' \otimes x''$)

$$S(x) = -x - \tilde{\sum} S(x')x''.$$

It turns out that, if one similarly defines a "twisted antipode" S_R^ϕ as a map $\mathcal{H} \to V$ by $S_R^\phi(1) = 1$ and

$$S_R^\phi(x) = -R\big(\phi(x) + \tilde{\sum} S_R^\phi(x')\phi(x'')\big),$$

then S_R^ϕ provides the counterterm and the convolution $S_R^\phi \star \phi = m_V(S_R \otimes \phi)\Delta$ solves the Bogoliubov recursion: $\phi_R = S_R^\phi \star \phi$ [25], [10]. Using this algebraic approach to the combinatorial intricacies of renormalization, many important questions in perturbative and non-perturbative quantum field theory can be treated from a convenient conceptual point of view, some of which will be reviewed in the following sections.

This picture translates rather easily to renormalization in coordinate space [1], as will be briefly discussed in subsection 1.6.

Before continuing the discussion of renormalization, we introduce some key algebraic notions. We will come back to the example of the graph Γ later on.

1.2 Basic definitions and notation

Let k be a field of characteristic zero. We consider k-bialgebras $(A, m, \mathbb{I}, \Delta, \epsilon)$ that are graded connected, that is

$$A = \bigoplus_{n=0}^{\infty} A_n, \quad A_0 \cong k, \quad A_m A_n \subseteq A_{m+n}, \quad \Delta(A_n) \subseteq \bigoplus_{l+m=n} A_l \otimes A_m.$$

By abuse of notation, we write $\mathbb{I}$ both for the unit and the unit map. Also, we sometimes consider ϵ as a map $A \to A_0$. We assume that $\Delta(\mathbb{I}) = \mathbb{I} \otimes \mathbb{I}$. It follows that $\epsilon(\mathbb{I}) = 1$ while $\epsilon(A_n) = 0$ for $n \neq 0$. The kernel of ϵ is called the augmentation ideal, and the map $P : A \to A$, $P = \mathrm{id} - \epsilon$, is called the projection onto the augmentation ideal. The coproduct Δ gives rise to another coassociative map: $\tilde{\Delta}$, defined by

$$\tilde{\Delta}(x) = \Delta(x) - \mathbb{I} \otimes x - x \otimes \mathbb{I}.$$

Recall that elements in the kernel of $\tilde{\Delta}$ are called *primitive*. We will occasionally use Sweedler's notation $\Delta(x) = \sum x' \otimes x''$ and also $\tilde{\Delta}(x) = \tilde{\sum} x' \otimes x''$.

It is a well known fact that connected graded bialgebras are Hopf algebras. Indeed, the sequence defined by the recursive relation

$$S(x) = -x - \tilde{\sum} S(x')x'' \text{ for } x \notin A_0, \quad S(\mathbb{I}) = \mathbb{I} \tag{1.2}$$

converges in $\mathrm{End}_k(A)$.

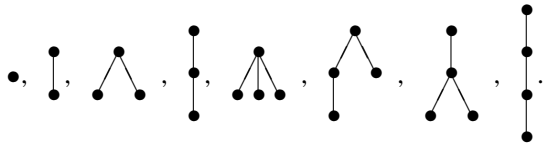
For a coalgebra (A, Δ) and an algebra (B, m), the vector space $\mathrm{Hom}_k(A, B)$ of linear maps $A \to B$ is equipped with a convolution product $\star$ by $(f, g) \mapsto f \star g = m(f \otimes g)\Delta$. Thus $(f \star g)(x) = \sum f(x')g(x'')$. Using the modified product $\star_P : (f, g) \mapsto f \star_P g = m(f \otimes g)(\mathrm{id} \otimes P)\Delta$, equations (1.2) can be rewritten

$$S(x) = -(S \star_P \mathrm{id})(x) \text{ for } x \notin A_0, \quad S(\mathbb{I}) = \mathbb{I}$$

which will be convenient later on.

1.3 The Hopf algebra of rooted trees

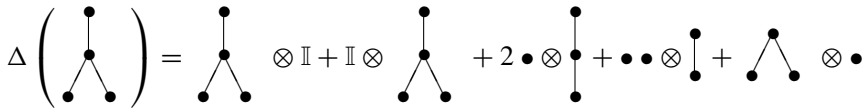
Now we give a more detailed construction of the Hopf algebra $\mathcal{H}$ of rooted trees [25], [10] that is in the center of all our considerations. An (undecorated, non-planar) rooted tree is a connected contractible finite graph with a distinguished vertex called the root. By convention, we will draw the root on top. We are only interested in isomorphism classes of rooted trees (an isomorphism of rooted trees being an isomorphism of graphs which maps the root to the root) which we, by abuse of language, simply call rooted trees again. As a graded algebra, $\mathcal{H}$ is the free commutative algebra generated by trees (including the empty tree which we consider the unit $\mathbb{I}$) with the weight grading: the weight of a tree is the number of its vertices. For instance, the trees of weight one to four are



A product of rooted trees is called a forest – obviously the weight of a forest is the sum of the weights of its trees. On $\mathcal{H}$ a coproduct Δ is introduced by

$$\Delta(\tau) = \mathbb{I} \otimes \tau + \tau \otimes \mathbb{I} + \sum_{\text{adm. c}} P_c(\tau) \otimes R_c(\tau) \tag{1.3}$$

where the sum goes over all *admissible cuts* of the tree τ. By a cut of τ we mean a nonempty subset of the edges of τ that are to be removed. The product of subtrees which "fall down" upon removal of those edges is called the *pruned part* and is denoted $P_c(\tau)$, the part which remains connected with the root is denoted $R_c(\tau)$. This makes sense only for certain "admissible" cuts: by definition, a cut $c(\tau)$ is admissible, if for each leaf l of τ it contains at most one edge on the unique path from l to the root. For instance,



The coassociativity of Δ is shown in [25]. $\mathcal{H}$ is obviously not cocommutative. Since the coproduct is compatible with the grading, $\mathcal{H}$ is a Hopf algebra. There is an

important linear endomorphism of $\mathcal{H}$, the grafting operator B_+ defined as follows:

$$B_+(\mathbb{I}) = \bullet, \qquad B_+(\tau_1 \dots \tau_n) = \underbrace{\quad}_{\tau_1 \ \dots \ \tau_n} \quad \text{for trees } \tau_i. \tag{1.4}$$

In words: B_+ creates a new root and connects it with each root of its argument. The special importance of B_+ will become evident in subsection 1.5: $B_+ \colon \mathcal{H} \to \mathcal{H}$ is a closed but not exact Hochschild 1-cochain.

The Hopf algebra $\mathcal{H}$ is the dual of a Hopf algebra considered earlier by Grossman and Larson [22], see [19]. It can also be described from the free pre-Lie algebra on one generator [8].

1.4 Tree-like structures and variations on a theme

Tree-like structures. From the Hopf algebra $\mathcal{H}$, defined in the previous subsection, several generalizations can be constructed: Hopf algebras of decorated trees, of planar trees, etc. This can be phrased most elegantly from a general point of view in terms of "tree-like structures", as for example introduced by Turaev in [37]: Consider the category of rooted trees and embeddings (an embedding $\tau' \to \tau$ is an isomorphism from τ' to a subtree of τ). A rooted tree-structure is then defined to be a contravariant functor from this category to the category of sets. For example, decorated (labelled) trees can be described by the functor ϕ which maps a tree onto a certain set its vertices and/or edges are decorated with. Being contravariant, ϕ maps embeddings of trees to the respective restrictions of decorations. Similarly, a planar structure is provided by a functor ϕ mapping a tree to the set of its topological embeddings into the real plane modulo orientation-preserving homeomorphisms of $\mathbb{R}^2$ onto itself. Now let ϕ be a rooted tree-structure. A rooted ϕ-tree is a pair (τ, s) where τ is a tree and s is an element of $\phi(\tau)$. The notions of isomorphisms and subtrees of rooted ϕ-trees are immediate.

Generalizations of $\mathcal{H}$. Using this convenient framework, we have immediately other Hopf algebras at hand: Let S be a set. The Hopf algebra $\mathcal{H}(S)$ is defined as in the previous subsection, replacing the word tree by S-decorated tree (for our purposes, we only decorate vertices, not edges). Similarly, $\mathcal{H}_{\mathrm{pl}}$ is the (noncommutative) Hopf algebra of planar rooted trees. In particular, for these Hopf algebras, the proofs of the coassociativity of Δ are verbatim the same. The planar Hopf algebra and its decorated versions $H_{\mathrm{pl}}(S)$ were extensively studied by Foissy [19]. He showed that they are self-dual and constructed isomorphisms to several other Hopf algebras on trees that have appeared in the literature.

The Hopf algebra of Feynman graphs. While rooted trees describe nested divergences in an obvious manner, the resolution of *overlapping* divergences into trees requires some care [38], [26], [18]. This problem exists only in momentum space.

By basing a Hopf algebra directly on Feynman graphs instead of trees, these issues can be avoided [26], [11]. As an algebra, let $\mathcal{H}_{\mathrm{CK}}$ be the free commutative algebra on 1PI Feynman graphs (of a given theory; the case of a non-scalar theory requires to take form factors (external structures) into account which we avoid here). The empty graph serves as a unit $\mathbb{I}$. In the following, a product of graphs is identified with the disjoint union of these graphs. On a graph, a coproduct is given [11] by

$$\Delta(\Gamma) = \mathbb{I} \otimes \Gamma + \Gamma \otimes \mathbb{I} + \sum_{\gamma \subsetneq \Gamma} \gamma \otimes \Gamma/\gamma$$

where the sum is over all 1PI superficially divergent proper subgraphs γ of Γ. A few examples are given in [11].

Still, thanks to the universal property mentioned at the end of subsection 1.5, Hopf algebras of rooted trees serve as an excellent role model for various questions and, moreover, yield most interesting links to different branches of mathematics [13], [21]. In the present paper, we will be mainly concerned with Hopf algebras of trees. In many cases, it is only a matter of notation to translate these results into the Hopf algebra of Feynman graphs, easily achieved by the practitioner of QFT [4].

In view of the preceding paragraphs, the reader might wish to try to describe $\mathcal{H}_{\mathrm{CK}}$ as a Hopf algebra of suitable tree-like structures, using the results of [26].

1.5 Hochschild cohomology of bialgebras

Definition. Let A be a bialgebra. We consider linear maps $L\colon A \to A^{\otimes n}$ as n-cochains and define a coboundary operator b by

$$bL := (\mathrm{id} \otimes L)\Delta + \sum_{i=1}^{n} (-1)^i \Delta_i L + (-1)^{n+1} L \otimes \mathbb{I} \tag{1.5}$$

where Δ denotes the coproduct and Δ_i the coproduct Δ applied to the i-th factor in $A^{\otimes n}$. The map $L \otimes \mathbb{I}$ is given by $x \mapsto L(x) \otimes \mathbb{I}$. It is essentially due to the coassociativity of Δ that b squares to zero, which gives rise to a cochain complex (C, b). Clearly (C, b) captures only information about the coalgebra structure of A. The cohomology of (C, b), denoted $HH^\bullet_\epsilon(A)$, is easily seen to be the dual (A considered as a bi*co*module rather than a bimodule over itself) notion of the Hochschild cohomology of algebras. Note that the right bicomodule action is here $(\mathrm{id} \otimes \epsilon)\Delta$ which explains the last summand in (1.5) and the subscript in $HH^\bullet_\epsilon$.

The role of $HH^1_\epsilon(\mathcal{H})$. For $n = 1$, the cocycle condition $bL = 0$ reduces to, for $L : A \to A$,

$$\Delta L = (\mathrm{id} \otimes L)\Delta + L \otimes \mathbb{I}. \tag{1.6}$$

Sometimes the following equivalent statement, using the map $\tilde{\Delta}$, is more convenient:

$$\tilde{\Delta} L = (\mathrm{id} \otimes L)\tilde{\Delta} + \mathrm{id} \otimes L(\mathbb{I}). \tag{1.7}$$

Let us now try to understand the space $HH^1_\epsilon(\mathcal{H})$ of "outer coderivations on $\mathcal{H}$". We first describe the 0-coboundaries ("inner coderivations"). They are of the form

$$L(\tau) = \sum \alpha_{\tau''} \tau' - \alpha_\tau \mathbb{I}$$

in Sweedler's notation, where α_τ is an element of k for each forest τ. For example, $L: \tau \mapsto \sum \tau' - \mathbb{I}$ is a 0-coboundary. Note that $\mathbb{I}$ is in the kernel of any 0-coboundary.

It is a crucial fact that the grafting operator B_+, introduced in subsection 1.3, is a 1-cocycle [10]:

Proposition 1.1.

$$\Delta B_+ = (\mathrm{id} \otimes B_+)\Delta + B_+ \otimes \mathbb{I}; \tag{1.8}$$

$$\tilde{\Delta} B_+ = (\mathrm{id} \otimes B_+)\tilde{\Delta} + \mathrm{id} \otimes \bullet. \tag{1.9}$$

Idea of proof. When looking at equation (1.9), the statement is rather immediate: Let τ be a forest. The first summand at the right side of (1.9) refers to cuts of $B_+(\tau)$ which affect at most all but one of the edges connecting the new root of $B_+(\tau)$ to the roots of τ, while the second summand takes care of the cut which completely separates the root of $B_+(\tau)$ from all its children. □

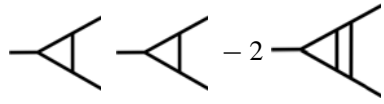
Since B_+ is a homogeneous linear endomorphism of degree 1, it is not a 0-coboundary – note that the coboundaries have no chance to increase the degree. Thus B_+ is a generator (among others) of $HH^1_\epsilon(\mathcal{H})$.

When looking for other generators L of $HH^1_\epsilon(\mathcal{H})$, the cocycle conditions (1.6), (1.7) immediately yield the requirement that $L(\mathbb{I})$ be a primitive element (and zero if L is exact). While $\bullet$ is up to scalar factors obviously the only primitive element in degree 1, there are plenty of primitives in higher degrees. For example,

(1.10)

is a primitive element in degree 2. Foissy [19] showed that $L \mapsto L(\mathbb{I})$ is a surjective map $HH^1_\epsilon(\mathcal{H}) \to \mathrm{Prim}(\mathcal{H})$ onto the set of primitive elements of $\mathcal{H}$. In the case of Hopf algebras of decorated rooted trees $\mathcal{H}(S)$ obviously any element $s \in S$ yields a homogeneous cocycle of degree 1 denoted B^s_+ which, applied to a forest, connects its roots to a new root decorated by s.

It should be clear that each 1PI Feynman graph which is free of subdivergences is a primitive element of $\mathcal{H}_{\mathrm{CK}}$. In general, there are primitive elements in higher degrees too, for example, cf. (1.10), the linear combination



in ϕ^3 theory in six dimensions.

Universal property. The category of objects (A, L) consisting of a commutative bialgebra A and a Hochschild 1-cocycle L on A with morphisms bialgebra morphisms commuting with the cocycles has the initial object $(\mathcal{H}, B_+)$. This is a result of [10]. Indeed, let (A, L) be such a pair. The map $\rho\colon \mathcal{H} \to A$ is simply defined by $\rho(\mathbb{I}) = \mathbb{I}$ and pushing forward along B_+ (and L) and the multiplication. The fact that ρ is a morphism of coalgebras is an easy consequence of (1.8).

Also it was shown in [1] that, conversely, the coproduct Δ of $\mathcal{H}$ is determined if one requires the map B_+ to be a 1-cocycle. This may serve to find different presentations of $\mathcal{H}$.

For any $\mathcal{H}$-bicomodule B, the higher Hochschild cohomology $HH^n(\mathcal{H}, B)$, $n \geq 2$, is trivial [19], thus in particular $HH^n_\epsilon(\mathcal{H}) = 0$.

1.6 Finiteness and locality from the Hopf algebra

We have now accumulated enough algebraic notions to come back to the original physical application already sketched in subsection 1.1. Given a specific quantum field theory, Hopf algebras $\mathcal{H}(S)$ and $\mathcal{H}_{\mathrm{CK}}$ are determined by its perturbative expansion into Feynman graphs. We denote this Hopf algebra generically by $\mathcal{H}$. Every divergent graph γ without subdivergences determines a Hochschild 1-cocycle B_+^γ, and any relevant tree or graph is in the range of a 1-cocycle of this kind. This ensures that any relevant term in the perturbative expansion is in the image of a Hochschild one-cocycle. This allows to prove locality: the all important Bogoliubov $\bar{R}$ operation on a character ϕ,

$$\phi \to \bar{R}[\phi] = m(S_R^\phi \otimes \phi)(\mathrm{id} \otimes P)\Delta,$$

P the projector in the augmentation ideal and ϕ some Feynman rules has the property that it only requires an overall Taylor subtraction at a fixed renormalization point to render it finite – implying that its divergences will not depend on logarithms of kinematical variables.

Momentum space. The next step is to choose a target algebra V and regularized Feynman rules $\phi\colon \mathcal{H} \to V$, and a renormalization scheme $R\colon V \to V$. The map ϕ is supposed to be a (unital) algebra homomorphism. We stick to the example $(V = \mathbb{C}[\epsilon^{-1}, \epsilon]], \phi)$ of dimensional regularization as in subsection 1.1, but stress once more that the reader can find suitable generalizations in the literature [16]. The minimal subtraction scheme where R is the projector onto the proper pole part is only one of many choices one can make. However, in any case we require R to preserve the UV divergent structure (i.e. the pole part) and to satisfy the Rota–Baxter equation

$$R(xy) + R(x)R(y) = R(xR(y)) + R(R(x)y). \tag{1.11}$$

It is easy to check that the minimal subtraction scheme satisfies (1.11). The Rota–Baxter equation is the algebraic key to the link between renormalization and Birkhoff decomposition, see for example [11], [15], [16]. It also guarantees that the renormalized Feynman rules are again an algebra homomorphism [27] as are the unrenormalized rules ϕ. Now the twisted antipode is defined by

$$S_R^\phi(\tau) = -R(S_R^\phi \star_P \phi)(\tau) \quad \text{for } \tau \notin \mathcal{H}_0, \quad S_R^\phi(\mathbb{I}) = 1, \tag{1.12}$$

equivalently, in Sweedler's notation

$$S_R^\phi(\tau) = -R\big(\phi(\tau) + \tilde{\sum} S_R^\phi(\tau')\phi(\tau'')\big) \quad \text{for } \tau \notin \mathcal{H}_0, \quad S_R^\phi(\mathbb{I}) = 1$$

where the term "twisted antipode" should be justified by a glance at the recursive expression (1.2) for the regular antipode. The map S_R^ϕ, as can be inferred from the example in Figures 1–3, yields the counterterm for ϕ. The complete renormalized evaluation function is then given by

$$\phi_R = S_R^\phi \star \phi. \tag{1.13}$$

One can find a non-recursive description of ϕ_R [25], [10] which shows the equivalence with Zimmermann's forest formula [38].

Example. In order to understand the twisted antipode, we come back to the example of subsection 1.1. On the relevant trees, the coproduct acts as follows:

$$\Delta\left(\begin{array}{ccc} & \gamma_1 & \\ \gamma_2 & & \gamma_3 \end{array}\right) = \mathbb{I} \otimes \begin{array}{ccc} & \gamma_1 & \\ \gamma_2 & & \gamma_3 \end{array} + \begin{array}{ccc} & \gamma_1 & \\ \gamma_2 & & \gamma_3 \end{array} \otimes \mathbb{I} + \bullet_{\gamma_2} \otimes \begin{array}{c} \gamma_1 \\ \gamma_3 \end{array} + \bullet_{\gamma_3} \otimes \begin{array}{c} \gamma_1 \\ \gamma_2 \end{array} + \bullet_{\gamma_2}\bullet_{\gamma_3} \otimes \bullet_{\gamma_1},$$
$$\Delta(\bullet_{\gamma_i}) = \mathbb{I} \otimes \bullet_{\gamma_i} + \bullet_{\gamma_i} \otimes \mathbb{I}. \tag{1.14}$$

According to (1.12) and (1.13), the algorithm for ϕ_R consists of the following steps:

- (F) apply the coproduct Δ to the tree under consideration;
- (C_n) apply the map $(\mathrm{id} \otimes P)\Delta \otimes \mathrm{id}^{\otimes n}$ (for $n = 1, \dots$) until each summand is of the form $\mathbb{I} \otimes \cdots$;
- (M) apply $\phi^{\otimes n}$ to go into $V^{\otimes n}$, as $\phi(\mathbb{I}) = S_R^\phi(\mathbb{I}) = 1$, the first factor $\mathbb{I}$ of each term is mapped to 1;
- (C_n') (for $n = \dots 1$) apply the map $-Rm \otimes \mathrm{id}^{\otimes n}$ until we end up in $V^{\otimes 2}$;
- (F') apply the map m to get into V.

For the tree $\begin{array}{ccc} & \gamma_1 & \\ \gamma_2 & & \gamma_3 \end{array}$ this algorithm is performed in Figures 1–3.

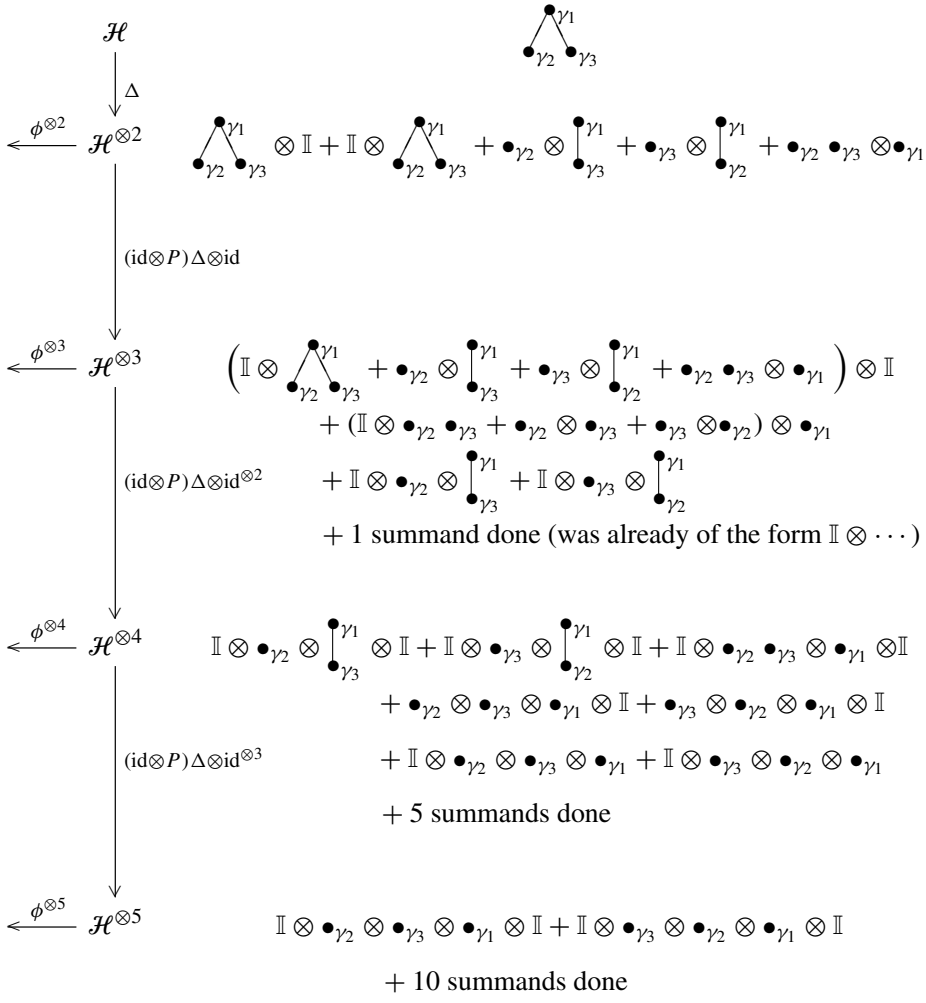


Figure 1. First part of the calculation of ϕ_R. Apply Δ and then $(\mathrm{id}\otimes P)\Delta\otimes\mathrm{id}^{\otimes n}$ until each summand is of the form $\mathbb{I}\otimes\cdots$.

While in our simple example of only two disjoint subdivergences, the Bogoliubov recursion could have been performed by hand without using the Hopf algebra, when going to higher loop orders, the Hopf algebra approach provides significant computational advantage [4], [5].

Locality of counterterms from Hochschild cohomology. Moreover, the Hopf algebra can be used to give a direct proof of finiteness of renormalization and locality of counterterms from a purely algebraic point of view. For a simple toy model, this has been done in a recent paper [28]. The basic observation is that the fact that every

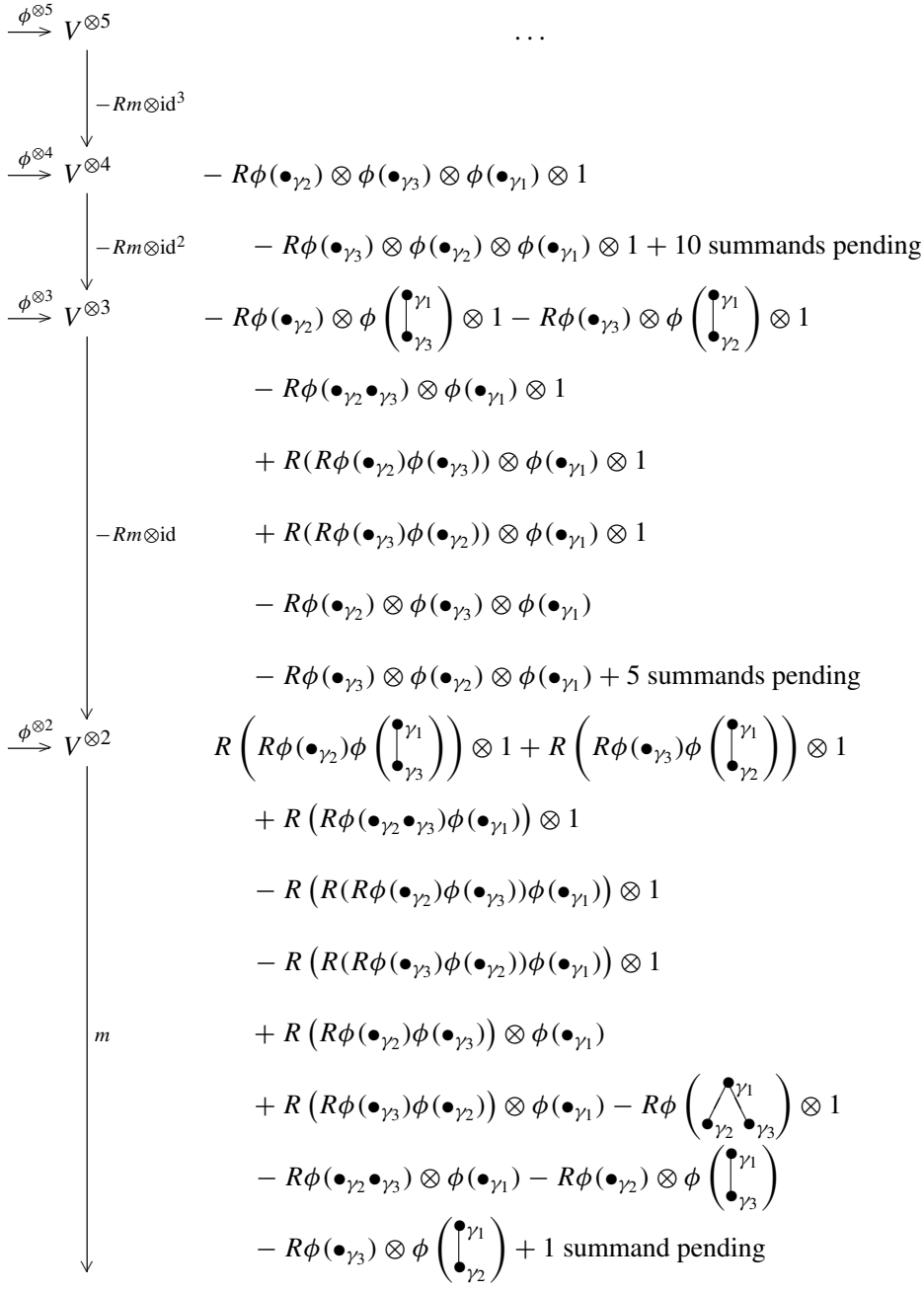


Figure 2. Second part of the calculation of ϕ_R. Apply $\phi^{\otimes n}$ and then $-Rm\otimes \mathrm{id}^{\otimes n}$ until arrival in $V^{\otimes 2}$. Then apply m to get into V.

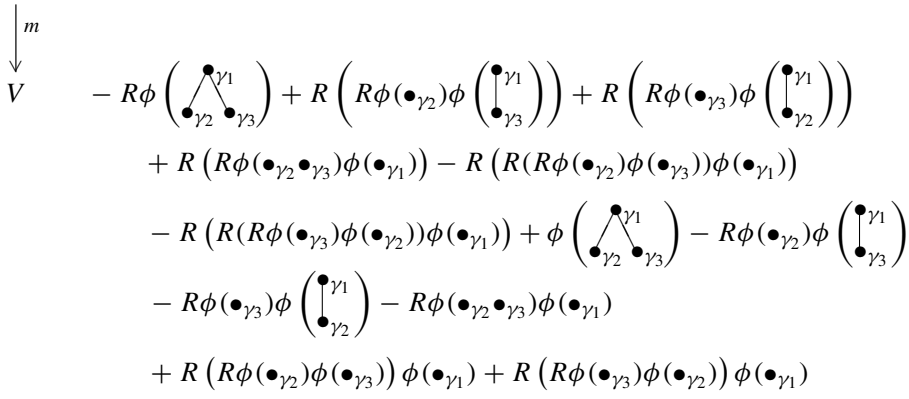


Figure 3. Third part of the calculation of ϕ_R. The reader should compare the result with (1.1). Using the fact that S_R^ϕ is an algebra homomorphism (if R is a Rota–Baxter map), the last step (C3) in Figure 1 and the first step (C3') in Figure 2 could have been avoided.

relevant tree or graph is in the range of a homogeneous Hochschild 1-cocycle of degree 1 allows for easy and clean inductive proofs of various statements for arbitrary loop number. This also holds on the level of graphs: the sum over all primitive graphs of given loop order n defines a 1-cocycle B_+^n such that every graph is generated in the range of these 1-cocycles. This allows, as observed in [29], [33], to prove locality in general.

Indeed, let B_+^n be a 1-cocycle, and let μ_+ be the measure defined by the n-loop integrand of $B_+^n(\mathbb{I})$. Let $\phi(B_+^n(X))$ be a Feynman amplitude defined by insertion of a collection of subdivergences X into those n-loop primitive graphs. We write $\phi(B_+^n(X)) = \int \phi(X) d\mu_+$, emphasizing that subgraphs become subintegrals under the Feynman rules.

Recall that P denotes the projection onto the augmentation ideal. Since $P(\mathbb{I}) = 0$, $PB_+^n = B_+^n$, we can write

$$m(S_R^\phi \otimes \phi P)\Delta(B_+^n(X)) = \int S_R^\phi \star \phi(X) d\mu_+.$$

This proves locality in a straightforward manner by induction over the augmentation degree, i.e. using the coradical filtration of the Hopf algebra.

Coordinate space. The language of rooted trees is especially suited for describing renormalization in coordinate space [1]. A particularly appealing approach to coordinate space renormalization is the work of Epstein and Glaser [17] (see also [36], [7]) who, starting from ideas of Bogoliubov [3] and others, extracted a set of axioms for time-ordered product and constructed such time-ordered products in terms of rigorous functional analysis. The result is completely equivalent to momentum space renor-

malization but has conceptual (albeit not computational) advantages. It is no surprise that the Hopf algebra picture fits equally nice into this framework [1], if one takes into account the specific features of Epstein–Glaser renormalization such as the absence of overlapping divergences and regularization parameters. In view of highly interesting mathematical ramifications such as a possible analogy to the Fulton–MacPherson compactification of configuration spaces [20], [34], it seems most appropriate to attack this problem using trees [1] rather than coordinate space Feynman diagrams.

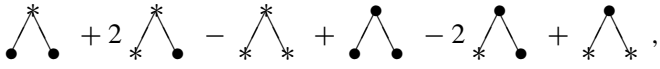
If there are no subdivergences, Epstein–Glaser renormalization amounts to a Taylor subtraction on test functions: Let ${}^0t$ be a distribution on some $\mathbb{R}^d - \{0\}$ with singularity at 0, for example ${}^0t = x^{-(d+1)}$. In order to extend ${}^0t$ onto all of $\mathbb{R}^d$, consider

$$t : f \mapsto {}^0t\Big(f - \sum_{|\alpha| \leq \rho} w_\alpha \partial^\alpha f(0)\Big) \tag{1.15}$$

where f is a test function and the w_α are auxiliary test functions with $\partial^\beta w_\alpha(0) = \delta_\alpha^\beta$. If ρ is large enough with respect to the degree of divergence of ${}^0t$ at 0, the modified distribution t is defined on all of $\mathbb{R}^d$. It is natural to consider the first summand in (1.15) as the unrenormalized contribution and the second summand as the counterterm.

Epstein and Glaser describe how to take care of distributions which may have an overall divergence and subdivergences, i.e. distributions which are not only singular on the thin diagonal $\{x_1 = \cdots = x_n\}$ of some M^n but on the fat diagonal $\{x_i = x_j \text{ for some } i, j\}$. The algorithm, with the above identification of unrenormalized part and counterterm, is structurally very similar to the Bogoliubov recursion and can thus be described by a twisted antipode [1].

Using some techniques of [1], notably the "cut product" $\odot$ of certain linear endomorphisms on $\mathcal{H}$ as a replacement for the convolution product, one can construct the map R as an algebra endomorphism on a Hopf algebra of trees with decorated vertices $\mathcal{H}(\{\bullet, *\})$ and consider the "twisted antipode" defined by $S_R = -R(S_R \odot_P \mathrm{id})$, and the renormalization map $S_R \odot \mathrm{id}$. Starting from the tree  , the map $S_R \odot \mathrm{id}$ yields



which should be compared to the last line in Figure 3 – vertices of type $\bullet$ mark unrenormalized contributions, vertices of type $\star$ the corresponding counterterms. These trees are then mapped into an appropriate space of operator-valued distributions: the above example describes terms needed for the renormalization of the fourth order time-ordered product. Using this somewhat modified approach, where the combinatorics happen entirely in the Hopf algebra (as opposed to between the Hopf algebra and the target ring), checking locality simply amounts to calculating the commutator of $S_R \odot \mathrm{id}$ and $B_{+\bullet}$:

$$(S_R \odot \mathrm{id}) B_{+\bullet} = (\mathrm{id} - R) B_{+\bullet} (S_R \odot \mathrm{id}).$$

Thus once the subdivergences are taken care of, it suffices to subtract the superficial divergence.

2 Hopf subalgebras and Dyson–Schwinger equations

Hopf subalgebras of the Hopf algebras of (decorated) rooted trees or Feynman graphs are in close relationship with Dyson–Schwinger equations. Indeed, any Dyson–Schwinger equation (to be defined below) gives rise to a Hopf subalgebra. This is a statement about self-similarity: a 1-cocycle like B_+ ensures that a product of trees is mapped to a tree, and this is a rather general phenomenon: the Green functions appear as functionals of themselves, the functionals being provided by the Dyson skeleton graphs which appear as the integral kernel $\phi(B_+^n(\mathbb{I}))$.

It will turn out in Theorem 2.2 that all Hopf subalgebras coming from a reasonably general class of Dyson–Schwinger equations are in fact isomorphic.

2.1 Hopf subalgebras of decorated rooted trees

For simplicity, we start our considerations in the Hopf algebra $\mathcal{H}$ of undecorated rooted trees. A full classification of their Hopf subalgebras is far beyond reach. However, we give a few examples the last of which will be directly related to Dyson–Schwinger equations.

Bounded fertility, finite parts, primitive elements. For $n \in \mathbb{N}$ let $\mathcal{H}_n$ be the subalgebra of $\mathcal{H}$ generated by trees whose vertices have fertility bounded from above by n. A glance at the definition of the coproduct (1.3) suffices to see that $\mathcal{H}_n$ is a Hopf subalgebra of $\mathcal{H}$. In particular, the Hopf algebra $\mathcal{H}_1$ with one generator in each degree is known as the Hopf algebra of *ladders*. It is closely related to iterated integrals [9], [27].

Similarly, the free commutative algebra generated by trees of degree $\leq n$ forms a Hopf subalgebra for any n since the coproduct respects the grading. Another example where there is nothing to check are subalgebras generated by an arbitrary collection of primitive elements of $\mathcal{H}$.

The Connes–Moscovici Hopf subalgebra. A less trivial example of a Hopf subalgebra of $\mathcal{H}$ arose in the work of Connes and Moscovici on local index formulas for transversally hypoelliptic operators on foliations [13], [10], [14]. In the case of a foliation of codimension 1, the relevant Hopf algebra $\mathcal{H}_T$ is defined by the generators X, Y, δ_n for $n \in \mathbb{N}$, the relations

$$[X, Y] = -X, \quad [X, \delta_n] = \delta_{n+1}, \quad [Y, \delta_n] = n\delta_n, \quad [\delta_n, \delta_m] = 0,$$

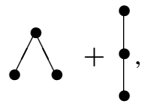
and the coproduct

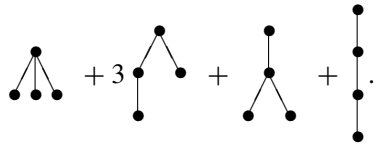
$$\Delta(X) = X \otimes \mathbb{I} + \mathbb{I} \otimes X + \delta_1, \quad \Delta(Y) = Y \otimes \mathbb{I} + \mathbb{I} \otimes Y, \quad \Delta(\delta_1) = \delta_1 \otimes \mathbb{I} + \mathbb{I} \otimes \delta_1.$$

Note that the relations above and the requirement that Δ be an algebra homomorphism determine Δ on the generators δ_n for $n \geq 2$ as well. Let N be the linear operator, called *natural growth operator*, on $\mathcal{H}$, defined on a tree τ by adding a branch to each vertex of τ and summing up the resulting trees, extended as a derivation onto all of $\mathcal{H}$. For example,

$N(\mathbb{I}) =$  ,

$N^2(\mathbb{I}) =$  ,

$N^3(\mathbb{I}) =$  , (2.1)

$N^4(\mathbb{I}) =$  .

Now identifying δ_1 with $\bullet$, and generally δ_n with $N^n(\mathbb{I})$, the commutative Hopf subalgebra of $\mathcal{H}_T$ generated by the δ_n can be embedded into $\mathcal{H}$ [10]. The resulting Hopf subalgebra is denoted $\mathcal{H}_{\mathrm{CM}}$. For example,

$$\tilde{\Delta}(\delta_1) = 0,$$
$$\tilde{\Delta}(\delta_2) = \delta_1 \otimes \delta_1,$$
$$\tilde{\Delta}(\delta_3) = 3\delta_1 \otimes \delta_2 + (\delta_2 + \delta_1^2) \otimes \delta_1.$$

The δ_n can be specified in a non-recursive manner:

$$\delta_n = \sum_{\tau \in \mathcal{T}_n} c_\tau \tau.$$

Here $\mathcal{T}_n$ is the set of trees of weight n. The integers c_τ, called *Connes–Moscovici weights*, have been computed in [27], [19] using the tree factorial

$$c_\tau = \frac{n!}{\tau! \operatorname{Sym}(\tau)}$$

where $\operatorname{Sym}(\tau)$ is the symmetry factor (rank of the group of symmetries) of τ.

A quadratic Dyson–Schwinger equation. Now we turn to the study of another source of Hopf subalgebras, the combinatorial Dyson–Schwinger equations. As a first example, we consider the equation

$$X = \mathbb{I} + \alpha B_+(X^2) \tag{2.2}$$

in $\mathcal{H}[[\alpha]]$. Using the ansatz

$$X = \sum_{n=0}^{\infty} \alpha^n c_n$$

one easily finds $c_0 = \mathbb{I}$ and

$$c_{n+1} = \sum_{k=0}^{n} B_+(c_k c_{n-k}) \tag{2.3}$$

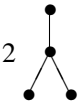
which determine X by induction. The first couple of c_n are easily calculated:

$c_0 = \mathbb{I}$,

$c_1 =$,

$c_2 = 2$  ,

$c_3 =$  $+ 4$  ,

$c_4 = 4$  $+ 2$  $+ 8$  .

We observe that c_n is a weighted sum of trees with vertex fertility bounded by 2 – this is due to the square of X in the Dyson–Schwinger equation (2.2). The reader should compare this to the Connes–Moscovici trees (2.1) discussed in the previous subsection. The recursive nature of (2.2) makes one suspect that the c_n generate a Hopf subalgebra of $\mathcal{H}$. Indeed, for each $n \geq 0$ and $k \leq n$ there is a polynomial P_k^n in the c_l for $l \leq n$ such that

$$\Delta c_n = \sum_{k=0}^{n} P_k^n \otimes c_k. \tag{2.4}$$

They are inductively determined by

$$P_{k+1}^{n+1} = \sum_{l=0}^{n-k} P_0^l P_k^{n-l} \tag{2.5}$$

and $P_0^{n+1} = c_{n+1}$. For a proof of this statement, see the more general Theorem 2.2 in the next subsection.

For the moment, we merely display the first P_k^n in an upper triangular matrix where columns are indexed by $n = 0, \ldots, 5$ and rows by $k = 0, \ldots, n$:

$$\begin{bmatrix} \mathbb{I} & c_1 & c_2 & c_3 & c_4 & c_5 \\ & \mathbb{I} & 2c_1 & 2c_2 + c_1^2 & 2c_3 + 2c_1c_2 & 2c_4 + 2c_1c_3 + c_2^2 \\ & & \mathbb{I} & 3c_1 & 3c_2 + 3c_1^2 & 6c_1c_2 + c_1^3 + 3c_3 \\ & & & \mathbb{I} & 4c_1 & 6c_1^2 + 4c_2 \\ & & & & \mathbb{I} & 5c_1 \\ & & & & & \mathbb{I} \end{bmatrix}.$$

The coefficients are basically multinomial coefficients as will become clear in the next subsection.

2.2 Combinatorial Dyson–Schwinger equations

Let A be any connected graded Hopf algebra which is free or free commutative as an algebra, and $(B_+^{d_n})_{n\in\mathbb{N}}$ a collection of Hochschild 1-cocycles on it (not necessarily pairwise distinct). The most general Dyson–Schwinger equation we wish to consider here is

$$X = \mathbb{I} + \sum_{n=1}^{\infty} \alpha^n w_n B_+^{d_n}(X^{n+1}) \tag{2.6}$$

in $A[[\alpha]]$. The parameter α plays the role of a coupling constant. The w_n are scalars in k. Again we decompose the solution

$$X = \sum_{n=0}^{\infty} \alpha^n c_n$$

with $c_n \in A$.

Lemma 2.1. *The Dyson–Schwinger equation* (2.6) *has a unique solution described by* $c_0 = \mathbb{I}$ *and*

$$c_n = \sum_{m=1}^{n} w_m B_+^{d_m}\Big(\sum_{\substack{k_1+\cdots+k_{m+1}=n-m, \\ k_i\geq 0}} c_{k_1} \ldots c_{k_{m+1}} \Big). \tag{2.7}$$

Proof. Inserting the ansatz into (2.6) and sorting by powers of α yields the result. Uniqueness is obvious. □

Theorem 2.2. *The elements* c_n *generate a Hopf subalgebra of* A:

$$\Delta(c_n) = \sum_{k=0}^{n} P_k^n \otimes c_k$$

where the P_k^n are homogeneous polynomials of degree $n-k$ in the c_l, $l \leq n$:

$$P_k^n = \sum_{l_1+\cdots+l_{k+1}=n-k} c_{l_1} \ldots c_{l_{k+1}}. \tag{2.8}$$

In particular, the P_k^n are independent of the w_n and $B_+^{d_n}$.

We emphasize that the main ingredient for the proof of this theorem is the fact that the $B_+^{d_n}$ are Hochschild 1-cocycles, the rest being a cumbersome but straightforward calculation.

Proof. We proceed by proving inductively the following statements.

(α_n) The theorem holds up to order n.

(β_n) For a given $m \in \{1, \ldots, n\}$ let $l_1 + \cdots + l_{m+1} =: p \in \{0, \ldots, n-m\}$, $l_i \geq 0$. Then the right hand sum

$$P(n-m, m, p) := \sum_{\substack{k_1+\cdots+k_{m+1}=n-m \\ k_i \geq l_i}} P_{l_1}^{k_1} \ldots P_{l_{m+1}}^{k_{m+1}} \tag{2.9}$$

does not depend on the single l_i but only on p, $n-m$ and m, justifying the notation $P(n-m, m, p)$.

(γ_n) In the above notation and for any $q \in \{1 \ldots n\}$, the term $P(n-m, m, q-m)$ does not depend on $m \in \{1, \ldots, q\}$.

To start the induction, we note that (α_0) is obvious. (β_1) is trivial as $m = 1$ enforces $l_1 = l_2 = 0$. Similarly, for (γ_1) only one m is in range and the statement thus trivially satisfied. We proceed to (α_n). By definition, and using (1.6) for the $B_+^{d_n}$,

$$\Delta(c_n) = \sum_{m=1}^{n} w_m((\mathrm{id} \otimes B_+^{d_m})\Delta + B_+^{d_m} \otimes \mathbb{I})\Big(\sum_{\substack{k_1+\cdots+k_{m+1}=n-m \\ k_i \geq 0}} c_{k_1} \ldots c_{k_{m+1}} \Big)$$

(using the induction hypothesis (α_{n-1}))

$$= c_n \otimes \mathbb{I} + \sum_{m=1}^{n} w_m (\mathrm{id} \otimes B_+^{d_m}) \sum_{\substack{k_1+\cdots+k_{m+1}=n-m \\ k_i \geq 0}} \sum_{l_1 \ldots l_{m+1}=0}^{k_1 \ldots k_{m+1}} P_{l_1}^{k_1} \ldots P_{l_{m+1}}^{k_{m+1}} \otimes c_{l_1} \ldots c_{l_{m+1}}$$

(by rearranging indices)

$$= c_n \otimes \mathbb{I} + \sum_{m=1}^{n} w_m \sum_{p=0}^{n-m} \sum_{l_1+\cdots+l_{m+1}=p} \sum_{\substack{k_1+\cdots+k_{m+1}=n-m \\ k_i \geq l_i}} P_{l_1}^{k_1} \ldots P_{l_{m+1}}^{k_{m+1}} \otimes B_+^{d_m}(c_{l_1} \ldots c_{l_{m+1}})$$

(by the induction hypothesis (β_n) and using the notation of (2.9))

$$= c_n \otimes \mathbb{I} + \sum_{m=1}^{n} w_m \sum_{p=0}^{n-m} P(n-m, m, p) \otimes \sum_{l_1+\cdots+l_{m+1}=p} B_+^{d_m}(c_{l_1} \dots c_{l_{m+1}})$$

(rearranging indices (q replaces $m+p$) and using (γ_n))

$$= c_n \otimes \mathbb{I} + \sum_{q=1}^{n} \sum_{m=1}^{q} w_m P(n-m, m, q-m) \otimes \sum_{l_1+\cdots+l_{m+1}=q-m} B_+^{d_m}(c_{l_1} \dots c_{l_{m+1}})$$

$$= c_n \otimes \mathbb{I} + \sum_{q=1}^{n} P(n-q, q, 0) \otimes \sum_{m=1}^{q} w_q \sum_{l_1+\cdots+l_{m+1}=q-m} B_+^{d_m}(c_{l_1} \dots c_{l_{m+1}}).$$

Since the right hand tensor factor is c_q, a glance at (2.9), using that $P_0^k = c_k$, verifies (α_n).

The items (β_n) and (γ_n) follow from (α_{n-1}):

$$\begin{aligned} P(n-m, m, p) &= \sum_{\substack{k_1+\cdots+k_{m+1}=n-m \\ k_i \geq l_i}} P_{l_1}^{k_1} \dots P_{l_{m+1}}^{k_{m+1}} \\ &= \sum_{\substack{k_1+\cdots+k_{m+1}=n-m \\ k_i \geq l_i}} \sum_{r_1^1+\cdots+r_{l_1+1}^1=k_1-l_1} \cdots \\ &\quad \cdots \sum_{r_1^{m+1}+\cdots+r_{l_{m+1}+1}^{m+1}=k_{m+1}-l_{m+1}} c_{r_1^1} \dots c_{r_{l_{m+1}+1}^{m+1}} \\ &= \sum_{r_1+\cdots+r_{m+p+1}=n-m-p} c_{r_1} \dots c_{r_{m+p+1}}, \end{aligned}$$

which is independent of any l_i whence (β_n). Substituting $p = q - m$ shows (γ_n). □

At first sight the fact that the coproduct on the c_i does not depend on the w_k and hence that all Dyson–Schwinger equations of this kind yield isomorphic Hopf subalgebras (provided there are no relations among the c_n) might well come as a surprise. The deeper reason for this is the recursiveness of (2.6) as will become more apparent in the next paragraphs.

Description in terms of trees. Now we specialize to the case $A = \mathcal{H}(S)$ where $(S = \dot{\bigcup} S_n, |\cdot|)$ is an arbitrary graded set of decorations such that $|d_n| = n$ for all n (one can even allow $d_n \subset S_n$ and define $B_+^{d_n} := \sum_{\delta \in d_n} B_+^{\delta}$). The maps $B_+^{d_n}$ are defined as in (1.4) where the newly created vertex is decorated by d_n. Using the following lemma, which gives an explicit presentation of the c_i in terms of trees, Theorem 2.2 can be proven in a more comprehensive way.

Lemma 2.3. *The solution of* (2.6) *can be described by* $c_0 = \mathbb{I}$ *and*

$$c_n = \sum_{\tau \in \mathcal{T}(S),\, |\tau| = n} \frac{\tau}{\mathrm{Sym}(\tau)} \prod_{v \in \tau^{[0]}} \gamma_v \tag{2.10}$$

where

$$\gamma_v = \begin{cases} w_{|\mathrm{dec}(v)|} \frac{(|\mathrm{dec}(v)|+1)!}{(|\mathrm{dec}(v)|+1-\mathrm{fert}(v))!} & \text{if } \mathrm{fert}(v) \le |\mathrm{dec}(v)| + 1, \\ 0 & \text{else.} \end{cases}$$

Here $\mathcal{T}(S)$ *denotes the set of* S*-decorated trees,* $\tau^{[0]}$ *the set of vertices of* τ*,* $\mathrm{dec}(v)$ *the decoration (in* S*) of* v*,* $|\tau|$ *the decoration weight of* τ*, i.e.* $|\tau| = \sum_{v \in \tau^{[0]}} |\mathrm{dec}(v)|$*, and* $\mathrm{fert}(v)$ *the fertility (number of outgoing edges) of the vertex* v*.*

Proof. This is an easy induction using the following argument: Let τ be a given tree in c_n and let its root o be decorated by something in degree m. According to (2.7), $\tau = B_+^{d_m}(\mathbb{I}^{k_0}\tau_1 \dots \tau_{m+1-k_0})$ where the τ_i are trees different from $\mathbb{I}$. The fertility of the root is thus $m + 1 - k_0$. We assume $\tau_1 \dots \tau_{m+1-k_0} = \sigma_1^{k_1} \dots \sigma_p^{k_p}$ where the σ_i are pairwise different trees. In (2.7), there are $C := \frac{(m+1)!}{k_0! \dots k_p!}$ choices to make which yield the tree τ. Since the γ_v are simply multiplied for all vertices v of a tree, it remains to see that for the only new vertex o in τ, we have

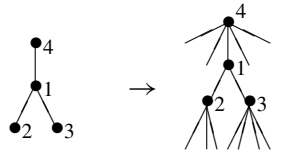
$$\gamma_o / w_m = \frac{(m+1)!}{k_0!} = C \frac{\mathrm{Sym}(\tau)}{\mathrm{Sym}(\tau_1) \dots \mathrm{Sym}(\tau_{m+1-k_0})}.$$

This however follows immediately from the definition of Sym. □

As a matter of fact, the coefficients

$$\prod_v \gamma_v \tag{2.11}$$

can be interpreted as follows: Consider each tree as an "operadic" object with $|\mathrm{dec}(v)| + 1 - \mathrm{fert}(v)$ inputs at each vertex v. For example,



Clearly, the total number of inputs is $n + 1$ for any tree of decoration weight n. Now the coefficient (2.11) is nothing but the number of planar embeddings of this operadic tree (where the trunk, i.e. the original tree is kept fixed). In other words, (2.11) counts the number of ways that the input edges can sway around the original tree. Using this idea, we obtain the following.

Operadic proof of Theorem 2.2. As a variation of (2.6) let us consider the operadic fixpoint equation

$$G(\alpha) = \mathbb{I} + \sum_n \alpha^n \mu_{n+1}(G(\alpha)^{\otimes(n+1)}).$$

Here, μ_j is a map $\in O^{[j]}\colon V^{\otimes j} \to V$ for some space V and $G(\alpha)$ is a formal series in α with coefficients in the $O^{[j]}$. We regard $\mathbb{I}\colon V \to V$ as the identity map. We write $G(\alpha) = \mathbb{I} + \sum_k \alpha^k \nu_k$. It follows easily by induction that $\nu_k \in O^{[k+1]}$. Clearly, $G(\alpha)$ is a sum (with unit weights) over all maps which we obtain by composition of some undecomposable maps μ_n.

The coproduct of decorated rooted trees acts on the ν_k in an obvious manner. A given monomial $\nu_{i_1}^{r_1} \cdots \nu_{i_l}^{r_l}$ (which lives in the PROP $V^{\otimes(r_1 i_1 + \cdots + r_l i_l + r)} \to V^{\otimes r}$, where $r = \sum r_i$) can be composed with any element in $O^{[r-1]}$ in

$$\frac{r!}{r_1! \ldots r_l!} \tag{2.12}$$

ways. Hence, as the ν_i sum over all maps with unit weight, this is the contribution to the term in the coproduct which has ν_k on the right hand side and the given monomial on the left hand side. Going back to the initial Dyson–Schwinger equation (2.6), we see that the same argument (2.12) also determines the coproduct on the c_k there, in agreement with (2.8):

$$P_k^n = \sum_{\substack{i_1 r_1 + \cdots + i_l r_l = n-k \\ 0 \le i_s < i_{s+1},\, \sum r_i = k+1}} \frac{(k+1)!}{r_1! \ldots r_l!} c_{i_1}^{r_1} \ldots c_{i_l}^{r_l}. \tag{2.13}$$

Indeed, the trees in c_k were weighted by a product over vertices (2.11), and the coproduct respects the planar structure. □

The coefficients in (2.11) and (2.13) arise thus in a completely natural way due to the transition from a noncommutative (planar) to a commutative (non-planar) setting.

Final remarks. Before we ultimately turn to the more analytical side of Dyson–Schwinger equations, let us mention that by Theorem 2.2, the Connes–Moscovici Hopf subalgebra presented in the preceding subsection is not generated by a Dyson–Schwinger equation of the form (2.6) if we restrict ourselves to one-cocycles mapping into the linear space of generators, as they typically appear in local quantum field theory.

Note that the Hopf algebras which appear as solutions of (2.6) are studied under the name Faa di Bruno algebras in [18], to which the Connes–Moscovici algebra can be related through an isomorphism.

In studying propagator insertions, one encounters the more general equation

$$X = \alpha B_+\left(\frac{\mathbb{I}}{\mathbb{I} - X}\right).$$

It yields a Hopf subalgebra in a similar way, see [6] for details.

Finally let us emphasize that the ladder Hopf algebra $\mathcal{H}_1$ introduced in the last subsection, can be generated by the *linear* Dyson–Schwinger equation

$$X = \mathbb{I} + \alpha B_+(X).$$

The Hopf algebra $\mathcal{H}_1$ plays a special role at the fixpoint of the renormalization group flow [35], see also the next subsection.

As opposed to the above example, we call Dyson–Schwinger equations of the form (2.6) (where some $w_n \neq 0$ for $n \geq 1$) *nonlinear*. They necessarily generate trees with sidebranchings.

2.3 Applications in physics and number theory

In physics, Dyson–Schwinger equations, usually derived by formal means using functional integrals, describe the loop expansion of Green functions in a recursive way. An alternative to derive these equations is given by the very existence of a Hopf algebra underlying perturbation theory. These Hopf algebras provide Hochschild 1-cocycles, and we can obtain the Dyson–Schwinger equations for them in a straightforward manner.

In the following, we first exhibit Dyson–Schwinger equations in three different contexts: as a source for transcendental numbers, as a manner to define a generating function for the polylogarithm, and as the equations of motion for a renormalizable quantum field theory. The presentation is by no means self-contained, and we refer the reader to the growing literature for more details [29], [33], [30], [31], [32], [24].

A simple toy model. Let us consider the equation we had before (2.2),

$$X_2 = \mathbb{I} + \alpha B_+(X_2^2),$$

and let us exhibit the difference between such an equation and the associated linear system

$$X_1 = \mathbb{I} + \alpha B_+(X_1).$$

We will study toy Feynman rules on these Hopf algebras, regarded as characters on the Hopf algebra. We explore that $\phi B_+(\mathbb{I})$ defines an integral kernel k, such that

$$\phi B_+(\mathbb{I})[z] = \int_0^\infty k(x, z)\, dx,$$

where the kernel is homogeneous: $k(ux, uz) = k(x, z)/u$. We regard the integral as the Fourier transform of the kernel with respect to the multiplicative group $\mathbb{R}_+$. To define our first set of renormalized Feynman rules, we simply set

$$\phi B_+(h)[z] = \int_0^\infty (k(x, z) - k(x, 1))\phi(h)[x]\, dx.$$

Note that we have $\phi(h_1 h_2) = \phi(h_1)\phi(h_2)$ which implies $\phi(\mathbb{I})[z] = 1$.

Let us define the transform $K(\gamma)$ of the kernel $k(x, 1)$ to be

$$K(\gamma) = \int_0^\infty k(x, 1) x^{-\gamma} dx.$$

This determines

$$\int_0^\infty k(x, z) x^{-\gamma} dx = z^{-\gamma} K(\gamma).$$

Let us now look at (2.3). Applying ϕ to both sides delivers an integral equation for the Green function

$$\phi(X_1)[z; \alpha] = 1 + \alpha \int_0^\infty \phi(X_1)[x; \alpha](k(x, z) - k(x, 1))\, dx.$$

Note that our choice of renormalized Feynman rules corresponds to the choice of a boundary condition for the Dyson–Schwinger equation, $\phi(X_1)[1; \alpha] = 1$. Omitting the subtraction of the kernel at $z = 1$ defines the unrenormalized Feynman rule ϕ_u, which reconstructs the renormalized one, $\phi = S_R^{\phi_u} \star \phi_u$, where R is the evaluation map at $z = 1$.

Equation (2.3) can be solved by an Ansatz $\phi(X_1)[z] = z^{-\gamma(\alpha)}$, which leads to

$$z^{-\gamma(\alpha)} = 1 + \alpha(z^{-\gamma(\alpha)} - 1) K(\gamma(\alpha)),$$

i.e. the series $\gamma(\alpha)$ is the solution of the equation $1 = \alpha K(\gamma(\alpha))$. The non-linear case (2.3) cannot be solved by such an *Ansatz*. Maintaining the same boundary condition we get the equation

$$\phi(X_2)[z; \alpha] = 1 + \alpha \int_0^\infty (\phi(X_2)[x; \alpha])^2 (k(x, z) - k(x, 1))\, dx.$$

At this moment, it is instructive to introduce a bit of quantum field theory wisdom. We observe that we can write this integral equation in the form

$$\phi(X_2)[z; \alpha] = 1 + \int_0^\infty \phi(X_2)[x; \alpha] \beth(x; \alpha)(k(x, z) - k(x, 1))\, dx,$$

where the running coupling $\beth(x; \alpha) = \alpha \phi(X_2)[x; \alpha]$ has been introduced. We see that we just modify the linear Dyson–Schwinger equation by this running coupling, which forces us to look for solutions not of the form

$$G(\alpha, z) = e^{-\gamma(\alpha)},$$

but instead of the more general form

$$G(\alpha, z) = e^{-\sum_{j=1}^\infty \gamma_j(\alpha) \ln^j z},$$

where the γ_j themselves are recursively defined through γ_1 thanks to the renormalization group.

Indeed, assume now that the running coupling is constant, $\partial_{\ln z} \beth = 0$. This turns the non-linear Dyson–Schwinger equation into the linear one (2.3). That is a general

phenomenon: the linear Dyson–Schwinger equation appears in the limit of a vanishing β-function, and signifies a possible fixpoint of the renormalization group.

This suggests a natural expansion in terms of the coefficients of the β-function, which will be presented elsewhere.

All this has a combinatorial counterpart:

$$\frac{\partial \ln X_2(\alpha)}{\partial \alpha} = S \star Y(X_2(\alpha_R)),$$

where Y is again the grading operator. Let us work this out in an example. We consider the solution $X_2(\alpha)$ of (2.3). Setting $X_2 = \mathbb{I} + \sum_{k=1}^{\infty} \alpha^k c_k$, we find to $O(\alpha^3)$,

$$\begin{aligned}
&\tilde{\Delta}(c_1) = 0 && S * Y(c_1) = c_1\\
&\tilde{\Delta}(c_2) = 2c_1 \otimes c_1 && S * Y(c_2) = 2c_2 - 2c_1^2\\
&\tilde{\Delta}(c_3) = 3c_1 \otimes c_2 + (2c_2 + c_1^2) \otimes c_1 && S * Y(c_3) = 3c_3 - 8c_1c_2 + 5c_1^3.
\end{aligned}$$

Furthermore,

$$\alpha\partial_\alpha \ln X_2(\alpha) = \alpha c_1 + 2\alpha^2(c_2 - \tfrac{1}{2}c_1^2) + 3\alpha^3(c_3 - c_1c_2 + \tfrac{1}{3}c_1^3).$$

Setting

$$\alpha = \frac{\alpha_R}{X(\alpha)},$$

and recursively replacing α by α_R,

$$\begin{aligned}
\alpha(\alpha_R) &= \frac{\alpha_R}{1 + \alpha(\alpha_R)c_1 + \alpha^2(\alpha_R)c_2 + \cdots}\\
&= \frac{\alpha_R}{1 + \frac{\alpha_R c_1}{1+\alpha_R c_1 + \cdots} + \alpha_R^2 c_2 + \cdots}\\
&= \frac{\alpha_R}{1 + \alpha_R c_1 + \alpha_R^2(c_2 - c_1^2) + \cdots},
\end{aligned}$$

confirms the result to that order. The general proof is a straightforward application of the results in [12].

Furthermore, the reader can check that

$$F_{c_m}(c_k) = (k - m + 1)c_{k-m}$$

for all $k \geq m$, which is at the heart of a recursive determination of the above coefficients $\gamma_j(\alpha)$ in (2.3). Here, F_{c_m} is the befooting operator

$$F_{c_m}(c_k) = \langle Z_{c_m} \otimes \mathrm{id}, \Delta(c_k) \rangle$$

of [6].

As a final remark, we mention that it is not the non-linearity which provides the major challenge in solving a non-linear Dyson–Schwinger equation, but the fact that the one-variable Fourier calculus presented above has to be replaced by a multi-variable calculus which leads to transcendental extensions [32] which is a fascinating topic in its own right.

Indeed, consider once more the linear equation (2.3), now with Feynman rules defined by a two-variable kernel $k(x, y, z) = 1/(x + y + z)^2$ with $k(xz, yz, z) = k(x, y, 1)/z^2$

$$\phi(B_+(h))(z) = \int dxdy \frac{(\phi(h)(x))^{q_2}(\phi(h)(y))^{q_1}}{(x + y + z)^2} - \Big|_{z=1},$$

and q_1, q_2 are two positive rational numbers which add to one. Comparing this system with the degenerate system where one of the q_i vanishes (and the other thus is unity) shows that the perturbative expansion in α provides coefficients which are transcendental extensions of the ones obtained in the degenerate case. This rather general phenomenon leads deeply into the transcendental structure of Green functions, currently under investigation.

Dyson–Schwinger equation for the polylog. Quantum field theory is concerned with the determination of correlators which we can regard as generating functions for a perturbative expansion of amplitudes. These correlators are solutions of our Dyson–Schwinger equations, the latter being typical fixpoint equations: the correlator equals a functional of the correlator. Such self-similarities appear in many branches of mathematics. Here, we want to exhibit one such appearance which we find particularly fascinating: the generating function for the polylog [31].

Following [31] consider the following $N \times N$ matrix once more borrowed from Spencer Bloch's function theory of the polylogarithm [2]:

$$\begin{array}{c} \alpha^0 \\ \alpha^1 \\ \alpha^2 \\ \alpha^3 \\ \cdots \end{array} \underbrace{\begin{pmatrix} +1 & | & 0 & | & 0 & | & 0 & | & \cdots \\ -\mathrm{Li}_1(z) & | & 2\pi i & | & 0 & | & 0 & | & \cdots \\ -\mathrm{Li}_2(z) & | & 2\pi i \ln z & | & [2\pi i]^2 & | & 0 & | & \cdots \\ -\mathrm{Li}_3(z) & | & 2\pi i \frac{\ln^2 z}{2!} & | & [2\pi i]^2 \ln z & | & [2\pi i]^3 & | & \cdots \\ \cdots & | & \cdots & | & \cdots & | & \cdots & | & \cdots \end{pmatrix}}_{\begin{array}{ccccc} u^0 & u^1 & u^2 & u^3 & \ldots \end{array}},$$

given up to $N = 4$. We assign an order in a small parameter α to each row, counting rows $0, 1, \ldots$ from top to bottom, similarly we count columns $0, 1, \ldots$ from left to right by a parameter u, and assign an order u^i to the i-th column. The polylog is defined by

$$\mathrm{Li}_n(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^n}$$

inside the unit circle and analytically continued with a branch cut along the real axis from one to plus infinity.

As the n-th polylog appears as the integral over the $(n-1)$-th polylog, we expect to be able to find a straightforward integral equation for its generating function which

resembles a Dyson–Schwinger equation. Consider

$$F(\alpha,u;z)=1-\frac{1}{1-z}+\frac{2\pi iu\alpha}{1-2\pi iu\alpha}+\alpha\Big(\int_0^z\frac{F(\alpha,0;x)}{x}\,dx+\int_1^z\frac{F(\alpha,u;x)-F(\alpha,0;x)}{x}\,dx\Big),$$

where we call $F(\alpha,u;z)$ a renormalized Green function, α the coupling (a small parameter, $0<\alpha<1$), and consider the perturbative expansion

$$F(\alpha,u;z)=1-\frac{1}{1-z}+\sum_{k=1}^{\infty}\alpha^k f_k(u;z).$$

We distinguished the lowest order term $f_0(z)=z/(z-1)$ (which corresponds to the term without quantum corrections in QFT) at order α^0 which here equals $-\operatorname{Li}_0(z)$. The limit $u\to 1$ can be taken in the above Dyson–Schwinger equation. We note that upon introducing a counterterm $Z(\alpha,u;\ln\rho)$, the above equation is the renormalized solution at $\rho\to 0$ of the equation

$$F_\rho(\alpha,u;z)=Z(\alpha,u;\ln\rho)-\frac{1}{1-z}+\frac{2\pi iu\alpha}{1-2\pi iu\alpha}+\alpha\int_0^z\frac{F_\rho(\alpha,u;x)}{x}\,dx.$$

We immediately confirm that, for $k>0$, the term of order $\alpha^k u^i$ in the renormalized solution of this Dyson–Schwinger equation is the entry (k,i) in the above matrix: the above matrix provides in its non-trivial entries the solution of the Dyson–Schwinger equation so constructed.

We now work with the cocommutative Hopf algebra $\mathcal{H}_1$ determined by the Dyson–Schwinger equation $X=\mathbb{I}+\alpha B_+(X)$, so $X=\sum_{k=0}^{\infty}\alpha^k t_k$ where the t_k are k-fold application of B_+ to $\mathbb{I}$, and let $Li\equiv Li(z)$ and $L\equiv L(z)$ be characters on the Hopf algebra defined by

$$-\phi(t_n)(z,0)\equiv Li(t_n)(z)=\operatorname{Li}_n(z),\ \ L(t_n)(z)=\frac{\ln^n(z)}{n!}.$$

We can regard the character Li as a Feynman rule and the transition $Li\to L$ as a renormalization map which leaves the behavior at infinity unchanged.

We know [2] that the elimination of all ambiguities due to a choice of branch lies in the construction of functions $a_p(z)=(2\pi i)^{-p}\,\widetilde{a}_p(z)$ where

$$\widetilde{a}_p(z):=\operatorname{Li}_p(z)-\cdots+(-1)^j\operatorname{Li}_{p-j}(z)\frac{\ln^j(z)}{j!}+\cdots+(-1)^{p-1}\operatorname{Li}_1(z)\frac{\ln^{p-1}(z)}{(p-1)!}.$$

This is now a very familiar equation:

Proposition 2.4. *For $z\in\mathbb{C}$,*

$$\tilde{a}_p(z)=m(L^{-1}\otimes Li)(\mathrm{id}\otimes P)\Delta(t_p),$$

where $L^{-1} = LS$, with S the antipode in $\mathcal{H}_1$, and P the projection onto the augmentation ideal.

Proof. Elementary combinatorics confirming that

$$LS(t_n/n!)(z) = (-\ln(z))^n/n!. \qquad \square$$

There is a strong analogy here to the Bogoliubov R operation in renormalization theory [29], [31], thanks to the fact that Li and L have matching asymptotic behavior for $|z| \to \infty$. Indeed, if we let R be defined to map the character Li to the character L, $R(Li) = L$, and P the projector onto the augmentation ideal of $\mathcal{H}_1$, then

$$LS = S_R^{Li} = -Rm(S_R^{Li} \otimes Li)(\mathrm{id} \otimes P)\Delta \equiv -R(\overline{Li}),$$

for example

$$S_R^{Li}(t_2) = -R\big(Li(t_2) + S_R^{Li}(t_1)Li(t_1)\big) = -L(t_2) + L(t_1)L(t_1) = \frac{\ln^2(z)}{2!},$$

where $\overline{Li}(t_2) = Li(t_2) - L(t_1)Li(t_1)$. Thus, a_p is the result of the Bogoliubov map

$$\overline{Li} = m(S_R^{Li} \otimes Li)(\mathrm{id} \otimes P)\Delta$$

acting on t_n. We have two completely equivalent mechanisms for the removal of ambiguities at this moment:

$$L^{-1} \star Li \ \text{ vs } \ S_R^{\phi} \star \phi.$$

This points towards an analogy between the structure of the polylog and QFT Green functions which very much suggests to explore QFT from the viewpoint of mixed Hodge structures in the future.

Dyson–Schwinger equations for full QFT. The quantum equations of motion, the Dyson–Schwinger equations of a full fledged quantum field theory, can be obtained in precisely the same manner as discussed above. They typically are of the form

$$\Gamma^{\underline{r}} = \mathbb{I} + \sum_{\substack{\gamma \in H_L^{[1]} \\ \mathrm{res}(\gamma) = \underline{r}}} \frac{\alpha^{|\gamma|}}{\mathrm{Sym}(\gamma)} B_+^{\gamma}(X_{\mathcal{R}}^{\gamma}) = \mathbb{I} + \sum_{\substack{\Gamma \in H_L \\ \mathrm{res}(\Gamma) = \underline{r}}} \frac{\alpha^{|\Gamma|}\Gamma}{\mathrm{Sym}(\Gamma)},$$

where the first sum is over a countable set of Hopf algebra primitives γ, $\mathrm{res}(\gamma) = \underline{r}$,

$$\Delta(\gamma) = \gamma \otimes \mathbb{I} + \mathbb{I} \otimes \gamma,$$

indexing the Hochschild 1-cocycles B_+^{γ} above, while the second sum is over all one-particle irreducible graphs contributing to the desired Green function, all weighted by their symmetry factors. In more traditional terms, the primitive graphs γ correspond to skeletons into which vertex and propagator corrections are to be inserted.

Here, $\Gamma^{\underline{r}}$ is to be regarded as a formal series

$$\Gamma^{\underline{r}} = \mathbb{I} + \sum_{k \geq 1} c^{\underline{r}}_k \alpha^k, \quad c^{\underline{r}}_k \in H.$$

These coefficients of the perturbative expansion deliver Hopf subalgebras in their own right, cf. Theorem 2.2. Indeed, the maps

$$B^{\underline{r},n}_+ = \sum_{\substack{\gamma \in H_L^{[1]} \\ \mathrm{res}(\gamma) = \underline{r},\, |\gamma| = n}} B^\gamma_+,$$

where the sum is over all primitive 1PI n-loop graphs γ with external leg structure $\underline{r}$, are 1-cocycles. They are implicitly defined by the second equality in (2.3), the remarkable feature is the fact that these maps can be shown to be Hochschild closed and hence ensure locality. A detailed account of this fact, which illuminates in particular the structure of gauge theories, is upcoming [24].

In (2.3), $X^\gamma_{\mathcal{R}}$ is of the form

$$X^\gamma_{\mathcal{R}} = \Gamma^{\mathrm{res}(\gamma)} (X_{\mathrm{coupl}})^{|\gamma|},$$

where X_{coupl} is the vertex function divided by the square roots of the inverse propagator functions. Under the Feynman rules X_{coupl} hence maps to the invariant charge.

As an example, consider QED. We have a set of residues (external leg structures)

$$\mathrm{Res} = \Big\{ \quad , \quad , \quad \Big\}.$$

We finish our paper by exhibiting the action of the Hochschild 1-cocycle B_+ on the order α expansion of

$$X \quad = \frac{\left(\Gamma \quad\right)^3}{\left(\Gamma \quad\right)^2 \left(\Gamma \quad\right)} = \Gamma \quad \left(X_{\mathrm{coupl}}\right)^2,$$

with

$$X_{\mathrm{coupl}} = \Gamma \quad \left(\Gamma \quad \sqrt{\Gamma \quad}\right)^{-1}.$$

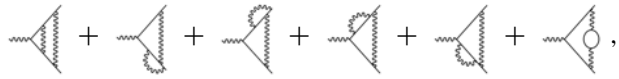
To order α, one finds

$$X \quad = 1 + \alpha \left(3 \quad + 2 \quad + \quad\right).$$

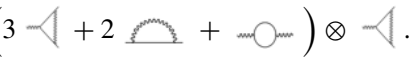
Hence, the non-primitive two-loop vertex graphs of QED are obtained as

$$B_+ \quad \left(3 \quad + 2 \quad + \quad\right).$$

Hochschild closedness demands that this equals



as then



In this manner one determines the Hochschild 1-cocycles for a renormalizable quantum field theory. This works particularly nice for gauge theories, as will be exhibited in [24].

2.4 Final remarks

There is a very powerful structure behind the above decomposition into Hopf algebra primitives – the fact that the sum over all Green functions $G^{\underline{n}}$ is indeed the sum over all 1PI graphs, and this sum, the effective action, can be written nicely as $\prod \frac{1}{1-\gamma}$, a product over "prime" graphs – graphs which are primitive elements of the Hopf algebra and which index the Hochschild 1-cocycles, delivering a complete factorization of the action. A single such Euler factor with its corresponding Dyson–Schwinger equation and Feynman rules was evaluated in [6], a calculation which was entirely in accordance with our study: an understanding of the weight of contributions $\sim \ln(z)$ from a knowledge of the weight of such contributions of smaller degree in α, dubbed propagator-coupling duality in [6]. Altogether, this allows to summarize the structure in QFT as a vast generalization of results summarized here. It turns out that even the quantum structure of gauge theories can be understood along these lines [28]. A full discussion is upcoming [24].

References

[1] C. Bergbauer and D. Kreimer, The Hopf algebra of rooted trees in Epstein–Glaser renormalization, *Ann. Henri Poincaré* 6 (2004), 343–367; `arXiv:hep-th/0403207`.

[2] S. Bloch, Function theory of polylogarithms, in *Structural properties of polylogarithms*, Math. Surveys Monogr. 37, Amer. Math. Soc., Providence, RI, 1991, 275–285.

[3] N. N. Bogolyubov and D. V. Shirkov, *Introduction to the theory of quantized fields*, 3rd ed., Wiley, New York 1980.

[4] D. J. Broadhurst and D. Kreimer, Renormalization automated by Hopf algebra, *J. Symb. Comput.* 27 (1999), 581–600; `arXiv:hep-th/9810087`.

[5] D. J. Broadhurst and D. Kreimer, Combinatoric explosion of renormalization tamed by Hopf algebra: 30-loop Padé-Borel resummation, *Phys. Lett. B* 475 (2000), 63–70; `arXiv:hep-th/9912093`.

[6] D. J. Broadhurst and D. Kreimer, Exact solutions of Dyson-Schwinger equations for iterated one-loop integrals and propagator-coupling duality, *Nucl. Phys. B* 600 (2001), 403–422; arXiv:hep-th/0012146.

[7] R. Brunetti and K. Fredenhagen, Microlocal analysis and interacting quantum field theories: renormalization on physical backgrounds, *Commun. Math. Phys.* 208 (2000), 623–661; arXiv:math-ph/9903028.

[8] F. Chapoton and M. Livernet, Pre-Lie algebras and the rooted trees operad, *Internat. Math. Res. Notices* 2001, 395–408; arXiv:math.QA/0002069.

[9] K. T. Chen, Iterated path integrals, *Bull. Amer. Math. Soc.* 83 (1977), 831–879.

[10] A. Connes and D. Kreimer, Hopf algebras, renormalization and noncommutative geometry, *Comm. Math. Phys.* 199 (1998), 203–242; arXiv:hep-th/9808042.

[11] A. Connes and D. Kreimer, Renormalization in quantum field theory and the Riemann-Hilbert problem I: The Hopf algebra structure of graphs and the main theorem, *Comm. Math. Phys.* 210 (2000), 249–273; arXiv:hep-th/9912092.

[12] A. Connes and D. Kreimer, Renormalization in quantum field theory and the Riemann-Hilbert problem. II: The beta-function, diffeomorphisms and the renormalization group, *Comm. Math. Phys.* 216 (2001), 215–241; arXiv:hep-th/0003188.

[13] A. Connes and H. Moscovici, Hopf algebras, cyclic cohomology and the transverse index theorem, *Comm. Math. Phys.* 198 (1998), 199–246.

[14] A. Connes and H. Moscovici, Background independent geometry and Hopf cyclic cohomology, 2005; arXiv:math.qa/0505475.

[15] K. Ebrahimi-Fard, L. Guo, and D. Kreimer, Integrable renormalization I: The ladder case, *J. Math. Phys.* 45 (2004), 3758–3769; arXiv:hep-th/0402095.

[16] K. Ebrahimi-Fard, L. Guo, and D. Kreimer, Integrable renormalization II: The general case, *Ann. Henri Poincaré* 6 (2005), 369–395; arXiv:hep-th/0403118.

[17] H. Epstein and V. Glaser, The role of locality in perturbation theory, *Ann. Inst. H. Poincaré A* 19 (1973), 211–295.

[18] H. Figueroa and J. M. Gracia-Bondia, Combinatorial Hopf algebras in quantum field theory. I, *Rev. Math. Phys.* 17 (2005), 881–976; arXiv:hep-th/0408145.

[19] L. Foissy, Les algèbres de Hopf des arbres enracinés décorés. I, II, *Bull. Sci. Math.* 126 (2002), 193–239; 249–288.

[20] W. Fulton and R MacPherson, A compactification of configuration spaces, *Ann. of Math.* (2) 139 (1994), 183–225.

[21] H. Gangl, A. Goncharov, and A Levin, Multiple logarithms, algebraic cycles and trees, 2005; arXiv:math.NT/0504552.

[22] R. Grossman and R. G. Larson, Hopf-algebraic structure of families of trees, *J. Algebra* 126 (1989), 184–210.

[23] K. Hepp, Proof of the Bogolyubov-Parasiuk Theorem on Renormalization, *Comm. Math. Phys.* 2 (1966), 301–326.

[24] D. Kreimer, Anatomy of a gauge theory, Ann. Phys. (online first) 2006; arXiv:hep-th/0509135.

[25] D. Kreimer, On the Hopf algebra structure of perturbative quantum field theories, *Adv. Theor. Math. Phys.* 2 (1998), 303–334; arXiv:q-alg/9707029.

[26] D. Kreimer, On overlapping divergences, *Comm. Math. Phys.* 204 (1999), 669–689; arXiv:hep-th/9810022.

[27] D. Kreimer, Chen's iterated integral represents the operator product expansion, *Adv. Theor. Math. Phys.* 3 (1999), 627–670; arXiv:hep-th/9901099.

[28] D. Kreimer, Factorization in quantum field theory: an exercise in Hopf algebras and local singularities, in *Frontiers in Number Theory, Physics, and Geometry II*, Les Houches School of Physics, March 2003, Springer-Verlag, 2006, to appear; arXiv:hep-th/0306020.

[29] D. Kreimer, New mathematical structures in renormalizable quantum field theories, *Ann. Phys.* 303 (2003), 179–202; arXiv:hep-th/0211136.

[30] D. Kreimer, Unique factorization in perturbative QFT, *Nucl. Phys. B Proc. Suppl.* 116 (2003), 392–396; arXiv:hep-ph/0211188.

[31] D. Kreimer, The residues of quantum field theory – numbers we should know (contributed to Workshop on Noncommutative Geometry and Number Theory, Bonn, Germany, 18–22 August 2003), 2004; arXiv:hep-th/0404090.

[32] D. Kreimer, What is the trouble with Dyson–Schwinger equations?, *Nucl. Phys. B Proc. Suppl.* 135 (2004), 238–242; arXiv:hep-th/0407016.

[33] D. Kreimer, Structures in Feynman graphs: Hopf algebras and symmetries, in *Graphs and patterns in mathematics and theoretical physics*, June 14-21, 2001, Proc. Symp. Pure Math. 73, Amer. Math. Soc., Providence, RI, 2005, 43–78; arXiv:hep-th/0202110.

[34] M. Markl, S. Shnider, and J. Stasheff, *Operads in algebra, topology and physics*, Math. Surveys Monogr. 96, Amer. Math. Soc., Providence, RI, 2002.

[35] I. Mencattini and D. Kreimer, The structure of the ladder insertion-elimination Lie algebra, *Comm. Math. Phys.* 259 (2005), 413–432; arXiv:math-ph/0408053.

[36] S. Popineau and R. Stora, A pedagogical remark on the Main Theorem of Perturbative Renormalization Theory, unpublished preprint.

[37] V. Turaev, Loops on surfaces, Feynman diagrams, and trees, *J. Geom. Phys.* 53 (2005), 461–482; arXiv:hep-th/0403266.

[38] W. Zimmermann, Convergence of Bogolyubov's method of renormalization in momentum space, *Comm. Math. Phys.* 15 (1969), 208–234.

Fonction ζ et matrices aléatoires

*Emmanuel Royer**

Institut de mathématiques et modélisation de Montpellier, UMR5149
Université Montpellier II
Case courrier 051
F-34095 Montpellier cedex 5
France
email : royer@math.univ-montp2.fr

Résumé. Depuis Euler puis Riemann, de nombreux liens ont été établis entre le comportement des nombres premiers et les propriétés analytiques de la fonction ζ de Riemann. Les points où ζ s'annulent ont une importance particulière justifiant leur étude fine. Depuis les années 70, un angle d'attaque de ces points d'annulation est apparu *via* les propriétés statistiques du spectre des matrices unitaires, qui se présentent comme un modèle de ζ. L'objet de ce texte de présentation est d'expliquer cet angle d'attaque et ses extensions à l'étude des fonctions L.

Abstract. Since Euler and Riemann, various links have been established between the behaviour of prime numbers and the analytical properties of the Riemann ζ function. The zeroes of ζ have a great importance that justifies their fine study. Since seventies, a rich tool has appeared for the study of zeroes: the statistical spectral properties of the unitary matrices that are a model for ζ. The aim of this survey is to explain the link between unitary matrices and ζ, and its extension to the L-functions.

Table des matières

*Partiellement financé par l'ACI jeunes chercheurs « arithmétique des fonctions L » et le réseau européen *Arithmetic Algebraic Geometry network*.

1 Le théorème de Montgomery

L'objet de cette partie est de rappeler les propriétés importantes de la fonction ζ de Riemann. Pour les détails, on renvoie à [75].

La fonction ζ de Riemann est définie, pour $\operatorname{Re} s > 1$ par

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$$

et cet objet analytique est relié à l'arithmétique de $\mathbb{Q}$ par le développement en produit eulérien

$$\zeta(s) = \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p^s}\right)^{-1} \tag{1.1}$$

en désignant par $\mathcal{P}$ l'ensemble des nombres premiers. On sait prolonger cette fonction en une fonction méromorphe sur $\mathbb{C}$ dont le seul pôle est simple et situé en $s = 1$. Le prolongement satisfait à une équation fonctionnelle

$$\Lambda(s) = \Lambda(1 - s) \tag{1.2}$$

où

$$\Lambda(s) = \zeta_\infty(s)\zeta(s) \quad \text{avec} \quad \zeta_\infty(s) = \pi^{-s/2}\Gamma_{\mathbb{R}}(s)$$

avec

$$\Gamma_{\mathbb{R}}(s) = \Gamma\left(\frac{s}{2}\right).$$

La fonction ζ s'annule en les entiers pairs strictement négatifs, ces points d'annulation sont appelés *zéros triviaux de* ζ. Le développement eulérien (1.1) et l'équation fonctionnelle (1.2) impliquent que tous les autres points d'annulation de ζ sont dans la bande $0 \leq \operatorname{Re} s \leq 1$ appelée *bande critique*. Grâce à l'équation fonctionnelle, les points d'annulation de ζ sont symétriques par rapport à la droite $\operatorname{Re} s = \frac{1}{2}$ appelée *droite critique* et Riemann a conjecturé l'hypothèse suivante.

Hypothèse de Riemann. *Les points d'annulation non triviaux de* ζ *sont sur la droite critique.*

Cette hypothèse est vérifiée[1] par Gourdon & Demichel *via* un calcul informatique pour les 10^{13} premiers points d'annulation classés par parties imaginaires positives croissantes [34]. D'autre part, Conrey a montré [13] que 40 % au moins des points d'annulation de ζ sont situés sur l'axe critique :

$$\lim_{T \to +\infty} \frac{\#\{\rho \in \zeta^{-1}(\{0\}) \colon 0 < \operatorname{Im} \rho < T,\ \operatorname{Re} \rho = 1/2\}}{\#\{\rho \in \zeta^{-1}(\{0\}) \colon 0 < \operatorname{Im} \rho < T\}} \geq 40\,\%.$$

[1] « En plus des erreurs possibles concernant la validité des résultats et algorithmes utilisés, [le calcul] est sujet a des erreurs qui ne sont pas aisément détectables (erreur humaine de codage, erreur de compilation, erreur de système, erreur de processeur *etc*.). Il est donc difficile de considérer un tel résultat comme "démontré" au sens fort de démonstration purement mathématiques. » (X. Gourdon *in* [34].)

L'hypothèse de Riemann équivaut au résultat suivant de distribution des nombres premiers : il existe $c > 0$ tel que

$$|\pi(x) - \mathrm{Li}(x)| \leq cx^{1/2}\log(x) \tag{1.3}$$

pour tout $x \geq 2$ avec

$$\pi(x) = \#\{p \in \mathcal{P} \colon p \leq x\} \quad \text{et} \quad \mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t}.$$

En conséquence, l'hypothèse de Riemann implique l'existence de $c > 0$ tel que tout intervalle

$$[x, x + cx^{\theta}\log^2(x)] \tag{1.4}$$

avec $x \geq 2$ et $\theta = 1/2$ contient au moins un nombre premier. Un moyen d'obtenir le résultat avec $1/2 < \theta < 1$ est de remplacer $x^{1/2}$ dans le terme de droite de (1.3) par x^{θ}. L'évaluation obtenue de π est alors équivalente à la « quasi-hypothèse de Riemann » affirmant que les points d'annulation non triviaux de ζ sont dans la bande $1 - \theta \leq \mathrm{Re}\, s \leq \theta$. Aucun résultat de ce type n'est connu mais il est remarquable que l'on sache démontrer que $\theta = 0,525$ est admissible dans (1.4) [2] (voir aussi la proposition 1.17 ci-dessous et [70]). Les premiers résultats sont dûs à Hoheisel en 1930 [39] (voir [46, §10.5]). Pour de plus amples informations concernant l'hypothèse de Riemann, on renvoie aux exposés de Bombieri [7] et Sarnak [73] et aux textes d'Ivic [44], [45]. Dans cet exposé, essentiellement pour simplifier l'exposition, on supposera que cette hypothèse est démontrée. La plupart des énoncés restent vrais si on ne fait pas cette hypothèse en prenant des fonctions de localisation (voir la définition 1.1) d'argument complexe.

Puisque $\overline{\zeta(s)} = \zeta(\bar{s})$, les points d'annulation de ζ sont symétriques par rapport à l'axe $\mathrm{Im}\, s = 0$. Le point $s = 1/2$ n'annule pas ζ puisque, par exemple à l'aide de la formule de Taylor-Mac Laurin, on a

$$\zeta(1/2) = -1,460354508809586812889499152515298012467229331012 6\ldots.$$

On note

$$Z(\zeta) = \{\beta_n + i\gamma_n,\ n \in \mathbb{Z}^*\}$$

l'ensemble des points d'annulation de ζ qui ne sont pas des zéros triviaux avec les conventions que $(\gamma_n)_{n\in\mathbb{Z}^*}$ croît et $\gamma_{-n} = -\gamma_n$ pour $n \in \mathbb{Z}^*$. Un éventuel point d'annulation multiple (on conjecture qu'il n'en existe pas) est répété autant de fois que sa multiplicité.

Notons

$$N(\zeta, T) = \#\{\gamma \in Z(\zeta) \colon 0 \leq \mathrm{Im}\, \gamma \leq T\}.$$

Alors

$$N(\zeta, T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi e}\right) + O(\log T)$$

pour tout $T \geq 3$ (la première ordonnée positive d'un point d'annulation de ζ est

$$\gamma_1 = 14,134725141734693790457251983562470270784257115 6\ldots).$$

Pour $T \geq 15$, on a donc

$$T = \frac{\gamma_T}{2\pi} \log\left(\frac{\gamma_T}{2\pi e}\right) + O(\log \gamma_T)$$

et, en posant

$$\tilde{\gamma}_n = \frac{\gamma_n}{2\pi} \log\left(\frac{|\gamma_n|}{2\pi e}\right)$$

on obtient

$$\frac{1}{N-1} \sum_{n=1}^{N-1} (\tilde{\gamma}_{n+1} - \tilde{\gamma}_n) = 1 + O\left(\frac{\log N}{N}\right). \tag{1.5}$$

Les points $\beta_n + i\tilde{\gamma}_n$ sont appelés *zéros normalisés* de ζ et (1.5) exprime qu'asymptotiquement, l'écart moyen entre zéros normalisés est 1. On note $\tilde{Z}(\zeta)$ l'ensemble des zéros normalisés de ζ. On va étudier la corrélation des zéros normalisés. Cette étude va dévoiler un lien prometteur entre les très mystérieux points d'annulation de ζ et les beaucoup moins mystérieuses valeurs propres de matrices unitaires. Pour la suite, on définit la notion de fonction de localisation.

Définition 1.1. Une *fonction de localisation* est une fonction de Schwartz sur $\mathbb{R}$ dont la transformée de Fourier est à support compact.

Pour construire une telle fonction, il suffit de prendre la transformée de Fourier inverse d'une fonction C^∞ à support compact.

Remarque 1.2. On choisit la normalisation suivante pour la transformée de Fourier :

$$\mathcal{F}[t \mapsto f(t)](\xi) = \hat{f}(\xi) = \int_{-\infty}^{+\infty} f(t) e(-\xi t)\, \mathrm{d}t$$

avec

$$e(x) = \exp(2i\pi x).$$

La transformée de Fourier inverse est alors

$$\mathcal{F}^{-1}[\xi \mapsto \hat{f}(\xi)](t) = \int_{-\infty}^{+\infty} \hat{f}(\xi) e(\xi t)\, \mathrm{d}\xi.$$

En 1973, Montgomery [64] a démontré un théorème qu'on énonce maintenant de la façon suivante [71].

Théorème 1.3. *Soit h la transformée de Fourier inverse d'une fonction C^∞ à support compact et ϕ une fonction de localisation telle que* $\operatorname{Supp}\hat{\phi} \subset]-1, 1[$. *Alors*

$$\lim_{T\to\infty} \frac{1}{N(\zeta, T)\|h\|_2^2} \sum_{\substack{(j,k)\in\mathbb{Z}^{*2}\\ j\neq k}} h\left(\frac{\gamma_j}{T}\right) h\left(\frac{\gamma_k}{T}\right) \phi\left[\frac{\log T}{2\pi}(\gamma_j - \gamma_k)\right] = \mathrm{COR}(2, \phi, \mathrm{univ})$$

avec

$$\mathrm{COR}(2, \phi, \mathrm{univ}) = \int_{-\infty}^{+\infty} \phi(t)\left[1 - \left(\frac{\sin(\pi t)}{\pi t}\right)^2\right] \mathrm{d}t.$$

Remarque 1.4. Le théorème 1.3 étudie la façon dont se répartissent les écarts entre zéros normalisés dans les intervalles de longueur environ 1. Voir la partie 2.

Remarque 1.5. Dans [71], le théorème 1.3 est énoncé avec $f\left(\frac{\log T}{2\pi}\gamma_j, \frac{\log T}{2\pi}\gamma_k\right)$ au lieu de $\phi\left[\frac{\log T}{2\pi}(\gamma_j - \gamma_k)\right]$, où f est symétrique et vérifie $f(x_1 + t, x_2 + t) = f(x_1, x_2)$ pour tout t réel. Pour f, on a pris $f(x_1, x_2) = \phi(x_2 - x_1)$ et il faudrait se restreindre à ϕ paire. Cependant, lorsque ϕ est impaire, le théorème 1.3 reste valable puisque les deux membres de l'égalité sont nuls. En écrivant une fonction ϕ quelconque comme la somme d'une fonction paire et d'une fonction impaire, on déduit donc qu'on peut enlever l'hypothèse de parité de ϕ.

On notera dCOR(2, univ) la densité de corrélation représentée figure 1 et définie par :

$$\mathrm{dCOR}(2, \mathrm{univ})(t) = 1 - \left(\frac{\sin(\pi t)}{\pi t}\right)^2.$$

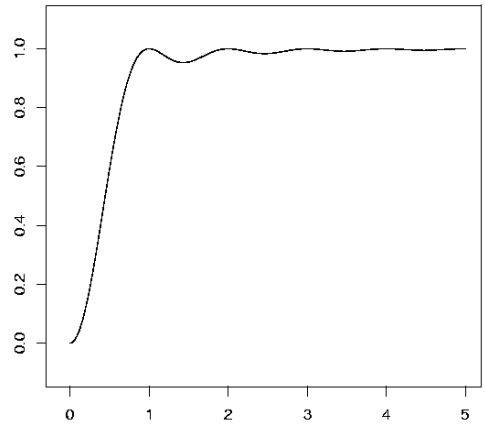


Figure 1. $x \mapsto \mathrm{dCOR}(2, \mathrm{univ})(x)$.

Le théorème 1.3 a été l'objet de nombreux travaux dans les années qui ont suivi sa publication. On peut, par exemple, se reporter à la bibliographie de [25].

Remarque 1.6. En prenant les mêmes hypothèses pour ϕ que dans l'énoncé du théorème 1.3, on peut montrer [71, Theorem 1.2] que

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{1\leq j\neq k\leq N}\phi(\tilde{\gamma}_j-\tilde{\gamma}_k)=\mathrm{COR}(2,\phi,\mathrm{univ}) \tag{1.6}$$

Remarque 1.7. Supposons que, dans l'équation (1.6), on puisse choisir pour ϕ la fonction caractéristique de $[-\delta,\delta]$. On obtient alors

$$\lim_{N\to+\infty}\frac{1}{N}\#\left\{\{\tilde{\gamma}_j,\tilde{\gamma}_k\}\subset\{\tilde{\gamma}_1,\dots,\tilde{\gamma}_N\}\colon|\tilde{\gamma}_j-\tilde{\gamma}_k|\leq\delta\right\}=\frac{1}{9}\pi^2\delta^3+O(\delta^5).$$

Autrement dit, le nombre de petits espaces entre zéros de ζ est très faible : on dit qu'il y a *répulsion*. En calculant environ $2\cdot10^9$ zéros autour de 10^{24}, Gourdon & Demichel [34] ont trouvé que le plus petit écart entre zéros est 0, 0002799. Pour les 10^{13} premiers zéros, le plus petit écart est 0, 0005330. Ce constat se traduit sur le tracé de la courbe de la fonction de distribution cumulative des écarts (voir la figure 2) : au voisinage de 0, cette courbe est aplatie.

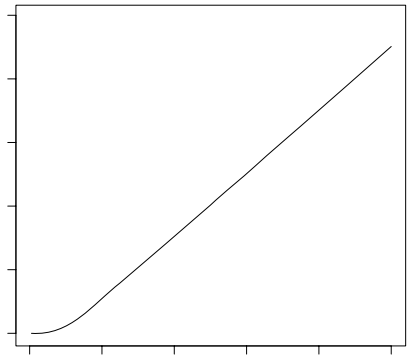


Figure 2. Fonction de distribution cumulative expérimentale des écarts entre zéros normalisés des zéros numérotés de 10^{22} à $10^{22}+10^4$ (d'après les données de Odlyzko [66]).

On explique le tracé de la figure 2. On a reporté les points de coordonnées $(a_i,\,p_i)$ pour $i\in\{1,\dots,200\}$. Les abscisses sont définies par

$$a_i=\frac{5}{200}i$$

de sorte qu'elles forment une suite de 200 points équidistribués de l'intervalle [0, 5]. Les ordonnées sont définies par

$$p_i = \frac{\#\{(j,k) \in \mathbb{N}^2 \colon 10^{22} < j < k < 10^{22} + 10^4,\ \hat{\gamma}_k - \hat{\gamma}_j < a_i\}}{10^4}.$$

Le graphe obtenu coïncide parfaitement avec celui des points de coordonnées (a_i, q_i) où $q_i = \mathrm{COR}(2, \phi_i, \mathrm{univ})$, la fonction ϕ_i étant la fonction caractéristique de $[0, a_i]$. On comparera avec la conjecture de Montgomery (page 177). Au passage, on remarque que l'observation de la distribution des écarts de 10^4 zéros consécutifs ne permet pas de déterminer le paquet de 10^4 zéros consécutifs étudié (*i.e.* de trouver n tel que ce paquet est l'ensemble des zéros numérotés de n à $n + 10^4$).

Remarque 1.8. La restriction sur le support de $\hat{\phi}$ peut sembler artificielle. Néanmoins, elle apparaît de façon essentielle dans la démonstration du théorème. Elle a aussi une signification analytique forte puisque les points d'abscisses ± 1 sont des points de discontinuité de $\hat{\phi}$. La formule de Parseval donne

$$\int_{-\infty}^{+\infty} \mathrm{dCOR}(2, \mathrm{univ})(t)\phi(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} \widehat{\mathrm{dCOR}(2, \mathrm{univ})}(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi$$

avec la transformée de Fourier de la densité représentée figure 3 et donnée par

$$\widehat{\mathrm{dCOR}(2, \mathrm{univ})}(\xi) = \begin{cases} \delta(\xi) + |\xi| - 1 & \text{si } |\xi| < 1 \\ 1 & \text{sinon.} \end{cases}$$

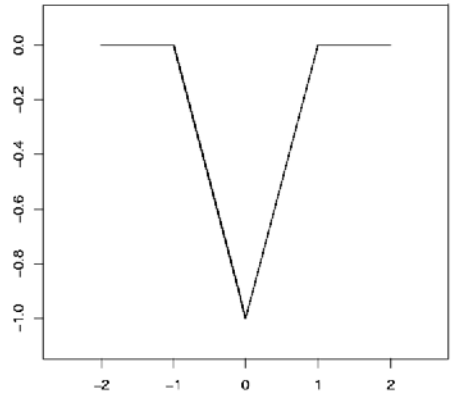


Figure 3. $x \mapsto \widehat{\mathrm{dCOR}(2, \mathrm{univ})}(x)$.

On a noté δ est la distribution de Dirac,

$$\int_{\mathbb{R}} \delta f = f(0).$$

Remarque 1.9. Un ingrédient majeur de la démonstration est la formule explicite de Riemann (voir [71, Proposition 2.1] ou [46, Theorem 5.12]). Cette formule utilise la fonction Λ de Von Mangoldt définie par

$$\Lambda(n) = \begin{cases} \log p & \text{si } n = p^\nu \text{ avec} \in \mathcal{P} \text{ et } \nu \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

Proposition 1.10. *Soit H une fonction de localisation. Alors,*

$$\sum_{j\in\mathbb{Z}^*} H(\gamma_j) = H\Big(-\frac{i}{2}\Big) + H\Big(\frac{i}{2}\Big) + \frac{1}{2\pi}\int_{-\infty}^{+\infty} H(t)\Omega_{\mathbb{R}}(t)\,\mathrm{d}t$$
$$- \frac{1}{2\pi}\sum_{n=1}^{+\infty}\frac{\Lambda(n)}{\sqrt{n}}\Big[\hat{H}\Big(\frac{\log n}{2\pi}\Big) + \hat{H}\Big(-\frac{\log n}{2\pi}\Big)\Big].$$

On a posé

$$\Omega_{\mathbb{R}}(t) = 2\,\mathrm{Re}\,\frac{\Gamma'_{\mathbb{R}}}{\Gamma_{\mathbb{R}}}\Big(\frac{1}{2}+it\Big)$$

avec

$$\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma\Big(\frac{s}{2}\Big).$$

Remarque 1.11. La formule de Riemann implique

$$N(\zeta, T)\|h\|_2^2 \sim \sum_{j\in\mathbb{Z}^*} h\Big(\frac{\gamma_j}{T}\Big)^2 \quad (T\to+\infty). \tag{1.7}$$

Pour démontrer (1.7), on utilise la formule de Riemann avec

$$H(t) = h^2\Big(\frac{t}{T}\Big), \quad \hat{H}(\xi) = T\widehat{h^2}(\xi T),$$

où $h\colon \mathbb{R}\to\mathbb{R}$ est telle que h^2 est une fonction de localisation. Grâce à une formule due à Mertens [46, §2.2]

$$\sum_{n\leq t}\frac{\Lambda(n)}{n} \sim \log t \quad (t\to+\infty)$$

(c'est aussi une conséquence directe du théorème des nombres premiers [74, Corollaire 8.1 du chapitre I.3]) et à une intégration par parties, on a

$$\frac{1}{2\pi}\sum_{n=1}^{+\infty}\frac{\Lambda(n)}{\sqrt{n}}\Big[\hat{H}\Big(\frac{\log n}{2\pi}\Big) + \hat{H}\Big(-\frac{\log n}{2\pi}\Big)\Big] = O(1).$$

D'autre part [35, 8.361.8],

$$\frac{\Gamma'_{\mathbb{R}}}{\Gamma_{\mathbb{R}}}(z) = \frac{1}{2}\log\frac{z}{2\pi} + \frac{1}{2}\int_0^{+\infty} \exp\left(-\frac{tz}{2}\right)\left(\frac{1}{t} - \frac{1}{1-e^{-t}}\right)\mathrm{d}t$$

de sorte que, si $|t|$ est borné, alors

$$\Omega_{\mathbb{R}}(tT) = \log T + O(1)$$

la constante de majoration ne dépendant que de la borne de $|t|$. On a donc

$$\begin{aligned}\frac{1}{2\pi}\int_{\mathbb{R}} H(t)\Omega_{\mathbb{R}}(t)\,\mathrm{d}t &= \frac{T}{2\pi}\int_{\mathbb{R}} h^2(t)\Omega_{\mathbb{R}}(tT)\,\mathrm{d}t \\ &\sim \frac{T}{2\pi}\log T\int_{\mathbb{R}} h^2(t)\,\mathrm{d}t \quad (T\to+\infty).\end{aligned}$$

Remarque 1.12. L'équation (1.7) implique

$$\begin{aligned}&\frac{1}{N(\zeta,T)\|h\|_2^2}\sum_{\substack{(j,k)\in\mathbb{Z}^{*2}\\ j\neq k}} h\left(\frac{\gamma_j}{T}\right)h\left(\frac{\gamma_k}{T}\right)\phi\left[\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right] \\ &\sim \frac{1}{N(\zeta,T)\|h\|_2^2}\sum_{(j,k)\in\mathbb{Z}^{*2}} h\left(\frac{\gamma_j}{T}\right)h\left(\frac{\gamma_k}{T}\right)\phi\left[\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right] - \phi(0) \quad (T\to+\infty).\end{aligned}$$

On va donc montrer

$$\begin{aligned}&\frac{1}{N(\zeta,T)\|h\|_2^2}\sum_{(j,k)\in\mathbb{Z}^{*2}} h\left(\frac{\gamma_j}{T}\right)h\left(\frac{\gamma_k}{T}\right)\phi\left[\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right] \\ &\qquad\sim \int_{\mathbb{R}} K(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi \quad (T\to\infty)\end{aligned}$$

avec

$$K(\xi) = 1 + \widehat{\mathrm{dCOR}(2,\mathrm{univ})}(\xi) = \begin{cases}\delta(\xi) + |\xi| & \text{if } |\xi|<1\\ 1 & \text{sinon.}\end{cases}$$

On donne quelques ingrédients de la démonstration. L'objectif n'est pas de répéter une démonstration très bien écrite dans [71] (voir aussi [46, Chapter 25]) mais d'expliquer la provenance des termes principaux. On sépare les zéros en calculant

$$\begin{aligned}&\sum_{(j,k)\in\mathbb{Z}^{*2}} h\left(\frac{\gamma_j}{T}\right)h\left(\frac{\gamma_k}{T}\right)\phi\left[\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right] \\ &\qquad= \sum_{(j,k)\in\mathbb{Z}^{*2}} h\left(\frac{\gamma_j}{T}\right)h\left(\frac{\gamma_k}{T}\right)\int_{\mathbb{R}}\delta\left(t-\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right)\phi(t)\,\mathrm{d}t.\end{aligned}$$

Ainsi, de

$$\mathcal{F}\left[t\mapsto\delta\left(t-\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right)\right](\xi) = \exp(-i\xi(\gamma_j-\gamma_k)\log T)$$

on déduit

$$\sum_{(j,k)\in\mathbb{Z}^{*2}} h\left(\frac{\gamma_j}{T}\right) h\left(\frac{\gamma_k}{T}\right) \phi\left[\frac{\log T}{2\pi}(\gamma_j-\gamma_k)\right] = \int_{\mathbb{R}} \Big| \sum_{j\in\mathbb{Z}^*} h\left(\frac{\gamma_j}{T}\right) e^{i\xi\gamma_j \log T}\Big|^2 \hat{\phi}(\xi)\,\mathrm{d}\xi.$$

La formule explicite, appliquée à

$$H(t) = h\Big(\frac{t}{T}\Big)e^{i\xi t\log T}, \quad \widehat{H}(u) = T\hat{h}\Big[\Big(\frac{\xi}{2\pi}\log T - u\Big)T\Big]$$

conduit alors à

$$\begin{aligned} &h\Big(\frac{\gamma_j}{T}\Big)e^{i\xi\gamma_j\log T} \\ &\quad = h\Big(\frac{i}{2T}\Big)T^{-\xi/2} + h\Big(\frac{-i}{2T}\Big)T^{\xi/2} + \frac{1}{2\pi}\int_{\mathbb{R}} h\Big(\frac{t}{T}\Big)\Omega_{\mathbb{R}}(t)e^{i\xi t\log T}\,\mathrm{d}t \\ &\quad\quad - \frac{T}{2\pi}\sum_{n=1}^{+\infty}\frac{\Lambda(n)}{\sqrt{n}}\Big\{\hat{h}\Big[(\xi\log T-\log n)\frac{T}{2\pi}\Big] + \hat{h}\Big[(\xi\log T+\log n)\frac{T}{2\pi}\Big]\Big\}. \end{aligned} \tag{1.8}$$

On prend le carré de la norme de cette égalité. On peut montrer que la contribution des doubles produits est négligeable. D'autre part, les termes $h(2/iT)T^{\pm\xi/2}$ ne contribuent pas en raison de la restriction sur le support de $\hat{\phi}$. On étudie alors la contribution de l'intégrale, elle sera la distribution de Dirac dans K. Cette contribution est

$$\begin{aligned} &\frac{1}{(2\pi)^2}\int_{\xi\in\mathbb{R}}\Big|\int_{t\in\mathbb{R}} h\Big(\frac{t}{T}\Big)\Omega_{\mathbb{R}}(t)e^{-i\xi t\log T}\,\mathrm{d}t\Big|^2\hat{\phi}(\xi)\,\mathrm{d}\xi \\ &\quad = \frac{1}{(2\pi)^2 T\log T}\int_{\alpha\in\mathbb{R}}\Big|\int_{t\in\mathbb{R}} h\Big(\frac{t}{T}\Big)\Omega_{\mathbb{R}}(t)e^{-i\alpha t/T}\,\mathrm{d}t\Big|^2\hat{\phi}\left(\frac{\alpha}{T\log T}\right)\mathrm{d}\alpha \\ &\quad = \frac{T}{(2\pi)^2\log T}\int_{\alpha\in\mathbb{R}}\Big|\int_{u\in\mathbb{R}} h(u)\Omega_{\mathbb{R}}(uT)e^{-i\alpha u}\,\mathrm{d}u\Big|^2\hat{\phi}\Big(\frac{\alpha}{T\log T}\Big)\,\mathrm{d}\alpha \\ &\quad \sim \frac{T}{(2\pi)^2\log T}\hat{\phi}(0)\int_{\alpha\in\mathbb{R}}\Big|\int_{u\in\mathbb{R}} h(u)\Omega_{\mathbb{R}}(uT)e^{-i\alpha u}\,\mathrm{d}u\Big|^2\,\mathrm{d}\alpha \quad (T\to\infty). \end{aligned}$$

Notons $g_T(\alpha)$ l'intégrale en u. Alors,

$$\int_{\alpha\in\mathbb{R}} g_T(\alpha)g_T(-\alpha)\,\mathrm{d}\alpha = \int_{\xi\in\mathbb{R}}\widehat{g_T}(\xi)\widehat{g_T}(-\xi)\,\mathrm{d}\xi$$

avec

$$\begin{aligned} \widehat{g_T}(\xi) &= \int_{u\in\mathbb{R}} h(u)\Omega_{\mathbb{R}}(uT)\mathcal{F}\Big[\alpha\mapsto e\Big(-\Big(\xi+\frac{u}{2\pi}\Big)\alpha\Big)\Big](0)\,\mathrm{d}u \\ &= 2\pi h(2\pi\xi)\Omega_{\mathbb{R}}(2\pi\xi T). \end{aligned}$$

Ainsi,

$$\frac{1}{(2\pi)^2}\int_{\xi\in\mathbb{R}}\Big|\int_{t\in\mathbb{R}}h\Big(\frac{t}{T}\Big)\Omega_{\mathbb{R}}(t)e^{-i\xi t\log T}\,\mathrm{d}t\Big|^2\hat{\phi}(\xi)\,\mathrm{d}\xi$$
$$\sim\frac{T}{2\pi\log T}\hat{\phi}(0)\int_{\xi\in\mathbb{R}}|h(\xi)|^2|\Omega_{\mathbb{R}}(\xi T)|^2\,\mathrm{d}\xi\sim\frac{T\log T}{2\pi}\|h\|_2^2\hat{\phi}(0)\quad(T\to\infty).$$

Ensuite, la contribution de la somme de (1.8) est

$$\Big(\frac{T}{2\pi}\Big)^2\sum_{m,n}\frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}}\int_{\xi\in\mathbb{R}}\sum_{(\varepsilon_1,\varepsilon_2)\in\{-,+\}^2}C^{\varepsilon_1,\varepsilon_2}(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi$$

avec

$$C^{\varepsilon_1,\varepsilon_2}(\xi)=\hat{h}\Big[(\xi\log T+\varepsilon_1\log n)\frac{T}{2\pi}\Big]\hat{h}\Big[(\xi\log T+\varepsilon_2\log n)\frac{T}{2\pi}\Big].$$

On examine la contribution de $C^{-,-}$. On a

$$\Big(\frac{T}{2\pi}\Big)^2\sum_{m,n}\frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}}\int_{\xi\in\mathbb{R}}C^{-,-}(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi$$
$$=\frac{T}{2\pi\log T}\sum_{m,n}\frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}}$$
$$\int_{u\in\mathbb{R}}\hat{h}(u)\hat{h}\Big[u+\frac{T}{2\pi}\log\Big(\frac{m}{n}\Big)\Big]\hat{\phi}\Big[\Big(\frac{2\pi}{T}u+\log n\Big)\frac{1}{\log T}\Big]\,\mathrm{d}u.$$

Par compacité du support de $\hat{h}$, les seuls entiers m et n contribuant vérifient

$$\log m-\log n\ll\frac{1}{T}.\tag{1.9}$$

Si $\mathrm{Supp}(\hat{\phi})=[-1+\delta,1-\delta]$ avec $\delta>0$, le seul entier n contribuant vérifie

$$\log n\le\log T^{1-\delta/2}\tag{1.10}$$

pour T assez grand. Les conditions (1.9) et (1.10) impliquent $m=n$ grâce au lemme suivant.

Lemme 1.13. *Soit $\delta>0$, $C>0$. Alors pour tout T assez grand, les conditions $|\log(m/n)|\le C/T$ et $\log n\le(1-\delta)\log T$ impliquent $m=n$.*

Démonstration. Si $m=n+u$ avec $u\ge1$ alors

$$\log m-\log n\ge\frac{u}{n}\ge\frac{1}{T^{1-\delta}}$$

ce qui n'est pas compatible avec

$$\frac{C}{T}\ge\log m-\log n.$$

Le cas $u<0$ est semblable. □

On a donc

$$\begin{aligned}\left(\frac{T}{2\pi}\right)^2 \sum_{m,n} \frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}} \int_{\xi\in\mathbb{R}} C^{-,-}(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi \\ = \frac{T}{2\pi\log T} \sum_n \frac{\Lambda(n)^2}{n} \int_{\mathbb{R}} \hat{h}(u)^2 \hat{\phi}\left[\left(\frac{2\pi}{T}u + \log n\right)\frac{1}{\log T}\right]\mathrm{d}u \qquad (1.11) \\ \sim \frac{T\|h\|_2^2}{2\pi\log T} \sum_n \frac{\Lambda(n)^2}{n} \hat{\phi}\left(\frac{\log n}{\log T}\right) \quad (T\to\infty)\end{aligned}$$

car $\|\hat{h}\|_2^2 = \|h\|_2^2$. Le théorème des nombres premiers avec reste conduit immédiatement à

$$\begin{aligned}\sum_{n\le x} \Lambda(n)^2 &= \sum_{p\le x} \log^2 p + O(x^{1/2}\log^2 x) \\ &= x\log x + O(x).\end{aligned}$$

Par intégration par parties, on a alors

$$\sum_{n\le x} \frac{\Lambda(n)^2}{n} \sim \frac{1}{2}\log^2 x \quad (x\to\infty).$$

Une nouvelle intégration par parties donne donc

$$\begin{aligned}\sum_n \frac{\Lambda(n)^2}{n} \hat{\phi}\left(\frac{\log n}{\log T}\right) &\sim \log^2 T \int_{v>0} \hat{\phi}(v) v\,\mathrm{d}v \\ &= \frac{1}{2}\log^2 T \int_{v\in\mathbb{R}} \hat{\phi}(v)|v|\,\mathrm{d}v \quad (T\to\infty).\end{aligned}$$

Finalement,

$$\left(\frac{T}{2\pi}\right)^2 \sum_{m,n} \frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}} \int_{\xi\in\mathbb{R}} C^{-,-}(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi \sim \frac{T\log T}{4\pi}\|h\|_2^2 \int_{\xi\in\mathbb{R}} \hat{\phi}(\xi)|\xi|\,\mathrm{d}\xi \quad (T\to\infty).$$

De même,

$$\left(\frac{T}{2\pi}\right)^2 \sum_{m,n} \frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}} \int_{\xi\in\mathbb{R}} C^{+,+}(\xi)\hat{\phi}(\xi)\,\mathrm{d}\xi \sim \frac{T\log T}{4\pi}\|h\|_2^2 \int_{\xi\in\mathbb{R}} \hat{\phi}(\xi)|\xi|\,\mathrm{d}\xi \quad (T\to\infty).$$

Pour $C^{\pm,\mp}$, la condition sur le support de $\hat{h}$ implique que $T|\log m + \log n|$ est bornée d'où $m = n = 1$ et il n'y a pas de contribution. La contribution arithmétique est donc

$$\frac{T\log T}{2\pi}\|h\|_2^2 \int_{\mathbb{R}} |\xi|\hat{\phi}(\xi)\,\mathrm{d}\xi$$

ce qui donne le terme $|\xi|$ dans K.

Remarque 1.14. La restriction du support de $\hat{\phi}$ à $]-1,1[$ permet, dans le calcul de (1.11), de ne prendre en compte que les termes diagonaux, *i.e.* ceux où $m = n$. Un élargissement du support de $\hat{\phi}$ impose donc de savoir évaluer les termes non diagonaux. Pour cela, il faut comprendre les sommes du type

$$\sum_{n,d} \frac{\Lambda(n)\Lambda(n+d)}{n},$$

c'est-à-dire la corrélation entre les nombres premiers. Plus précisément, on peut montrer (voir [31], [33]) que si on connaît l'équivalent

$$\int_1^X \Big[\sum_{x<n\le x+h} \Lambda(n) - h \Big]^2 \mathrm{d}x \sim hX \log \frac{X}{h} \quad (1 \le h \le X^{1-\varepsilon},\ X \to +\infty)$$

alors, on peut étendre le support de $\hat{\phi}$ autant qu'on veut.

La remarque précédente conduit à la conjecture de Montgomery.

Conjecture de Montgomery. *Soit ϕ une fonction de localisation. Alors*

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{1\le j,k\le N} \phi(\tilde{\gamma}_j - \tilde{\gamma}_k) = \mathrm{COR}(2,\phi,\mathrm{univ}).$$

La conjecture de Montgomery a d'intéressantes conséquences.

Proposition 1.15. (1) *En admettant uniquement l'hypothèse de Riemann, au moins* $2/3$ *des points d'annulation de ζ sont d'ordre* 1: *il existe une fonction ε de limite nulle en $+\infty$ telle que*

$$\frac{1}{N(\zeta,T)} \#\big\{\gamma \in Z(\zeta) : 0 < \mathrm{Im}\,\gamma < T,\ \zeta'(\gamma) \ne 0\big\} \ge \frac{2}{3} + \varepsilon(T).$$

(2) *Si de plus, la conjecture de Montgomery est vraie, les points d'annulation de ζ sont presque tous d'ordre* 1.

Démonstration. Pour ϕ, on choisit la fonction

$$\phi(x) = \Big[\frac{\sin(\pi\alpha x)}{\pi\alpha x}\Big]^2$$

avec $\alpha \in]0,1[$ à optimiser. La transformée de Fourier de ϕ est à support compact dans $]-\alpha,\alpha[$ et caractérisée sur son support par

$$\hat{\phi}(\xi) = \frac{1}{\alpha}\Big(1 - \frac{|\xi|}{\alpha}\Big) \quad (|\xi| < \alpha).$$

Ainsi,

$$\int_{\mathbb{R}} \phi\,\mathrm{dCOR}(2,\mathrm{univ}) = \int_{\mathbb{R}} \hat{\phi}\,\widehat{\mathrm{dCOR}(2,\mathrm{univ})} = \frac{\alpha}{3} - 1.$$

Grâce au théorème 1.3, et compte-tenu des remarques 1.6 et 1.12, on a donc

$$\frac{1}{N(\zeta,T)}\sum_{\substack{j\in\mathbb{Z}^*\\ 0<\operatorname{Im}\gamma_j<T}}\sum_{\substack{k\in\mathbb{Z}^*\\ 0<\operatorname{Im}\gamma_k<T}}\left[\frac{\sin\left(\frac{\alpha\log T}{2}(\gamma_j-\gamma_k)\right)}{\frac{\alpha\log T}{2}(\gamma_j-\gamma_k)}\right]^2=\frac{1}{\alpha}+\frac{\alpha}{3}+\varepsilon(T).$$

En ne gardant, dans la somme sur k, que les valeurs de k telles que $\gamma_k=\gamma_j$, on a alors

$$\frac{1}{N(\zeta,T)}\sum_{\substack{j\in\mathbb{Z}^*\\ 0<\operatorname{Im}\gamma_j<T}}v_{\gamma_j}(\zeta)\leq\frac{1}{\alpha}+\frac{\alpha}{3}+\varepsilon(T)$$

en notant $v_\gamma(\zeta)$ l'ordre d'annulation de ζ en γ. En ne sommant que sur les points d'annulation, et non sur ces points comptés avec multiplicité, on a donc

$$\frac{1}{N(\zeta,T)}\sum_{\substack{\gamma\in Z(\zeta)\\ 0<\operatorname{Im}\gamma<T}}v_\gamma(\zeta)^2\leq\frac{1}{\alpha}+\frac{\alpha}{3}+\varepsilon(T).$$

Puisque

$$(2-v)v\begin{cases}\leq 0 & \text{si } v\geq 2,\\ =v & \text{si } v=1,\end{cases}$$

on a

$$\frac{1}{N(\zeta,T)}\sum_{\substack{\gamma\in Z(\zeta)\\ 0<\operatorname{Im}\gamma<T}}\big(2-v_\gamma(\zeta)\big)\,v_\gamma(\zeta)\leq\frac{\#\left\{\gamma\in Z(\zeta)\colon 0<\operatorname{Im}\gamma<T,\ \zeta'(\gamma)\neq 0\right\}}{N(\zeta,T)}$$

ou

$$\frac{1}{N(\zeta,T)}\sum_{\substack{\gamma\in Z(\zeta)\\ 0<\operatorname{Im}\gamma<T}}\big(2-v_\gamma(\zeta)\big)\,v_\gamma(\zeta)\geq 2-\frac{1}{\alpha}-\frac{\alpha}{3}+\varepsilon(T),$$

d'où

$$\frac{\#\left\{\gamma\in Z(\zeta)\colon 0<\operatorname{Im}\gamma<T,\ \zeta'(\gamma)\neq 0\right\}}{N(\zeta,T)}\geq\frac{2}{3}+\varepsilon(T)$$

avec le choix optimal (compte-tenu des contraintes) $\alpha\to 1^-$. Si la conjecture de Montgomery est vraie, on peut choisir $\alpha>1$. Il faut alors remplacer $\frac{1}{\alpha}+\frac{\alpha}{3}$ par $1+\frac{1}{3\alpha^2}$ et on obtient que presque tous les points d'annulation de ζ sont simples avec $\alpha\to+\infty$. □

Remarque 1.16. À l'aide d'un procédé d'optimisation décrit, par exemple dans [47, Appendix A], on peut remplacer le choix de ϕ dans la démonstration de la proposition 1.15 par le choix optimal

$$\phi(x)=\frac{1}{1-\cos(\sqrt{2})}\left[\frac{\sin\left(\frac{\sqrt{2}-2\pi x}{2}\right)}{\sqrt{2}-2\pi x}+\frac{\sin\left(\frac{\sqrt{2}+2\pi x}{2}\right)}{\sqrt{2}+2\pi x}\right]^2$$

dont la transformée de Fourier, de support $[-1, 1]$ est caractérisée par

$$\hat{\phi}(\xi) = \frac{1}{1-\cos(\sqrt{2})}\left[\frac{1}{2\sqrt{2}}\sin(\sqrt{2}(1-|\xi|)) + \frac{1-|\xi|}{2}\cos(\sqrt{2}\xi)\right] \quad (|\xi| \leq 1).$$

Le résultat est légèrement meilleur ($\frac{3}{2} - \frac{\sqrt{2}}{2}$ cotan $\frac{\sqrt{2}}{2} \simeq 0,6725$ au lieu de $2/3$) [65], [10]. Cependant, en faisant, en plus de l'hypothèse de Riemann, une hypothèse sur les moments d'ordre 6 des fonctions L de caractères de Dirichlet (cette hypothèse étant conséquence de l'hypothèse de Lindelöf sur ces fonctions L), Conrey, Ghosh & Gonek montrent qu'on peut remplacer $\frac{2}{3}$ par $\frac{19}{27} \simeq 0,7$ [17]. Leur méthode est différente de celle exposée ici. Notons que pour obtenir ce résultat avec la méthode de Montgomery, il suffirait de connaître la conjecture de Montgomery pour ϕ telle que $\operatorname{Supp}\hat{\phi} \subset]-\frac{3}{2}, \frac{3}{2}[$.

Parmi les problèmes délicats de théorie des nombres, il y a ceux mélangeant les structures additives et multiplicatives des entiers. Par exemple, une conjecture affirme que, pour tout entier $n > 0$, il existe un nombre premier dans $]n^2, (n+1)^2[$. Cette conjecture est impliquée par la majoration $p' - p \leq p^A$ avec $A = 1/2$ pour tout nombre premier p où p' désigne le plus petit nombre premier strictement supérieur à p. Un résultat de Goldston & Heath-Brown [36] fait un pas vers cette conjecture (voir aussi [68], [59]).

Proposition 1.17. *Si la conjecture de Montgomery est vraie, alors*

$$\lim_{\substack{p\to+\infty \\ p\in\mathcal{P}}} \frac{\min\{p' \in \mathcal{P} : p' > p\} - p}{\sqrt{p \log p}} = 0.$$

Pour terminer cette partie, on reproduit, figure 4, un graphique de Gourdon [34] montrant l'erreur entre les données expérimentales et la conjecture de Montgomery. Plus précisément, Gourdon & Demichel ont calculé 2×10^9 zéros de hauteur d'ordre 10^{16} puis 2×10^9 zéros de hauteur d'ordre 10^{20} et 2×10^9 zéros de hauteur d'ordre 10^{24}. Pour chacune de ces trois séries, ils ont calculé la densité de probabilité des espacements entre zéros normalisés puis soustrait à cette densité la densité conjecturée dCOR(2, univ). Autrement dit, ils représentent (vraisemblablement avec d'autres nombres que 4 et 200) les points de coordonnées $(a_i, d_i^{\exp} - d_i^{\text{th}})$ pour $i \in \{1, \ldots, 200\}$ avec

$$a_i = \frac{4i}{200},$$

$$d_i^{\exp} = \frac{\#\{(j,k) \in \mathbb{N}^2 : h < j < k \leq h + 2\cdot 10^9,\ a_i < \hat{\gamma}_k - \hat{\gamma}_j < a_i + \delta\}}{2\cdot 10^9}$$

pour $h \in \{10^{16}, 10^{20}, 10^{24}\}$, δ petit et

$$d_i^{\text{th}} = \text{dCOR}(2, \text{univ})(a_i) \approx \int_{\mathbb{R}} \phi_i \, \text{dCOR}(2, \text{univ})$$

avec ϕ_i la fonction caractéristique de $[a_i, a_i + \delta]$.

La figure 4 reproduit le résultat obtenu.

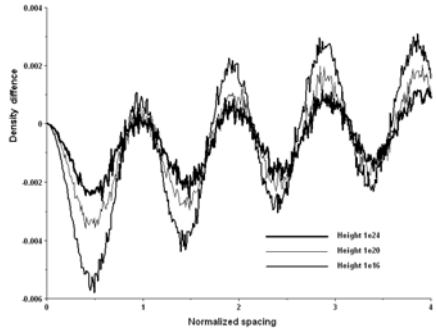


Figure 4. Différence entre densités de corrélation expérimentale et théorique.

2 Corrélations d'ordres supérieurs

Le théorème de Montgomery a pour objet la corrélation des paires de zéros normalisés : il permet l'étude de la statistique des écarts entre les zéros normalisés. On introduit la *corrélation de niveau n*. Elle permet d'étudier la distribution des sous-ensembles de taille n d'une suite donnée. On va restreindre notre étude à la distribution des écarts entre les éléments des sous-ensembles de taille n d'une suite donnée.

Soit $n \geq 2$ un entier et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction vérifiant les conditions suivantes

(1) f est symétrique,

(2) $f(x_1 + t, x_2 + t, \ldots, x_n + t) = f(x_1, x_2, \ldots, x_n)$ pour tout $t \in \mathbb{R}$.

Soit $\mathbf{u} = (u_n)_{n\in\mathbb{Z}}$ une suite réelle. La première condition permet de voir f comme une fonction sur les sous-ensembles de $\mathbf{u}(\mathbb{Z})$ de cardinal n :

$$f(\mathbf{u}(S)) = f(u_{i_1}, \ldots, u_{i_n}) \quad \text{si} \quad S = \{i_1, \ldots, i_n\}.$$

La deuxième condition permet d'exprimer le fait que f ne dépend que des différences successives entre ses arguments : si $\ell_i = x_{i+1} - x_i$ pour tout $i \in \{1, \ldots, n-1\}$, le choix de $t = -x_1$ conduit en effet à

$$f(x_1, \ldots, x_n) = f(0, \ell_1, \ell_1 + \ell_2, \ldots, \ell_1 + \ell_2 + \cdots + \ell_{n-1}).$$

On peut, par exemple, construire f à partir de n'importe quelle fonction $g : \mathbb{R} \to \mathbb{R}$ en posant

$$f(x_1, \ldots, x_n) = \prod_{1\leq i<j\leq n} g(|x_i - x_j|). \tag{2.1}$$

La *corrélation de niveau n* de la suite $\mathbf{u}$ est alors la limite de

$$R_n[M;\, f](\mathbf{u}) = \frac{n!}{M} \sum_{\substack{S\subset\{1,\dots,M\}\\ \#S=n}} f\left(\mathbf{u}(S)\right) \tag{2.2}$$

lorsque $M \to +\infty$ dans $\mathbb{N}$ (et lorsque cette limite existe). On a

$$\begin{aligned} R_n[M;\, f](\mathbf{u}) &= \frac{n!}{M} \sum_{1\le i_1<\cdots<i_n\le M} f(u_{i_1},\dots,u_{i_n}) = \frac{1}{M} \sum_{\substack{1\le i_1,\dots,i_n\le M\\ \text{distincts}}} f(u_{i_1},\dots,u_{i_n}) \\ &= \frac{n!}{M} \sum_{1\le i_1<\cdots<i_n\le M} f(0, u_{i_2}-u_{i_1},\dots,u_{i_n}-u_{i_1}). \end{aligned}$$

Remarque 2.1. Pour la compréhension du terme $1/M$ en facteur de la somme, lire la remarque 3.5.

On remplace la suite $\mathbf{u}$ par la suite $\tilde{Z}(\zeta)$ des zéros normalisés de ζ. Rudnick & Sarnak [71, Theorem 1.1] ont calculé les corrélations de tous niveaux de cette suite.

Théorème 2.2. *Soit $n \ge 2$ un entier. Soit h la transformée de Fourier inverse d'une fonction C^∞ à support compact et f une fonction de Schwartz dont la transformée de Fourier*

$$\xi \mapsto \int_{\mathbb{R}^n} f(x) e(-\xi\cdot x)\,\mathrm{d}x$$

est à support dans le domaine

$$\{(\xi_1,\dots,\xi_n)\in\mathbb{R}^n : |\xi_1|+\cdots+|\xi_n|\le 2\}.$$

On suppose que les trois conditions suivantes sont satisfaites :

- *la fonction f est symétrique,*
- *pour tout $t\in\mathbb{R}$, on a $f(x_1+t,\dots,x_n+t) = f(x_1,\dots,x_n)$,*
- *la fonction f tend rapidement vers 0 quand $|x|, \to \infty$ dans l'hyperplan d'équation $x_1+\cdots+x_n=0$.*

Alors

$$\begin{aligned} &\lim_{T\to+\infty} \frac{1}{N(\zeta,T)} \sum_{\substack{(j_1,\dots,j_n)\in\mathbb{Z}^n\\ \textit{distincts}}} h\left(\frac{\gamma_{j_1}}{T}\right)\dots h\left(\frac{\gamma_{j_n}}{T}\right) f\left(\frac{\log T}{2\pi}(\gamma_{j_1},\dots,\gamma_{j_n})\right) \\ &\qquad = \left(\int_{\mathbb{R}} h^n\right) \mathrm{COR}(n,\, f, \mathrm{univ}) \end{aligned}$$

avec

$$\mathrm{COR}(n,\, f, \mathrm{univ}) = \int_{\mathbb{R}^{n-1}} f(0,x_2,\dots,x_n)\,\mathrm{dCOR}(n,\mathrm{univ})(0,x_2,\dots,x_n)\,\mathrm{d}x_2\dots\mathrm{d}x_n$$

et

$$\mathrm{dCOR}(n, \mathrm{univ})(x_1, x_2, \ldots, x_n) = \det\left(\frac{\sin[\pi(x_i - x_j)]}{\pi(x_i - x_j)}\right)_{1\leq i,j\leq n}.$$

On peut en déduire les corrélations de $\tilde{Z}(\zeta)$ [71, Theorem 1.2].

Corollaire 2.3. *Soit $n \geq 2$ un entier. Soit f comme dans l'énoncé du théorème* 2.2. *Alors*

$$\lim_{N\to+\infty} R_n[N; f]\left(\tilde{Z}(\zeta)\right) = \mathrm{COR}(n, f, \mathrm{univ}).$$

Bomolgony & Keating [5], [6] donnent des arguments heuristiques (de nombreuses approximations sont faites sans preuve) permettant de traîter les termes non diagonaux apparaissant dans le calcul des corrélations (voir remarque 1.14) à partir d'une évaluation conjecturale de

$$\pi(m, n, k, N) = \#\{(p, q) \in \mathcal{P} \times \mathcal{P} : p < N,\ mp - nq = k\}.$$

Leurs arguments, faisant intervenir de délicats calculs combinatoires, permettent de penser que dans le corollaire 2.3, on peut supprimer la condition de support de $\hat{f}$.

3 Le lien avec les matrices aléatoires

3.1 Statistiques discriminantes

On note I_N la matrice identité $N \times N$. Lorsque la taille est implicite, on supprime l'indice et on note plutôt I. Les groupes de matrices qu'on va utiliser sont les suivants :

$\mathrm{U}(N)$	groupe des matrices A de taille $N \times N$ telles $A\,{}^t\bar{A} = I$,
$\mathrm{SO}(N)$	sous-groupe de $\mathrm{U}(N)$ constitué des matrices réelles de déterminant 1,
$\mathrm{USp}(2N)$	sous-groupe de $\mathrm{U}(2N)$ formé des matrices A vérifiant

$$A\begin{pmatrix} 0 & I_N \\ I_N & 0 \end{pmatrix} {}^tA = \begin{pmatrix} 0 & I_N \\ I_N & 0 \end{pmatrix}.$$

Ces groupes sont munis d'une mesure de Haar (mesure de probabilité invariante par translation à gauche). Si A est une matrice de $\mathrm{U}(N)$, son polynôme caractéristique se factorise sous la forme

$$\det(XI - A) = \prod_{j=1}^{N}\left(X - e^{i\phi_j(A)}\right)$$

avec $j \mapsto \phi_j(A)$ croissante à valeurs dans $[0, 2\pi[$. On note $\boldsymbol{\phi}_{\mathrm{U}}(A)$ le multiensemble des angles propres de A dans $[0, 2\pi[$:

$$\boldsymbol{\phi}_{\mathrm{U}}(A) = \{\phi_j(A)\}_{1\leq j\leq N} \subset [0, 2\pi[^N.$$

Si A est une matrice de $\mathrm{SO}(2N+1)$, son polynôme caractéristique se factorise sous la forme

$$\det(XI - A) = (X-1)\prod_{j=1}^{N}(X - e^{i\phi_j(A)})(X - e^{-i\phi_j(A)})$$

avec $j \mapsto \phi_j(A)$ croissante à valeurs dans $[0, \pi]$. On note $\boldsymbol{\phi}_{\mathrm{SO}^-}(A)$ le multiensemble des angles propres de A dans $[0, \pi]$, où la valeur propre 1 est comptée avec multiplicité diminuée de 1 :

$$\boldsymbol{\phi}_{\mathrm{SO}^-}(A) = \{\phi_j(A)\}_{1\le j\le N} \subset [0,\pi]^N.$$

Enfin, si A est une matrice de $\mathrm{SO}(2N)$ ou $\mathrm{USp}(2N)$, son polynôme caractéristique se factorise sous la forme

$$\det(XI - A) = \prod_{j=1}^{N}(X - e^{i\phi_j(A)})(X - e^{-i\phi_j(A)})$$

avec $j \mapsto \phi_j(A)$ croissante à valeurs dans $[0, \pi]$. On note $\boldsymbol{\phi}_G(A)$, avec $G = \mathrm{SO}$ ou $G = \mathrm{USp}$, le multiensemble des angles propres de A dans $[0, \pi]$:

$$\boldsymbol{\phi}_G(A) = \{\phi_j(A)\}_{1\le j\le N} \subset [0,\pi]^N.$$

Une fonction f de $\mathrm{GL}(N)$ est dite *centrale* si $f(h^{-1}gh) = f(g)$ pour toutes matrices g et h de $\mathrm{GL}(N)$. La valeur de $f(g)$ ne dépend alors que du multiensemble des valeurs propres de g. Par abus de langage, si G est l'un des groupes $\mathrm{U}(N)$, $\mathrm{SO}(2N)$, $\mathrm{SO}(2N+1)$, $\mathrm{USp}(2N)$ on dira qu'une fonction est *centrale sur* G si, pour toute matrice $g \in G$, la valeur $f(g)$ ne dépend que du multiensemble des valeurs propres de g. Soit f une fonction centrale sur $\mathrm{U}(N)$, si $A \in \mathrm{U}(N)$, la valeur $f(A)$ ne dépend que du multiensemble $\boldsymbol{\phi}_{\mathrm{U}}(A)$. On peut donc associer à f une fonction symétrique $\tilde{f}$ sur $[0, 2\pi[^N$. On a alors

$$\int_{\mathrm{U}(N)} f(A)\,\mathrm{d}_{\mathrm{U}}A = \int_{[0,2\pi[^N} \tilde{f}(\phi_1,\dots,\phi_N)\,P_U(\phi_1,\dots,\phi_N)\prod_{j=1}^{N}\mathrm{d}\phi_j$$

où $\mathrm{d}_{\mathrm{U}}A$ est la mesure de Haar sur $\mathrm{U}(N)$ et

$$P_U(\phi_1,\dots,\phi_N) = \frac{1}{(2\pi)^N N!}\prod_{1\le m<n\le N}|e^{i\phi_m} - e^{i\phi_n}|^2 \tag{3.1}$$

[79, Theorem 7.4.B]. Soit f une fonction centrale sur $G(N) \in \{\mathrm{SO}(2N), \mathrm{SO}(2N+1), \mathrm{USp}(2N)\}$, si $A \in G(N)$, la valeur $f(A)$ ne dépend que du multiensemble $\boldsymbol{\phi}_G(A)$ avec $G = \mathrm{SO}^+$ si $G(N) = \mathrm{SO}(2N)$, $G = \mathrm{SO}^-$ si $G(N) = \mathrm{SO}(2N+1)$ et $G = \mathrm{USp}$ si $G(N) = \mathrm{USp}(2N)$. On peut donc associer à f une fonction symétrique $\tilde{f}$ sur $[0,\pi]^N$. On a alors

$$\int_{G(N)} f(A)\,\mathrm{d}_G A = \int_{[0,\pi]^N} \tilde{f}(\phi_1,\dots,\phi_N)\,P_G(\phi_1,\dots,\phi_N)\prod_{j=1}^{N}\mathrm{d}\phi_j \tag{3.2}$$

où $\mathrm{d}_G A$ est la mesure de Haar sur $G(N)$ et

$$P_{\mathrm{SO}^+}(\phi_1,\dots,\phi_N) = \frac{2}{(2\pi)^N N!} \prod_{1\le m<n\le N} |2\cos\phi_m - 2\cos\phi_n|^2$$

$$P_{\mathrm{SO}^-}(\phi_1,\dots,\phi_N) = \frac{2^N}{\pi^N N!} \prod_{1\le m<n\le N} (2\cos\phi_m - 2\cos\phi_n)^2 \prod_{m=1}^{N} \sin^2\frac{\phi_m}{2}$$

$$P_{\mathrm{USp}}(\phi_1,\dots,\phi_N) = \frac{2^N}{\pi^N N!} \prod_{1\le m<n\le N} (2\cos\phi_m - 2\cos\phi_n)^2 \prod_{m=1}^{N} \sin^2\phi_m$$

[79, équations (9.15), (9.7) et Theorem 7.8.B]. Les formules d'intégration (3.1) et (3.2) sont classiquement appelées *formules de Weyl*.

Dans le but d'étudier la distribution des espaces entre angles propres de matrices unitaires, on introduit les corrélations des suites de ces angles propres. Pour comparer les différentes situations, on normalise les angles de façon à avoir un écart moyen entre angles consécutifs de 1. Si $A \in \mathrm{U}(N)$, ses valeurs propres $e^{i\phi_j(A)}$ parcourent le cercle unité dans le sens trigonométrique lorsque j croît de 1 à N : ainsi, la suite des espacements entre angles propres consécutifs est-elle la suite $\{s_j(A)\}_{1\le j\le N}$ avec

$$s_j(A) = \begin{cases} \phi_{j+1}(A) - \phi_j(A) & \text{si } 1 \le j \le N-1, \\ 2\pi + \phi_1(A) - \phi_N(A) & \text{si } j = N. \end{cases}$$

La moyenne de cette suite étant $\frac{2\pi}{N}$, on appellera angles normalisés les angles

$$\widetilde{\phi}_j(A) = \frac{N}{2\pi}\phi_j(A) \quad (A \in \mathrm{U}(N))$$

et on note $\widetilde{\boldsymbol{\phi}}_U(A)$ la suite des N angles normalisés. Pour $A \in \mathrm{SO}(2N)$ ou $A \in \mathrm{USp}(2N)$, la moyenne des écarts entre les angles consécutifs de

$$\{e^{i\phi_1(A)},\dots,e^{i\phi_N(A)},e^{-i\phi_N(A)},\dots,e^{-i\phi_1(A)}\}$$

le long du cercle unité est $\frac{2\pi}{2N}$. On appellera donc angles normalisés les angles

$$\widetilde{\phi}_j(A) = \frac{N}{\pi}\phi_j(A) \quad (A \in \mathrm{SO}(2N) \text{ ou } A \in \mathrm{USp}(2N))$$

et on note $\widetilde{\boldsymbol{\phi}}_G(A)$ la suite des N angles normalisés avec $G = \mathrm{SO}^+$ ou $G = \mathrm{USp}$ selon que $A \in \mathrm{SO}(2N)$ ou $A \in \mathrm{USp}(2N)$. Enfin, pour $A \in \mathrm{SO}(2N+1)$, la moyenne des écarts entre les angles consécutifs de $\{1, e^{i\phi_1(A)},\dots,e^{i\phi_N(A)},e^{-i\phi_N(A)},\dots,e^{-i\phi_1(A)}\}$ le long du cercle unité est $\frac{2\pi}{2N+1}$. On appellera donc angles normalisés les angles

$$\widetilde{\phi}_j(A) = \frac{N+1/2}{\pi}\phi_j(A) \qquad (A \in \mathrm{SO}(2N+1))$$

et on note $\widetilde{\boldsymbol{\phi}}_{\mathrm{SO}^-}(A)$ la suite des N angles normalisés.

Gaudin [26] (voir aussi les travaux de Dyson [22–24]) a calculé la limite, lorsque N tends vers $+\infty$, des corrélations des angles normalisés pour $\mathrm{U}(N)$. Katz & Sarnak ont montré que les corrélations étaient les mêmes pour les autres groupes [51, Proposition 5.10.3] et ont minoré la vitesse de convergence.

Théorème 3.1. *Soit $n \geq 2$ un entier. Soit $F\colon \mathbb{R}^n \to \mathbb{R}$ une fonction vérifiant les conditions suivantes* :

(1) *F est symétrique,*

(2) *$F(x_1 + t, x_2 + t, \ldots, x_n + t) = F(x_1, x_2, \ldots, x_n)$ pour tout $t \in \mathbb{R}$,*

(3) *il existe un réel $\alpha \geq 0$ tel que*

$$\sup_{1\leq i,j\leq n} |x_i - x_j| > \alpha \Rightarrow F(x_1, \ldots, x_n) = 0.$$

Soit $(G(N))_N$ l'une des suites de groupes

$$(\mathrm{U}(N))_N,\ (\mathrm{USp}(2N))_N,\ (\mathrm{SO}(2N))_N,\ (\mathrm{SO}(2N+1))_N.$$

On note $\lambda = 0$ sauf si $G(N) = \mathrm{SO}(2N+1)$ où $\lambda = 1/2$. Alors

$$\lim_{N\to+\infty} \int_{G(N)} R_n[N+\lambda; F](\widetilde{\boldsymbol{\phi}}_G(A))\, \mathrm{d}_G A = \mathrm{COR}(n, F, \mathrm{univ}).$$

Plus précisément, on a le résultat suivant [51, Proposition 5.10.3 et Proposition 5.11.2,]

Théorème 3.2. *Avec les notations du théorème* 3.1,

$$\Big| \int_{G(N)} R_n[N+\lambda; F](\widetilde{\boldsymbol{\phi}}_G(A))\, \mathrm{d}_G A - \mathrm{COR}(n, F, \mathrm{univ}) \Big| \leq n!(8\alpha)^{n-1} \|F\|_\infty \frac{\log N + (\pi\alpha)^2 + \alpha + 1}{N}$$

et

$$\int_{G(N)} \Big| R_n[N+\lambda; F](\widetilde{\boldsymbol{\phi}}_G(A)) - \int_{G(N)} R_n[N+\lambda; F](\widetilde{\boldsymbol{\phi}}_G(B))\, \mathrm{d}_G B \Big|^2 \mathrm{d}_G A \leq n![3(8\alpha)^{n-1} + 65(8\alpha)^{n-2}] \frac{\|F\|_\infty^2}{N}.$$

La comparaison des théorèmes 2.2 et 3.1, semble indiquer que les zéros normalisés de ζ ont même comportement statistique que les valeurs propres de matrices unitaires. L'universalité des densités $\mathrm{dCOR}(n, \mathrm{univ})$ (au regard des sous-groupes de $\mathrm{U}(N)$) ne permet cependant pas de déterminer le groupe qui permettrait de « mimer » ζ.

Les fonctions de corrélations permettent de définir d'autres statistiques. Considérons, par exemple, l'espacement. Soit $k \geq 1$ un entier. Si f est une fonction continue

à support compact dans $\mathbb{R}^+$, la k^{e} mesure d'espacement des angles propres normalisés d'une matrice $A \in G(N)$ pour $G(N) \in \{\mathrm{U}(N), \mathrm{SO}(2N), \mathrm{SO}(2N+1), \mathrm{USp}(2N)\}$ est

$$\mu_k(A; f) = \frac{1}{N-k} \sum_{i=1}^{N-k} f(\widetilde{\phi}_{i+k}(A) - \widetilde{\phi}_i(A)).$$

Katz & Sarnak [51, Theorem 1.2.3] démontrent alors le résultat suivant.

Théorème 3.3. *Soit $k \geq 1$ un entier. Il existe une mesure de probabilité, $\mu(k, \mathrm{univ})$, dont la fonction de distribution cumulative est continue, telle que pour toute fonction continue à support compact f, si $(G(N))_N$ est l'une des suites de groupes $(\mathrm{U}(N))_N$, $(\mathrm{USp}(2N))_N$, $(\mathrm{SO}(2N))_N$, $(\mathrm{SO}(2N+1))_N$ alors,*

$$\lim_{N \to +\infty} \int_{G(N)} \mu_k(A, f) \, \mathrm{d}_G A = \int_0^{+\infty} f \, \mathrm{d}\mu(k, \mathrm{univ}).$$

On connaît une minoration de la vitesse de convergence grâce au résultat suivant [51, 1.2.6].

Théorème 3.4. *Avec les notations du théorème* 3.3*, on a, pour tout $\varepsilon > 0$ la majoration*

$$\int_{G(N)} \sup_f \left[\mu_k(A; f) - \int_0^{+\infty} f \, \mathrm{d}\mu(k, \mathrm{univ})\right] \mathrm{d}_G A \ll_\varepsilon N^{-1/6+\varepsilon}$$

la borne supérieure étant prise sur toutes les fonctions continues à support compact dont la valeur absolue est majorée par 1.

On écrit ensuite quelques statistiques discriminantes. Soit $k \geq 1$ un entier. Si f est une fonction de $\mathbb{R}^+$ à support compact, la *mesure de répartition de la k^e valeur propre* dans $G(N)$ est

$$\nu_k(G(N); f) = \int_{G(N)} f(\widetilde{\phi}_k(A)) \, \mathrm{d}_G A.$$

Ensuite, si f est symétrique sur $\mathbb{R}^k$, continue et à décroissance rapide, la *distribution de répartition d'ordre k* des valeurs propres normalisées de $G(N)$ est

$$W_k(G(N); f) = \int_{G(N)} \sum_{1 \leq i_1 < \cdots < i_k \leq N} f(\widetilde{\phi}_{i_1}(A), \ldots, \widetilde{\phi}_{i_k}(A)) \, \mathrm{d}_G A.$$

Remarque 3.5. La différence entre la corrélation d'ordre k et la répartition d'ordre k réside essentiellement dans le choix de la fonction test f. La corrélation permet de mesurer la distribution des paquets de k angles propres dans des intervalles de longueurs d'ordre $1/N$, la répartition d'ordre k permet de mesurer la distribution des « petits angles propres », *i.e.* ceux se trouvant dans un intervalle centré en 0 de longueur environ $1/N$. En se souvenant que l'écart moyen entre deux angles propres est $1/N$, on s'attend à ce qu'il y ait N couples (i, j) tels que la différence $|\phi_i - \phi_j|$ soit

inférieure à $1/N$ et un angle dans l'intervalle $[0, 1/N[$. C'est la justification du terme $1/N$ en facteur de la somme dans la corrélation n'apparaissant pas dans le répartition d'ordre k.

Katz & Sarnak démontrent le théorème suivant [51, Proposition 7.5.6, Theorem AD.2.2].

Théorème 3.6. (1) *Soit $k \geq 1$ un entier. Il existe trois fonctions continues ν_k, $\nu_{k,+}$ et $\nu_{k,-}$ telles que, pour toute fonction f continue à croissance polynomiale,*

$$\lim_{N\to+\infty} \nu_k(G(N);\, f) = \begin{cases} \int_{[0,+\infty[} f(x)\nu_k(x)\,\mathrm{d}x & \text{si } (G(N))_N = (\mathrm{U}(N))_N, \\ \int_{[0,+\infty[} f(x)\nu_{k,+}(x)\,\mathrm{d}x & \text{si } (G(N))_N = (\mathrm{SO}(2N))_N, \\ \int_{[0,+\infty[} f(x)\nu_{k,-}(x)\,\mathrm{d}x & \text{si } (G(N))_N = (\mathrm{USp}(2N))_N \\ & \text{ou } (G(N))_N = (\mathrm{SO}(2N+1))_N. \end{cases}$$

(2) *Soit $k \geq 1$ un entier. Il existe trois fonctions, W_k, $W_{k,+}$ et $W_{k,-}$, telles que pour toute fonction continue à support compact f, on ait*

$$\lim_{N\to+\infty} W_k(G(N);\, f) = \begin{cases} \int_{[0,+\infty[^k} f(x)W_k(x)\,\mathrm{d}x & \text{si } (G(N))_N = (\mathrm{U}(N))_N, \\ \int_{[0,+\infty[^k} f(x)W_{k,+}(x)\,\mathrm{d}x & \text{si } (G(N))_N = (\mathrm{SO}(2N))_N, \\ \int_{[0,+\infty[^k} f(x)W_{k,-}(x)\,\mathrm{d}x & \text{si } (G(N))_N = (\mathrm{USp}(2N))_N \\ & \text{ou } (G(N))_N = ((2N+1))_N \end{cases}$$

avec

$$W_k(x_1,\dots,x_k) = \det\left(\frac{\sin[\pi(x_i-x_j)]}{\pi(x_i-x_j)}\right)_{1\leq i,j\leq n},$$

$$W_{k,\pm}(x_1,\dots,x_k) = \det\left(\frac{\sin[\pi(x_i-x_j)]}{\pi(x_i-x_j)} \pm \frac{\sin[\pi(x_i+x_j)]}{\pi(x_i+x_j)}\right)_{1\leq i,j\leq n}.$$

On peut déterminer les fonctions ν_k et $\nu_{k,\pm}$ de la façon suivante [51, Proposition 7.5.5]. Avec les mêmes notations que dans le point 2) du théorème 3.6 on pose

$$E(T,t) = 1+\sum_{k=1}^{+\infty} \frac{T^k}{k!} \int_{[0,t]^k} W_k(x)\,\mathrm{d}x, \quad E_\pm(T,t) = 1+\sum_{k=1}^{+\infty} \frac{T^k}{k!} \int_{[0,t]^k} W_{k,\pm}(x)\,\mathrm{d}x.$$

En développant ces fonctions autour de $T=-1$, on obtient

$$E(T,t) = \sum_{k=1}^{+\infty} (1+T)^k E_k(t), \quad E_\pm(T,t) = \sum_{k=1}^{+\infty} (1+T)^k E_{k,\pm}(t).$$

Les fonctions de distributions cumulatives cherchées sont alors

$$\int_0^t \nu_k(x)\,\mathrm{d}x = 1 - \sum_{k=1}^{n-1} E_k(t), \quad \int_0^t \nu_{k,\pm}(x)\,\mathrm{d}x = 1 - \sum_{k=1}^{n-1} E_{k,\pm}(t).$$

En particulier, ces formules peuvent être utilisées pour déterminer les premiers termes des développements limités de ν_1 et $\nu_{1,\pm}$ [61], [51, Appendix : Graphs]. On trouve

$$\lim_{N\to+\infty}\frac{1}{x}\operatorname{Haar}\{A\in \mathrm{U}(N)\colon \widetilde{\phi}_1(A)\in[0,x]\}=1-\frac{\pi^2}{36}x^3+O(x^5),$$
$$\lim_{N\to+\infty}\frac{1}{x}\operatorname{Haar}\{A\in \mathrm{SO}(2N)\colon \widetilde{\phi}_1(A)\in[0,x]\}=2-\frac{2\pi^2}{9}x^2+O(x^4),$$
$$\lim_{N\to+\infty}\frac{1}{x}\operatorname{Haar}\{A\in \mathrm{SO}(2N+1)\colon \widetilde{\phi}_1(A)\in[0,x]\}=\frac{2\pi^2}{9}x^2+O(x^4),$$
$$\lim_{N\to+\infty}\frac{1}{x}\operatorname{Haar}\{A\in \mathrm{USp}(2N)\colon \widetilde{\phi}_1(A)\in[0,x]\}=\frac{2\pi^2}{9}x^2+O(x^4).$$

Dans le cas de $\mathrm{SO}(2N+1)$ et $\mathrm{USp}(2N)$, il y a donc très peu de petites valeurs propres non triviales. Il y a répulsion (voir la remarque 1.7). Ce phénomène n'existe pas dans le cas de $\mathrm{U}(N)$ ou $\mathrm{SO}(2N)$. Dans le cas de $\mathrm{SO}(2N)$, il y a même deux fois plus de petites valeurs propres que dans le cas de $\mathrm{U}(N)$.

Enfin, on introduit une dernière statistique, très proche de la répartition de premier ordre. L'intérêt de cette nouvelle statistique est de permettre une discrimination de tous les groupes qu'on a introduit : l'observation de cette statistique sur les valeurs propres d'un des groupes $\mathrm{U}(N)$, $\mathrm{USp}(2N)$, $\mathrm{SO}(2N)$ ou $\mathrm{SO}(2N+1)$ permettra de retrouver le groupe étudié. On introduit une normalisation légèrement dífférente des angles propres. Soit $A\in\mathrm{U}(N)$, on étend son spectre en posant

$$\phi_{j+\ell N}(A)=\phi_j+2\ell\pi\quad (A\in\mathrm{U}(N),\ j\in\{1,\dots,N\},\ \ell\in\mathbb{Z}).$$

Si f est à décroissance rapide, on pose

$$D(f,\mathrm{U}(N),A)=\sum_{n\in\mathbb{Z}} f\Big(\frac{N}{2\pi}\phi_n(A)\Big).$$

Si $K\subset\mathrm{U}(N)$ est un sous-groupe compact, on définit

$$D(K;f)=\int_K D(f,\mathrm{U}(N),A)\,\mathrm{d}_K A.$$

Cette statistique joue le même rôle que la répartition d'ordre 1. Pour l'en distinguer, on lui réserve le nom de statistique des *petits angles propres*. Katz & Sarnak [51, Theorem AD.12.6] démontrent le résultat suivant.

Théorème 3.7. *Soit f une fonction à décroissance rapide. Alors*

$$\lim_{N\to+\infty} D(\mathrm{U}(N), f) = \int_{\mathbb{R}} f(t) D[U](t)\,\mathrm{d}t \quad \textit{avec } D[U](t) = 1,$$

$$\lim_{N\to+\infty} D(\mathrm{USp}(2N), f) = \int_{\mathbb{R}} f(t) D[\mathrm{USp}](t)\,\mathrm{d}t$$

$$\textit{avec } D[Usp](t) = 1 - \frac{\sin(2\pi t)}{2\pi t},$$

$$\lim_{N\to+\infty} D(\mathrm{SO}(2N), f) = \int_{\mathbb{R}} f(t) D[\mathrm{SO}^+](t)\,\mathrm{d}t$$

$$\textit{avec } D[\mathrm{SO}^+](t) = 1 + \frac{\sin(2\pi t)}{2\pi t},$$

$$\lim_{N\to+\infty} D(\mathrm{SO}(2N+1), f) = \int_{\mathbb{R}} f(t) D[\mathrm{SO}^-](t)\,\mathrm{d}t$$

$$\textit{avec } D[\mathrm{SO}^-](t) = 1 - \frac{\sin(2\pi t)}{2\pi t} + \delta(t).$$

Hughes & Rudnick [41, Theorem 6.2 et Theorem 6.5] ont calculé les premiers moments centrés réduits de la variable aléatoire limite de $D_N : A \mapsto D(f, \mathrm{U}(N), A)$. Posons

$$\sigma[\phi] = \sqrt{\int_{-1}^{1} \min(1, |u|)\hat{\phi}(u)^2\,\mathrm{d}u}.$$

Théorème 3.8. *Soit $m \geq 2$ un entier et f une fonction de localisation telle que* $\mathrm{Supp}\,\hat{\phi} \subset [-2/m, 2/m]$. *Alors,*

$$\lim_{N\to+\infty} \int_{\mathrm{U}(N)} \left[\frac{D(f, \mathrm{U}(N), A) - D(\mathrm{U}(N), f)}{\sigma[f]}\right]^m \mathrm{d}A = \begin{cases} \frac{m!}{2^{m/2}(m/2)!} & \textit{si } m \textit{ est pair,} \\ 0 & \textit{sinon.} \end{cases}$$

Il résulte de ce théorème que les premiers moments de la variable limite de D_N sont ceux de la loi normale. Il est assez rare, dans les théorèmes concernant les matrices aléatoires, de voir une contrainte sur le support de la fonction de localisation. Elle est ici essentielle : si pour f on prend la fonction caractéristique de $[-1, 1]$, la variable D_N est une fonction de comptage, donc une variable aléatoire discrète. Sa variable aléatoire limite ne peut donc pas suivre la loi normale et le théorème 3.8 ne peut pas s'étendre simultanément à tous les moments et toutes les fonctions de localisation. La variable aléatoire limite de D_N a un comportement « faussement normal »[2]. Pour les autres groupes, on pourra se reporter à [42].

[2]Hughes & Rudnick utilisent le terme « mock-Gaussian ». On pourrait aussi dire « se gaussant de la loi de Gauss ».

3.2 Moments de polynômes caractéristiques

Soit $G(N)$ l'un des groupes $\mathrm{U}(N)$, $\mathrm{USp}(2N)$, $\mathrm{SO}(2N)$ et $\mathrm{SO}(2N+1)$. Pour $s \in \mathbb{C}$, et $A \in G(N)$, on définit

$$L(A,s) = \det(I - As)$$

puis le moment complexe d'ordre s de $G(N)$ par

$$M[G(N);s] = \int_{G(N)} |L(A,1)|^s \, \mathrm{d}_G A.$$

Puisque $M[\mathrm{SO}(2N+1);s]$ est nul à cause de la valeur propre 1, on étudie plutôt :

$$M'[\mathrm{SO}(2N+1);s] = \int_{\mathrm{SO}(2N+1)} |L'(A,1)|^s \, \mathrm{d}A.$$

Keating & Snaith [52], [53] ont calculé ces moments.

Théorème 3.9. *Soit $s \in \mathbb{C}$, alors*

$$\begin{aligned}
M[\mathrm{U}(N);s] &= \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+s)}{\Gamma(j+s/2)^2} = g_{\mathrm{U}}(s) N^{(s/2)^2}[1+o(1)],\\
M[\mathrm{USp}(2N);s] &= 2^{Ns} \prod_{j=1}^{N} \frac{\Gamma(j+N+1)\Gamma(j+s+1/2)}{\Gamma(j+1/2)\Gamma(j+N+s+1)}\\
&= g_{\mathrm{USp}}(s) N^{s(s+1)/2}[1+o(1)],\\
M[\mathrm{SO}(2N);s] &= 2^{2Ns} \prod_{j=1}^{N} \frac{\Gamma(j+N-1)\Gamma(j+s-1/2)}{\Gamma(j-1/2)\Gamma(j+N+s-1)}\\
&= g_{\mathrm{SO}^+}(s) N^{s(s-1)/2}[1+o(1)],\\
M'[\mathrm{SO}(2N+1);s] &= \frac{1}{2} M[\mathrm{SO}(2N);s+1]
\end{aligned}$$

avec

$$\begin{aligned}
g_{\mathrm{U}}(s) &= \frac{G_2(1+s/2)^2}{G_2(1+s)},\\
g_{\mathrm{USp}}(s) &= 2^{s^2/2} \frac{G_2(1+s)\sqrt{\Gamma(1+s)}}{\sqrt{G_2(1+2s)\Gamma(1+2s)}},\\
g_{\mathrm{SO}^+}(s) &= 2^{s^2/2} \frac{G_2(1+s)\sqrt{\Gamma(1+2s)}}{\sqrt{G_2(1+2s)\Gamma(1+s)}}.
\end{aligned}$$

Remarque 3.10. La fonction G_2 est la fonction de Barnes[3], inverse de la fonction double gamma Γ_2. De façon générale, Vignéras [77, §2] montre que, pour tout entier $n \geq 1$, il existe une unique fonction G_n méromorphe telle que

[3] Puisqu'elle fut étudiée par le futur évêque Barnes [3].

(1) pour tout z complexe, $G_n(z+1) = G_{n-1}(z)G_n(z)$,

(2) $G_n(1) = 1$,

(3) G_n est indéfiniment dérivable sur l'intervalle réel $[1, +\infty[$ et

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} \ln G_n(x) \geq 0,$$

(4) $G_0(x) = x$ pour tout réel x.

La démonstration repose sur un théorème de Dufresnoy & Pisot [19] qui, en plus de l'existence et de l'unicité de G_n fournit le développement en série de Weierstraß. On trouve ainsi

$$G_2(1+z) = (2\pi)^{z/2} e^{-[(1+\gamma)z^2+z]/2} \prod_{n=1}^{+\infty} \left[\left(1 + \frac{z}{n}\right)^n e^{-z+z^2/(2n)} \right]$$

où

$$\gamma = \lim_{N\to+\infty} \sum_{n=1}^{N} \frac{1}{n} - \log N$$

est la constante d'Euler. La fonction G_2 est entière et tous ses zéros sont les entiers $-n \leq 0$ avec multiplicité $n+1$. Son logarithme est une fonction génératrice des valeurs de ζ aux entiers positifs puisque, pour $|z| < 1$, on a

$$\log G_2(1+z) = [\log(2\pi) - 1]\frac{z}{2} - (1+\gamma)\frac{z^2}{2} + \sum_{n=3}^{+\infty} (-1)^{n-1} \zeta(n-1) \frac{z^n}{n}.$$

Compte-tenu de $G_2(1) = 1$ et de la relation de récurrence, on a $G_2(2) = 1$ et

$$G_2(n) = \prod_{k=1}^{n-2} (k!) \quad \text{pour } n \geq 3. \tag{3.3}$$

Barnes a montré que

$$G_2\left(\frac{1}{2}\right) = A^{-3/2} \pi^{-1/4} e^{1/8} 2^{1/24} \tag{3.4}$$

où A est la constante de Glaisher définie par

$$\log A = \lim_{N\to+\infty} \left[\log \prod_{k=1}^{N} k^k - \left(\frac{N^2+N}{2} + \frac{1}{12} \right) \log N + \frac{N^2}{4} \right].$$

Glaisher [28] (voir aussi [76, Theorem 4.6] et [78, Appendix]) a montré que

$$\begin{aligned} A &= \exp\left(\frac{1}{12} - \zeta'(1)\right) = (2\pi)^{1/12} \exp\left(\frac{\gamma}{12} - \frac{\zeta'(2)}{2\pi^2}\right) \\ &= 1,28242712910062263687534256888\ldots. \end{aligned}$$

Adamchik [1, §6] a donné des méthodes de calculs avec grande précision de cette constante. On déduit de l'expression 3.4 et de la relation de récurrence que

$$G_2\left(n+\frac{1}{2}\right) = \pi^{n/2-1/4}2^{1/24-n(n-1)/2}e^{1/8}A^{-3/2}\prod_{j=1}^{n-1}[(2j-1)!!] \qquad (3.5)$$

pour tout entier $n \geq 1$ où le produit vide vaut 1 par convention et où $(2j-1)!! = 3\times 5\times\cdots\times(2j-1)$. La relation (3.3) conduit à

$$g_{\mathrm{U}}(2k) = \prod_{j=0}^{k-1}\frac{j!}{(k+j)!}$$

$$g_{\mathrm{U}}(2k+1) = \pi^{k+1/2}2^{1/12-3k^2/2-k/2}e^{1/4}A^{-3}k!\prod_{j=1}^{k}\frac{(2j-1)!!}{(2j)!j!}$$

$$g_{Usp}(2k) = \prod_{j=1}^{k}\frac{1}{(2j-1)!!}$$

$$g_{\mathrm{SO}^+}(2k) = 2^k\prod_{j=1}^{k-1}\frac{1}{(2j-1)!!}$$

pour tout entier $k \geq 1$.

Tab. 1. Quelques valeurs de $g_{\mathrm{U}}(k)$.

k	1	2	3	4
$g_{\mathrm{U}}(k)$	$\frac{\sqrt{\pi}e^{1/4}2^{1/12}}{A^3}$	1	$\frac{\pi^{3/2}2^{1/12}e^{1/4}}{8A^3}$	$\frac{1}{12}$
k	5	6	7	8
$g_{\mathrm{U}}(k)$	$\frac{\pi^{5/2}2^{1/12}e^{1/4}}{2048A^3}$	$\frac{1}{8640}$	$\frac{\pi^{7/2}2^{1/12}e^{1/4}}{50331648A^3}$	$\frac{1}{870912000}$

Plutôt que d'étudier les moments des polynômes caractéristiques, on peut étudier l'autocorrélation. Cela signifie qu'au lieu d'étudier l'espérance sur $\mathrm{U}(N)$ de la variable aléatoire $A \mapsto L(A,1)$ multipliée par elle-même, on calcule l'espérance sur $\mathrm{U}(N)$ d'un produit de variables aléatoires obtenues à partir de la variable aléatoire $A \mapsto L(A,1)$ par déformation. Ici, la déformation se fait en remplaçant la valeur 1 par un nombre complexe de module 1. C'est ce qu'on fait Conrey, Farmer, Keating, Rubinstein & Snaith [15] pour les sous-groupes compacts de $\mathrm{U}(N)$. On ne donne que le résultat concernant $\mathrm{U}(N)$ qui est essentiellement dû à Basor & Forrester [4] (voir aussi [9]).

Théorème 3.11. *Pour tout* $(\alpha_1, \dots, \alpha_{2k}) \in \mathbb{R}^{2k}$,

$$\int_{\mathrm{U}(N)} \prod_{j=1}^{k} L(A, e^{-\alpha_j}) \prod_{j=1}^{k} L({}^{t}\overline{A}, e^{\alpha_{k+j}})\, \mathrm{d}A$$
$$= \exp\Big(-\frac{N}{2}\sum_{j=1}^{k}\alpha_j + \frac{N}{2}\sum_{j=1}^{k}\alpha_{k+j}\Big)$$
$$\times \sum_{\sigma\in\Xi} \exp\Big(\frac{N}{2}\sum_{j=1}^{k}\alpha_{\sigma(j)} - \frac{N}{2}\sum_{j=1}^{k}\alpha_{\sigma(k+j)}\Big) \prod_{\substack{1\le\ell\le k\\ k+1\le m\le 2k}} \frac{1}{1-\exp[\alpha_{\sigma(m)}-\alpha_{\sigma(\ell)}]}.$$

L'ensemble Ξ *est l'ensemble des permutations* σ *de* $\{1, \dots, 2k\}$ *telles que*

$$\sigma(1) < \sigma(2) < \cdots < \sigma(k) \text{ et } \sigma(k+1) < \sigma(k+2) < \cdots < \sigma(2k).$$

3.3 Modélisation de ζ par les polynômes caractéristiques

En montrant que les variables aléatoires sur $\mathrm{U}(N)$ muni de la mesure de Haar

$$A \mapsto \operatorname{Re}\log L(A,1) \quad \text{et} \quad A \mapsto \operatorname{Im}\log L(A,1)$$

tendent vers deux variables indépendantes de loi normales d'espérance 0 et variance 1, Keating & Snaith [52] montrent que

$$\lim_{N\to+\infty} \operatorname{Haar}\Big\{A\in\mathrm{U}(N)\colon \frac{\log L(A,1)}{\sqrt{\frac{1}{2}\log N}} \in B\Big\} = \frac{1}{2\pi}\iint_B \exp\Big(-\frac{x^2+y^2}{2}\Big)\,\mathrm{d}x\,\mathrm{d}y$$

pour tout rectangle B de $\mathbb{R}^2$. D'autre part, Selberg[4] a montré

$$\lim_{T\to+\infty} \#\Big\{t\in[T,2T]\colon \frac{\log\zeta\left(\frac{1}{2}+it\right)}{\sqrt{\frac{1}{2}\log\log T}} \in B\Big\} = \frac{1}{2\pi}\iint_B \exp\Big(-\frac{x^2+y^2}{2}\Big)\,\mathrm{d}x\,\mathrm{d}y.$$

Cela, ajouté aux calculs de corrélations, invite à modéliser ζ par les polynômes caractéristiques de $\mathrm{U}(N)$ en faisant l'identification

$$N \longleftrightarrow \log T. \tag{3.6}$$

D'autre part, $s \mapsto L(A,s)$ est une fonction entière. De plus, on a l'équation

$$s^{-N/2}L(A,s) = \varepsilon(A)\Big(\frac{1}{s}\Big)^{-N/2} L\Big({}^{t}\bar{A}, \frac{1}{s}\Big)$$

où $\varepsilon(A) = (-1)^N \det A$ est de module 1. Tous les points d'annulation de $L(A,s)$ sont sur le cercle unité, ce qui peut être comparé à l'hypothèse de Riemann. L'espacement entre les points d'annulation de $L(A,s)$ est $\frac{2\pi}{N}$ à comparer à l'espacement moyen $\frac{2\pi}{\log T}$ des points d'annulation de ζ à hauteur T. Là encore, l'identification se fait *via* (3.6).

[4] Dans un travail non publié. Voir [49] pour une démonstration et une intéressante extension.

Cette modélisation néglige évidemment les propriétés arithmétiques de ζ (ou bien, elle les met en valeur en creux). Keating & Snaith [52] montrent ainsi

$$\int_{\mathrm{U}(N)} [\mathrm{Im}\log L(A,1)]^2 \,\mathrm{d}A = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + o(1)$$

alors que Goldoston [30, (1.10)], en supposant vraie la conjecture de Montgomery, a montré

$$\frac{1}{T}\int_0^T \left[\mathrm{Im}\log\zeta\left(\frac{1}{2}+it\right)\right]^2 \mathrm{d}t = \frac{1}{2}\log\log\frac{T}{2\pi} + \frac{1}{2}(\gamma+1) + \sum_{m=2}^{+\infty}\sum_{p\in\mathcal{P}}\frac{1+m}{m^2}\frac{1}{p^m} + o(1).$$

Cependant, on va voir dans la partie suivante que cette modélisation est riche.

4 Moments de ζ

La conjecture de Lindelöf concerne la taille de $\zeta\left(\frac{1}{2}+it\right)$ pour t grand. Cette conjecture est que pour tout réel $\varepsilon > 0$, il existe une constante $C > 0$ telle que pour tout réel t tel que $|t| \geq 1$, on ait

$$\left|\zeta\left(\frac{1}{2}+it\right)\right| \leq Ct^{\varepsilon}.$$

Une façon d'étudier cette conjecture est d'étudier les moments d'ordres pairs :

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{2k} \mathrm{d}t.$$

Dès 1918, Hardy & Littlewood ont montré

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{2} \mathrm{d}t = \log(T)(1+o(1)).$$

On trouve une démonstration dans [75, Theorem 7.4] (voir aussi [43, Chapter 15]). En 1926, Ingham a calculé le moment d'ordre 4 suivant :

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{4} \mathrm{d}t = \frac{1}{2\pi^2}\log^4(T(1+o(1)).$$

On trouve une démonstration dans [43, Chapter 5]. Depuis, aucun moment d'ordre supérieur n'a été précisément évalué. Il a même fallu attendre jusqu'à très récemment pour avoir des conjectures complètes sur les moments supérieurs. Ainsi, en 1998, Conrey & Ghosh [16] ont conjecturé

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{6} \mathrm{d}t = \frac{1}{8640}\prod_{p\in\mathcal{P}}\left[\left(1-\frac{1}{p}\right)^4\left(1+\frac{4}{p}+\frac{1}{p^2}\right)\right]\log^9 T. \quad (4.1)$$

En 2000, Conrey & Gonek [18] ont conjecturé

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^8 \mathrm{d}t = \frac{1}{870912000}\prod_{p\in\mathcal{P}}\left[\left(1-\frac{1}{p}\right)^9\left(1+\frac{9}{p}+\frac{9}{p^2}+\frac{1}{p^3}\right)\right]\log^{16} T. \tag{4.2}$$

On donne les très grandes lignes du cheminement conduisant à cette conjecture. Si $z\in\mathbb{C}$, on définit les coefficients de Dirichlet de ζ^z en écrivant

$$\zeta(s)^z = \sum_{n=1}^{+\infty}\frac{d_z(n)}{n^s} \quad \text{pour } \mathrm{Re}\, s > 1.$$

En utilisant le produit eulérien, on voit que d_z est la fonction multiplicative définie pour toute puissance de nombre premier p^j par

$$d_z(p^j) = \frac{\Gamma(k+j)}{\Gamma(k)j!} \quad \text{pour tous } p\in\mathcal{P} \text{ et } j\in\mathbb{N}.$$

Pour $\mathrm{Re}\, s \in]0,1[$, on utilise une « équation fonctionnelle approchée[5] ». On obtient une égalité de la forme

$$\zeta(s)^k = D_{k,N}(s) + \omega_k(s)D_{k,M}(1-s) + E_k(s).$$

Dans cette équation, on a

$$D_{k,N}(s) = \sum_{n=1}^{+\infty}\frac{d_k(n)}{n^s}$$

et

$$MN = \left(\frac{t}{2\pi}\right)^k, \quad \omega_k(s) = \left[\pi^{s-1/2}\frac{\Gamma\left(\frac{1-s}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\right]^k$$

et $E_k(s)$ est un terme d'erreur. Conrey & Gonek conjecturent alors

$$\int_T^{2T}\left|\zeta\left(\frac{1}{2}+it\right)\right|^{2k} \sim \int_T^{2T}\left|D_{k,N}\left(\frac{1}{2}+it\right)\right|^2 + \int_T^{2T}\left|D_{k,M}\left(\frac{1}{2}+it\right)\right|^2 \tag{4.3}$$

pour $MN = [T/(2\pi)]^k$ et M et N supérieurs à $1/2$. *Via* l'étude des intégrales de polynômes de Dirichlet

$$\int_T^{2T}\left|\sum_{n=1}^{N} a_n n^{it}\right|^2 \mathrm{d}t$$

(voir, par exemple, [46, Chapter 9]), on en déduit

$$\int_T^{2T}\left|D_{k,N}\left(\frac{1}{2}+it\right)\right|^2 = \sum_{n=1}^{N}\frac{d_k(n)^2}{n}[T+o(n)] \sim \frac{a_k}{\Gamma(k^2+1)}T\log^{k^2} N$$

[5] Par cette expression, il faut comprendre qu'on utilise l'équation fonctionnelle pour transformer une intégrale contour de type transformée de Mellin autour de s en deux intégrales sur des droites verticales du demi-plan à droite de la bande critique ([62, Lecture I] ou [43, Chapter 4] pour la mise en œuvre).

avec

$$a_k = \prod_{p\in\mathcal{P}} \left[\left(1-\frac{1}{p}\right)^{k^2} \sum_{r=0}^{+\infty} \frac{d_k(p^r)^2}{p^r} \right] = \prod_{p\in\mathcal{P}} \left(1-\frac{1}{p}\right)^{(k-1)^2} \sum_{j=0}^{k-1} \binom{k-1}{j}^2 \frac{1}{p^j} \quad (4.4)$$

pour $N \ll T$. Cette dernière restriction oblige à avoir $k \leq 2$. Pour autoriser l'exploration de moments d'ordres supérieurs, Conrey & Gonek utilisent un résultat de Goldston & Gonek [32] permettant l'évaluation d'intégrales du type

$$\int_T^{2T} \left| \sum_{n=1}^{N} a_n n^{-1/2-it} \right|^2 \mathrm{d}t$$

dès lors qu'on sait évaluer

$$\sum_{n\leq x} a_n \quad \text{et} \quad \sum_{n\leq x} a_n a_{n+h}.$$

En utilisant la méthode du symbole δ de Duke, Friedlander & Iwaniec, Conrey & Gonek établissent une conjecture portant sur le cas où $a_k = d_k$. Cela les conduit à faire la conjecture suivante.

Conjecture 1. *Soit $N = T^{1+\eta}$ avec $\eta \in [0, 1]$. Alors,*

$$\int_T^{2T} \left| D_{k,N}\left(\frac{1}{2}+it\right)\right|^2 \mathrm{d}t \sim w_k(\eta) \frac{a_k}{\Gamma(k^2+1)} T \log^{k^2} T$$

où a_k est défini comme en (4.4) *et*

$$w_k(\eta) = (1+\eta)^{k^2} \left\{ 1 - k^3 \left(1 - \frac{1}{1+\eta}\right) - \sum_{n=1}^{k^2-1} \binom{k^2}{n+1} \gamma_k(n) \left[1 - \frac{1}{(1+\eta)^{n+1}} \right] \right\}$$

avec

$$\gamma_k(n) = (-1)^n \sum_{i=0}^{k} \sum_{j=0}^{k} \binom{k}{i} \binom{k}{j} \binom{n-1}{i-1, j-1, n-i-j+1}.$$

Remarque 4.1. Le coefficient multinomial $\binom{a}{b, c, a-b-c}$ est nul si l'un des ses coefficients est strictement négatif et défini par

$$\binom{a}{b, c, a-b-c} = \frac{a!}{b!c!(a-b-c)!} \quad \text{sinon.}$$

Pour $k = 3$ (donc pour le moment d'ordre 6), on a

$$w_3(\eta) = 1 + 9\,\eta + 36\,\eta^2 + 84\,\eta^3 + 126\,\eta^4 - 630\,\eta^5 + 588\,\eta^6 - 180\,\eta^7 + 9\,\eta^8 - 2\,\eta^9$$

et de $w_3(\eta) + w_3(1-\eta) = 42$, on déduit (4.1). Pour $k = 4$ (donc pour le moment d'ordre 8), on a

$$\begin{aligned} w_4(\eta) = 1 &+ 16\eta + 120\eta^2 + 560\eta^3 + 1820\eta^4 + 4368\eta^5 + 8008\eta^6 + 11440\eta^7 \\ &+ 12870\eta^8 + 11440\eta^9 - 152152\eta^{10} + 179088\eta^{11} - 78260\eta^{12} \\ &+ 14000\eta^{13} - 1320\eta^{14} + 16\eta^{15} - 3\eta^{16} \end{aligned}$$

et de $2w_4(1) = 24024$, on déduit (4.2).

Le coefficient numérique en facteur du coefficient a_k semble assez mystérieux. On va voir que la théorie des matrices aléatoires fournit ce coefficient. Pour la suite, on pose donc (en supposant l'existence de la limite)

$$g_k = \lim_{T\to+\infty} \frac{1}{a_k T \log^{k^2} T} \int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{2k} \mathrm{d}t.$$

Le coefficient a_k est apparu assez naturellement dans ce qui précède : sa nature est arithmétique ; le coefficient g_k va apparaître naturellement dans ce qui suit : sa nature est spectrale.

En utilisant la partie 3.3, on modélise la fonction ζ par les polynômes caractéristiques de matrices unitaires. L'exemple donné en fin de cette partie montre que cette modélisation ne prend pas en compte la spécificité arithmétique de ζ. Keating & Snaith ont alors énoncé la conjecture suivante.

Conjecture 2. *Soit $k \geq 1$ un entier. Alors*

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{2k} \mathrm{d}t \sim g_k a_k T \log^{k^2} T$$

lorsque $T \to +\infty$, avec

$$g_k = g_{\mathrm{U}}(2k) = \prod_{j=0}^{j-1} \frac{j!}{(k+j)!}$$

et

$$a_k = \left(1-\frac{1}{p}\right)^{(k-1)^2} \sum_{j=0}^{k-1} \binom{k-1}{j}^2 \frac{1}{p^j}.$$

Dans cette conjecture, on peut aisément remplacer k par n'importe quel nombre complexe de partie réelle $\mathrm{Re}\, k > -\frac{1}{2}$. La table 1 montre que cette conjecture est compatible avec celles de Conrey & Ghosh (équation(4.1)) et Conrey & Gonek (équation (4.2)) obtenues par des méthodes indépendantes de la théorie des matrices aléatoires.

Hughes a lui étudié les moments discrets et, *via* la modélisation par les matrices aléatoires, il obtient la conjecture suivante [40, Conjecture 5].

Conjecture 3. *Soit $\lambda \in \mathbb{C}$ tel que $\operatorname{Re}\lambda > -\frac{1}{2}$. Alors, uniformément en α tel que $|\alpha| \leq \frac{1}{2\pi}\log\frac{T}{2\pi}$ on a*

$$\lim_{T\to+\infty} \frac{2\pi}{T\log T\left(\log\frac{T}{2\pi}\right)^{\lambda^2}} \sum_{0<\gamma_n\leq T} \left|\zeta\left[\frac{1}{2}+i\left(\gamma_n+\frac{2\pi\alpha}{\log\frac{T}{2\pi}}\right)\right]\right|^{2\lambda} = g_{\mathrm{U}}(2\lambda)a(\lambda)F_\lambda(2\pi\alpha)$$

avec

$$F_\lambda(x) = \frac{x^2}{4}\left[j_\lambda\left(\frac{x}{2}\right)\right]^2 + \frac{x^2}{4}\left[j_{\lambda-1}\left(\frac{x}{2}\right)\right]^2 - \lambda x j_\lambda\left(\frac{x}{2}\right) j_{\lambda-1}\left(\frac{x}{2}\right),$$

où j_n est la fonction de Bessel sphérique de première espèce et niveau n.

Remarque 4.2. Soit J_n la fonction de Bessel de première espèce et de niveau n [35, § 8.4 et § 8.5] alors

$$j_n(z) = \sqrt{\frac{\pi}{2z}} J_{n+1/2}(z).$$

Par [35, 8.461] on a

$$\begin{aligned} j_n(z) = {} & \frac{1}{z}\sin\left(z-\frac{\pi}{2}n\right)\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\frac{(-1)^k(n+2k)!}{(2k)!(n-2k)!}\frac{1}{(2z)^{2k}} \\ & + \frac{1}{z}\cos\left(z-\frac{\pi}{2}n\right)\sum_{k=0}^{\lfloor\frac{n-1}{2}\rfloor}\frac{(-1)^k(n+2k+1)!}{(2k+1)!(n-2k-1)!}\frac{1}{(2z)^{2k+1}}. \end{aligned}$$

On en tire

$$j_n(z) = \begin{cases} (-1)^{(n+1)/2}\frac{\cos z}{z} + O\left(\frac{1}{z^2}\right) & n \text{ impair}, \\ (-1)^{n/2}\frac{\sin z}{z} + O\left(\frac{1}{z^2}\right) & n \text{ pair} \end{cases} \qquad (|z|\to\infty).$$

Par [35, 8.402] on a

$$j_n(z) = 2^n \sum_{m=0}^{+\infty} \frac{(-1)^m(n+m)!}{m!(2n+2m+1)!}\, z^{n+2m}.$$

Enfin, en lien avec le théorème 3.11, Conrey, Farmer, Keating, Rubinstein & Snaith [14, équation (2.2.15) et theorem 2.4.1] ont conjecturé l'autocorrélation associée à ζ.

Conjecture 4. *Soit* $(\alpha_1, \dots, \alpha_{2k}) \in]-\frac{1}{2}, \frac{1}{2}[^{2k}$. *Alors,*

$$\int_0^T \prod_{j=1}^{k} \zeta\Big(\frac{1}{2}+\alpha_j+it\Big)\prod_{j=1}^{k}\zeta\Big(\frac{1}{2}-\alpha_{k+j}-it\Big)\,\mathrm{d}t$$
$$= \int_0^T \exp\Big(-\frac{1}{2}\log\frac{t}{2\pi}\sum_{j=1}^{k}\alpha_j+\frac{1}{2}\log\frac{t}{2\pi}\sum_{j=1}^{k}\alpha_{k+j}\Big)$$
$$\times \sum_{\sigma\in\Xi}\exp\Big(\frac{1}{2}\log\frac{t}{2\pi}\sum_{j=1}^{k}\alpha_{\sigma(j)}-\frac{1}{2}\log\frac{t}{2\pi}\sum_{j=1}^{k}\alpha_{\sigma(k+j)}\Big)$$
$$\times A_k(\alpha_{\sigma(1)},\dots,\alpha_{\sigma(2k)})\prod_{1\le\ell,m\le k}\zeta(1-\alpha_{\sigma(k+m)}+\alpha_{\sigma(\ell)})[1+O(t^{-1/2+\varepsilon})]\,\mathrm{d}t$$

avec

$$A_k(u_1,\dots,u_{2k})$$
$$= \prod_{p\in\mathcal{P}}\prod_{1\le i,j\le k}\Big(1-\frac{1}{p^{1+u_i-u_{j+k}}}\Big)\int_0^1\prod_{j=1}^{k}\Big(1-\frac{e^{2i\pi\theta}}{p^{1/2+u_j}}\Big)^{-1}\Big(1-\frac{e^{-2i\pi\theta}}{p^{1/2-u_{k+j}}}\Big)^{-1}\mathrm{d}\theta.$$

La structure de l'autocorrélation de ζ est, de façon évidente, la même que celle de l'autocorrélation des polynômes caractéristiques de $\mathrm{U}(N)$ (voir théorème 3.11), au facteur arithmétique A_k près.

5 Fonctions L et matrices aléatoires

5.1 Un rapide survol des fonctions L

L'universalité des corrélations sur les sous-groupes de $\mathrm{U}(N)$ a son pendant sur les fonctions L que nous commençons par brièvement introduire. On renvoie à [62, Lecture I] et [11, Lecture 4] pour plus d'informations. Des exemples plus « concrets » de fonctions L seront étudiés plus en détails dans les parties suivantes et cette partie peut être omise en première lecture.

On va introduire la fonction L associée à une représentation irréductible, automorphe et cuspidale π de $\mathrm{GL}_m(\mathbb{Q})$. Pour $m = 1$, on obtient la fonction ζ et les fonctions L de Dirichlet associée à des caractères primitifs. Pour $m = 2$, on obtient les fonctions L des formes modulaires paraboliques primitives et des formes de Maass paraboliques primitives. Pour $m \ge 1$, ces fonctions L ont pour forme

$$L(\pi, s) = \prod_{p\in\mathcal{P}} L(\pi_p, s)$$

avec

$$L(\pi_p, s) = \prod_{j=1}^{m} [1 - \alpha_{\pi,j}(p) p^{-s}]^{-1}.$$

Pour tout p, le nombre complexe $\alpha_{\pi,j}(p)$ vérifie,

$$|\alpha_{\pi,j}(p)| < \sqrt{p}.$$

On associe à π, un entier $Q_\pi > 0$ appelé son conducteur. Lorsque p ne divise pas Q_π, les complexes complexe $\alpha_{\pi,j}(p)$ sont non nuls et la conjecture de Ramanujan–Petersson prédit $|\alpha_{\pi,j}(p)| = 1$. Cette conjecture est un résultat immédiat lorsque $m = 1$. Pour $m = 2$, dans le cas particulier des formes modulaires (holomorphes), c'est un résultat de Deligne très difficile. En toute généralité, le meilleur résultat à ce jour est

$$(p, Q_\pi) = 1 \Rightarrow |\alpha_{\pi,j}(p)| \leq p^{1/2-1/(m^2+1)}.$$

De même que pour ζ, on sait compléter $L(\pi, s)$. On associe à π un ensemble de m nombres complexes $(\mu_{\pi,j})_{1\leq j\leq m}$ et le facteur à l'infini

$$L(\pi_\infty, s) = \prod_{j=1}^{m} \Gamma_{\mathbb{R}}(s + \mu_{\pi,j}).$$

Remarque 5.1. Souvent, on peut regrouper les facteurs $\Gamma_{\mathbb{R}}$ grâce à la relation de Legendre

$$\Gamma\left(\frac{z}{2}\right)\Gamma\left(\frac{z+1}{2}\right) = \frac{\sqrt{\pi}}{2^{z-1}}\Gamma(z).$$

On a

$$\operatorname{Re}\mu_{\pi,j} > -\frac{1}{2}.$$

La fonction L complétée de π est

$$\Lambda(\pi, s) = L(\pi_\infty, s)L(\pi, s).$$

Cette fonction admet un prolongement en fonction entière (sauf si c'est ζ), elle est bornée dans les bandes verticales et satisfait à l'équation

$$\Lambda(\pi, s) = \tau_\pi Q_\pi^{-s} \Lambda(\tilde{\pi}, 1 - s)$$

où τ_π est un complexe non nul et $\tilde{\pi}$ est la contragrédiente de π. Les paramètres de $\tilde{\pi}$ se déduisent de ceux π par conjuguaison complexe,

$$\begin{aligned}
\alpha_{\tilde{\pi},j}(p) &= \overline{\alpha_{\pi,j}(p)} \\
\mu_{\tilde{\pi},j} &= \overline{\mu_{\pi,j}} \\
Q_{\tilde{\pi}} &= Q_\pi \\
\tau(\tilde{\pi}) &= Q_\pi/\tau(\pi).
\end{aligned}$$

Par définition, un zéro non trivial de $L(\pi, s)$ est un point d'annulation de $\Lambda(\pi, s)$. On note $1/2 + i\gamma_{\pi,j}$ ces zéros et l'hypothèse de Riemann pour $L(\pi, s)$ est $\gamma_{\pi,j} \in \mathbb{R}$.

On donne maintenant une formule explicite à la Riemann. Dans ce but, on introduit des coefficients apparaissant dans la dérivée logarithmique de $L(\pi, s)$. On pose

$$a_\pi(p^k) = \sum_{j=1}^{m} \alpha_{\pi,j}(p)^k.$$

On définit δ_ζ comme étant 1 si $L(\pi, s) = \zeta(s)$ et 0 sinon. On a alors

Proposition 5.2. *Soit H une fonction de localisation. Alors*

$$\begin{aligned}\sum_{j\in\mathbb{Z}^*} H(\gamma_{\pi,j}) = \delta_\zeta\Big[H\Big(-\frac{i}{2}\Big) + H\Big(\frac{i}{2}\Big)\Big] \\ + \frac{1}{2\pi}\int_{-\infty}^{+\infty} H(t)\Big\{\log Q_\pi + \sum_{j=1}^{m}\Big[\frac{\Gamma'_{\mathbb{R}}}{\Gamma_{\mathbb{R}}}\Big(\frac{1}{2} + \mu_{\pi,j} + it\Big) \\ + \frac{\Gamma'_{\mathbb{R}}}{\Gamma_{\mathbb{R}}}\Big(\frac{1}{2} + \overline{\mu_{\pi,j}} - it\Big)\Big]\Big\}\,\mathrm{d}t \\ - \frac{1}{2\pi}\sum_{n=1}^{+\infty}\Big[\frac{\Lambda(n)a_\pi(n)}{\sqrt{n}}\widehat{H}\Big(\frac{\log n}{2\pi}\Big) + \frac{\Lambda(n)\overline{a_\pi(n)}}{\sqrt{n}}\widehat{H}\Big(-\frac{\log n}{2\pi}\Big)\Big].\end{aligned}$$

Une conséquence de ce résultat est que

$$\#\Big\{\gamma \in [0, T]\colon \Lambda\Big(\frac{1}{2} + i\gamma, \pi\Big) = 0\Big\} \sim \frac{m}{2\pi} T\log T \quad (T \to \infty).$$

On note $N(\pi, T)$ la quantité de gauche. Un ingrédient de la démonstration du théorème de Montgomery était

$$\sum_{n\le x} \frac{\Lambda(n)^2}{n} \sim \frac{1}{2}\log^2 x \quad (x \to \infty).$$

Pour $L(\pi, s)$, on a l'équivalence

$$\sum_{n\le x} \frac{|\Lambda(n)a_\pi(n)|^2}{n} \sim \frac{1}{2}\log^2 x \quad (x \to \infty). \tag{5.1}$$

Cependant, cette équivalence est conditionnelle. Sa démonstration utilise l'hypothèse

$$\mathcal{H}\colon \text{ soit } k \ge 2, \text{ alors } \sum_{p\in\mathcal{P}} \frac{\log(p)a_\pi(p^k)}{p^k} < \infty.$$

Cette hypothèse est démontrée pour $m \le 3$. En toute généralité, elle est conséquence de la conjecture de Ramanujan–Petersson. Rudnick & Sarnak [71] ont démontré le résultat suivant qui étend le théorème 2.2.

Théorème 5.3. *Soit $n \geq 2$ un entier. Soit π une représentation unitaire irréductible automorphe cuspidale de $\mathrm{GL}_m(\mathbb{Q})$. Si $m > 3$, on admet l'hypothèse $\mathcal{H}$.*

Soit h la transformée de Fourier inverse d'une fonction C^∞ à support compact et f une fonction de Schwartz dont la transformée de Fourier

$$\xi \mapsto \int_{\mathbb{R}^n} f(x)e(-\xi \cdot x)\,\mathrm{d}x$$

est à support dans le domaine

$$\{(\xi_1, \dots, \xi_n) \in \mathbb{R}^n : |\xi_1| + \cdots + |\xi_n| \leq 2\}.$$

On suppose que les trois conditions suivantes sont satisfaites :

- *la fonction f est symétrique,*
- *pour tout $t \in \mathbb{R}$, on a $f(x_1 + t, \dots, x_n + t) = f(x_1, \dots, x_n)$,*
- *la fonction f tend rapidement vers 0 quand $|x| \to \infty$ dans l'hyperplan d'équation $x_1 + \cdots + x_n = 0$.*

Alors

$$\lim_{T\to+\infty} \frac{1}{N(\pi, T)} \sum_{\substack{(j_1,\dots,j_n)\in\mathbb{Z}^n \\ \text{distincts}}} h\left(\frac{\gamma_{\pi,j_1}}{T}\right)\cdots h\left(\frac{\gamma_{\pi,j_n}}{T}\right) f\left(\frac{\log T}{2\pi}(\gamma_{\pi,j_1}, \dots, \gamma_{\pi,j_n})\right)$$
$$= \left(\int_{\mathbb{R}} h^n\right) \mathrm{COR}(n, f, \mathrm{univ})$$

avec $\mathrm{COR}(n, f, \mathrm{univ})$ *définie au théorème* 2.2.

Ainsi, les corrélations ne permettent pas de distinguer les fonctions L. La raison essentielle est l'universalité de la formule (5.1). L'universalité des corrélations est donc conséquence du peu d'information arithmétique prise en considération.

5.2 Caractères de Dirichlet

L'objet de cette partie est un bref survol de la théorie des caractères de Dirichlet. Pour les détails, on renvoie à [46, Chapter 3 et §4.6]. Si q est un entier, un *caractère de Dirichlet modulo q* est un morphisme du groupe multiplicatif $(\mathbb{Z}/q\mathbb{Z})^\times$ dans le groupe multiplicatif $\mathbb{C}^\times$. On considère un tel caractère comme une fonction totalement multiplicative de $\mathbb{Z}$ dans $\mathbb{C}$ en posant

$$\begin{cases} \chi(n) = \chi(n + q\mathbb{Z}) & \text{si } (n, q) = 1, \\ \chi(n) = 0 & \text{sinon.} \end{cases}$$

Le *caractère principal modulo q* est le caractère χ_0 défini par $\chi_0(n) = 1$ si $(n, q) = 1$ et $\chi_0(n) = 0$ sinon. Le *conducteur* d'un caractère modulo q est le plus petit diviseur q^* de q tel que $\chi = \chi_0\chi^*$ où χ_0 est le caractère principal modulo q et χ^* est un caractère de module q^*. Un caractère modulo q est dit *primitif* si $q = q^*$. L'ensemble

des caractères modulo q a

$$\varphi(q) = q \prod_{\substack{p\in\mathcal{P}\\ p|q}} \left(1 - \frac{1}{p}\right)$$

éléments. Parmi ceux-ci, le nombre de caractères primitifs est

$$\varphi^*(q) = q \prod_{\substack{p\in\mathcal{P}\\ p|q\\ p^2\nmid q}} \left(1 - \frac{2}{p}\right) \prod_{\substack{p\in\mathcal{P}\\ p^2|q}} \left(1 - \frac{1}{p}\right)^2.$$

Soit $q > 1$ et χ un caractère primitif modulo q. Sa fonction L est

$$L(\chi, s) = \sum_{n=1}^{+\infty} \chi(n) n^{-s} = \prod_{p\in\mathcal{P}} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

Cette fonction se prolonge en une fonction entière. On pose

$$L(\chi_\infty, s) = L(\overline{\chi_\infty}, s) = \pi^{-s/2}\Gamma\left(\frac{s+\kappa}{2}\right)$$

avec

$$\kappa = \begin{cases} 0 & \text{si } \chi(-1) = 1 \\ 1 & \text{sinon.} \end{cases}$$

La fonction $L(\chi, s)$ satisfait alors l'équation fonctionnelle

$$L(\chi_\infty, s)L(\chi, s) = \varepsilon(\chi) q^{-s+1/2} L(\overline{\chi_\infty}, 1-s) L(\overline{\chi}, 1-s)$$

où

$$\varepsilon(\chi) = \frac{1}{i^\kappa \sqrt{q}} \sum_{n=0}^{q-1} \chi(n) \exp\left(2i\pi \frac{n}{q}\right)$$

est un nombre complexe de module 1.

De même que pour la fonction ζ de Riemann, on conjecture que les points d'annulation de $L(\chi, s)$ à l'intérieur de la bande critique $\operatorname{Re} s \in [0, 1]$ sont situés sur l'axe critique $\operatorname{Re} s = \frac{1}{2}$. Cette conjecture s'appelle l'hypothèse de Riemann pour $L(\chi, s)$. La suite des points d'annulation de $L(\chi, s)$ est

$$\bigcup_{j\in\mathbb{Z}^*} \left\{\frac{1}{2} + i\gamma_{\chi,j}\right\}$$

avec $\operatorname{Re} \gamma_{\chi,j} \geq 0$ pour tout $j > 0$. On retrouve évidemment le cas de la partie 5.1 restreinte à $m = 1$.

5.3 Familles de fonctions L de caractères de Dirichlet et matrices aléatoires

Dans cette partie, on admet l'hypothèse de Riemann pour les fonctions L de caractères de Dirichlet.

Si q est premier, l'ensemble $X^*(q)$ des caractères de Dirichlet primitifs modulo q est l'ensemble de tous les caractères modulo q sauf le caractère principal. Son cardinal est donc $q-2$ et on définit sur $X^*(q)$ un opérateur de moyenne en posant

$$\mathbb{E}_q(X) = \frac{1}{\#X^*(q)} \sum_{\chi \in X^*(q)} X(\chi)$$

pour toute $X\colon X^*(q) \to \mathbb{R}$. On l'appelle opérateur de moyenne car

$$\lim_{q\to+\infty} \mathbb{E}_q(\mathbf{1}) = 1$$

où $\mathbf{1}$ est l'application constante égale à 1 sur $X^*(q)$. On appelle alors moment centré réduit d'ordre m le terme

$$R_q^{(m)}(X) = \mathbb{E}_q\left(\left(\frac{X-\mathbb{E}_q(X)}{\sigma_q(X)}\right)^m\right)$$

où $\sigma_q(X)$ est la variance

$$\sigma_q(X) = \sqrt{\mathbb{E}_q((X-\mathbb{E}_q(X))^2)}.$$

Alors que les moments permettent d'étudier la distribution d'une variable aléatoire, les moments centrés réduits permettent, en normalisant l'espérance à 0 et la variance à 1, de comparer différentes distributions.

Pour $\chi \in X^*(q)$ et ϕ une fonction de localisation, on introduit

$$D[\phi](\chi) = \sum_{j\in\mathbb{Z}^*} \phi\left[\frac{\log(q)}{2\pi}\gamma_{\chi,j}\right].$$

Cette fonction compte les « petits zéros » de $L(\chi, s)$, *i.e.* ceux qui se trouvent dans un intervalle centré en $\frac{1}{2}$ et de longueur l'espacement moyen entre zéros. Le théorème ci-dessous, de Hughes & Rudnick [41, Theorem 3.1], implique que les « petits zéros » des fonctions L de caractères de Dirichlet se comportent comme les « petits angles propres » des matrices unitaires.

Théorème 5.4. *Soit ϕ une fonction de localisation telle que* $\operatorname{Supp}\hat{\phi} \in [-2,2]$*, alors lorsque q parcourt les nombres premiers,*

$$\lim_{q\to\infty} \mathbb{E}_q(D[\phi]) = \int_{\mathbb{R}} \phi(t) D[U](t)\,\mathrm{d}t.$$

On rappelle que $D[U]$ a été définie au théorème 3.7. Hughes & Rudnick calculent aussi la variance [41, Theorem 4.1].

Théorème 5.5. *Soit ϕ une fonction de localisation telle que* $\operatorname{Supp}\hat{\phi} \in [-1, 1]$*, alors lorsque q parcourt les nombres premiers,*

$$\lim_{q\to+\infty} \sigma_q(D[\phi]) = \sqrt{\int_{-1}^{1} |u|\hat{\phi}(u)^2 \,\mathrm{d}u}.$$

Hughes & Rudnick [41, Theorem 5.1] vont au delà du calcul de la moyenne et de la variance en calculant les premiers moments centrés réduits.

Théorème 5.6. *Soit $\alpha > 0$ et ϕ une fonction de localisation telle que* $\operatorname{Supp}\hat{\phi} \in [-\alpha, \alpha]$*. Soit $m < 2/\alpha$ un entier. Lorsque q parcourt les nombres premiers,*

$$\lim_{q\to+\infty} R_q^{(m)}(D[\phi]) = \begin{cases} \frac{m!}{2^{m/2}(m/2)!} & \text{si } m \text{ est pair,} \\ 0 & \text{sinon.} \end{cases}$$

Les premiers moments de la variable aléatoire limite centrée réduite de $D[\phi]$ sont donc ceux d'une variable aléatoire normale d'espérance 0 et variance 1. On trouve le même phénomène que pour les matrices de $\mathrm{U}(N)$ (voir théorème 3.8).

Si, pour ϕ, on prend la fonction indicatrice de l'intervalle $[-1, 1]$, la variable aléatoire $D[\phi]$ compte le nombre de zéros des fonctions L de caractères de Dirichlet dans $[-\log(q)/(2\pi), \log(q)/(2\pi)]$. La variable aléatoire limite lorsque $q \to +\infty$ est donc discrète, contrairement à la loi normale. Il est donc impossible que le théorème 5.6 s'étende à tous les moments et toutes les fonctions de localisation simultanément. On retrouve le même comportement « faussement normal » que pour les matrices aléatoires.

En corollaire du théorème 5.4, Hughes & Rudnick [41, Theorem 8.3] prouvent qu'il existe une proportion non nulle de caractères primitifs ayant un premier zéro plus petit que ce à quoi on peut s'attendre (*i.e.* l'espacement moyen $2\pi/\log(q)$).

Corollaire 5.7. *Admettons l'hypothèse Riemann pour les fonctions L de caractères primitifs. Soit $\beta \geq 0,633$. Lorque q parcourt les nombres premiers,*

$$\liminf_{q\to\infty} \frac{1}{q-2} \#\left\{\chi \neq \chi_0 \colon \gamma_{\chi,1} < \beta \frac{2\pi}{\log q}\right\} \geq P(\beta)$$

avec

$$P(\beta) = \frac{11\pi^2 - 3 - 72\beta^2 - 88\pi^2\beta^2 - 48\beta^4 + 176\pi^2\beta^4}{12\pi^2(4\beta^2-1)^2} > 0.$$

5.4 Formes modulaires

L'objet de cette partie est un bref survol de la théorie des formes modulaires[6]. Pour les détails, on renvoie à [63] et [60]. Soit $N \in \mathbb{N}$, on note

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z});\, c \equiv 0 \pmod{N} \right\}.$$

Ce groupe agit sur le demi-plan de Poincaré

$$\mathcal{H} = \{z \in \mathbb{C} :\ \mathrm{Im}\, z > 0\}$$

par

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d}.$$

Soit $k \geq 0$ fixé dans toute la suite, $\mathrm{SL}(2,\mathbb{Z})$ agit sur les fonctions holomorphes sur $\mathcal{H}$ par l'action

$$f_{|\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}(z) = (cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right).$$

Une fonction vérifie la *condition modulaire* de poids k sur $\Gamma_0(N)$ si

$$f_{|\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} = f \quad \text{pour tout } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N). \tag{5.2}$$

Remarques 5.8. (1) Soit $M \,|\, N$ et f vérifiant la condition modulaire de poids k sur $\Gamma_0(M)$. Alors, elle vérifie la condition modulaire de poids k sur $\Gamma_0(N)$.

(2) La matrice $-I$ est dans $\Gamma_0(N)$. Si f vérifie l'équation (5.2) on a donc, pour tout z dans le demi-plan de Poincaré

$$f_{|-I}(z) = f(z)$$

c'est-à-dire

$$(-1)^k f(z) = f(z).$$

Si k est impair, on a donc $f = 0$ et c'est la raison pour laquelle on se restreint à k pair.

(3) Si on s'autorise d'autres groupes que $\Gamma_0(N)$ et si on en choisit un ne contenant pas $-I$, il n'est pas nécessaire de choisir k pair.

Soit f vérifiant la condition modulaire de poids k sur $\Gamma_0(N)$. Si $M \in \mathrm{SL}(2,\mathbb{Z})$, la fonction $f_{|M}$ est périodique de période

$$u_M = \inf \left\{ u \in \mathbb{N}^* :\ \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \in M^{-1}\Gamma_0(N)M \right\}.$$

[6] Pour nous, de telles formes sont nécessairement holomorphes

On a donc un développement en série de Fourier

$$f_{|M}(z) = \sum_{n\in\mathbb{Z}} \hat{f}_M(n) \exp\left(\frac{2i\pi n}{u_M} z\right). \tag{5.3}$$

Remarques 5.9. (1) La condition modulaire de poids k implique que f_M ne dépend que de la classe de M modulo $\Gamma_0(N)$. Il n'y a donc qu'un nombre fini de développements de Fourier distincts construits à partir de f comme dans l'équation (5.3).

(2) Pour $M \in \Gamma_0(N)$, on notera $\hat{f}$ au lieu de $\hat{f}_M$. On a $u_M = 1$. Le développement obtenu est appelé *développement de f à l'infini*.

On dit que f est *nulle aux pointes* si

$$\hat{f}_M(n) = 0 \quad \text{pour tout } n \leq 0,\ M \in \mathrm{SL}(2,\mathbb{Z}).$$

Une *forme modulaire parabolique de poids k et de niveau N* est une fonction

(1) holomorphe sur $\mathcal{H}$,

(2) qui vérifie la condition modulaire de poids k sur $\Gamma_0(N)$,

(3) qui est nulle aux pointes.

L'ensemble des formes modulaires paraboliques de poids k et de niveau N est un espace vectoriel complexe, noté $S_k(N)$ et on peut vérifier qu'il est de dimension finie. On a l'approximation : pour tout $\varepsilon > 0$, il existe $c > 0$ tel que pour tout N,

$$\left|\dim S_k(N) - \frac{k-1}{12} N \prod_{p|N}\left(1+\frac{1}{p}\right)\right| \leq cN^{1/2+\varepsilon}.$$

On muni $S_k(N)$ d'une structure d'espace hermitien grâce au produit

$$(f,g) = \int_{\Gamma_0(N)\backslash\mathcal{H}} f(z)\overline{g(z)} y^k \frac{\mathrm{d}x\,\mathrm{d}y}{y^2}.$$

On cherche alors à construire une base orthogonale « naturelle ».

On sait construire des applications linéaires, dites *opérateurs de Hecke T_n* pour tout entier n. On définit

$$\begin{aligned} T_n :\ S_k(N) &\longrightarrow S_k(N) \\ \sum_m \hat{f}(m)\exp(2i\pi mz) &\longmapsto \sum_m \Big[\sum_{\substack{d|(m,n)\\(d,N)=1}} d^{k-1}\hat{f}\Big(\frac{mn}{d^2}\Big)\Big]\exp(2i\pi mz). \end{aligned}$$

Ces opérateurs vérifient la relation

$$T_m \circ T_n = \sum_{\substack{d|(m,n)\\(d,N)=1}} d^{k-1} T_{\frac{mn}{d^2}}$$

de sorte qu'ils *commutent*. D'autre part, ils sont presque tous *autoadjoints* :

$$\text{si } (n,N) = 1 \text{ alors } (T_n f, g) = (f, T_n g).$$

Soit f un vecteur propre de T_n (on dit une *forme propre*). On a

$$T_n f = t_f(n) f.$$

En comparant les développements de Fourier des deux membres de l'équation précédente, on obtient pour tout m l'égalité

$$\sum_{\substack{d|(m,n)\\(d,N)=1}} d^{k-1} \hat{f}\left(\frac{mn}{d^2}\right) = t_f(n)\hat{f}(m)$$

puis en choisissant $m = 1$ on en déduit

$$\hat{f}(n) = t_f(n)\hat{f}(1).$$

Se posent alors deux questions :

(1) Sait-on si $\hat{f}(1) \neq 0$?

(2) Peut-on construire des formes propres de tous les opérateurs de Hecke ?

On va faire appel au résultat classique suivant.

Proposition 5.10. *Soit $\mathcal{E}$ un espace hermitien sur $\mathbb{C}$ et $\mathcal{F}$ une famille d'opérateurs normaux commutant deux à deux. Alors il existe une base orthonormale $\mathcal{B}$ de $\mathcal{E}$ formée de vecteurs propres de tous les opérateurs de $\mathcal{F}$.*

Corollaire 5.11. *Il existe une base orthogonale de l'espace $S_k(N)$ formée de vecteurs propres de tous les opérateurs de Hecke T_n tels que $(n, N) = 1$.*

Malheureusement, cela ne répond de façon satisfaisante ni à la première, ni à la seconde question. On peut trouver des formes f, vecteurs propres de tous les opérateurs de Hecke T_n, $(n, N) = 1$ telles que $\hat{f}(1) = 0$: soit L et M entiers tels que $L \neq 1$ et $LM = N$. Soit $f \in S_k(M)$ et

$$f_L : z \mapsto f(Lz) = \sum_n \hat{f}(n) e^{2i\pi Lnz} = \sum_{n \equiv 0 \pmod{L}} \hat{f}\left(\frac{n}{L}\right) e^{2i\pi nz}.$$

On vérifie

$$\begin{cases} f_L \in S_k(N), \\ T_n f = t_f(n) f \Rightarrow T_n f_L = t_f(n) f_L, \\ \hat{f}_L(1) = 0. \end{cases}$$

On définit alors l'espace des *formes anciennes* par

$$S_k^a(N) = \mathrm{Vect}\{f_L : f \in S_k(M),\ L \neq 1,\ LM \mid N,\ M \neq N\}.$$

On vérifie que cet espace est stable par les opérateurs de Hecke. On pose

$$S_k^n(N) = S_k^a(N)^{\perp}$$

et cet espace est appelé espace des *formes nouvelles*. On va voir qu'on peut répondre aux deux questions précédentes en se restreignant aux formes nouvelles.

Théorème 5.12 (Théorème d'Atkin & Lehner). *Soit $f \in S_k(N)$. On suppose que*

$$(n, N) = 1 \Longrightarrow \hat{f}(n) = 0.$$

Alors, $f \in S_k^a(N)$.

On en déduit du théorème d'Atkin & Lehner le théorème de multiplicité un.

Théorème 5.13 (Théorème de multiplicité un). *Soit $f \in S_k^n(N)$ avec $T_n f = t(n) f$ pour tous $(n, N) = 1$. Si $g \in S_k^n(N)$ et $T_n g = t(n) g$ pour tous $(n, N) = 1$ alors $g \in \mathbb{C} f$.*

On peut maintenant considérer une base orthogonale $(f_1, \ldots, f_d)$ de $S_k^n(N)$ formée de vecteurs propres de tous les opérateurs de Hecke T_n tels que $(n, N) = 1$ avec $\hat{f}_i(1) \neq 0$. On peut alors diviser chaque forme f_i par $\hat{f}_i(1)$ pour obtenir une base appelée *base de formes primitives* de $S_k(N)$. On note $\mathrm{H}_k^*(N)$ cette base dont on peut montrer qu'elle est unique.

Proposition 5.14. *Soit $f \in \mathrm{H}_k^*(N)$. Pour tout $n \in \mathbb{N}$, il existe $t_f(n)$ tel que*

$$T_n f = t_f(n) f.$$

On a donc pour $f \in \mathrm{H}_k^*(N)$

$$f(z) = \sum_{n=1}^{+\infty} t_f(n) e^{2i\pi nz}.$$

On déduit ensuite de la relation de multiplicativité des opérateurs de Hecke une relation de multiplicativité pour les coefficients de Fourier des formes primitives :

$$\hat{f}(m)\hat{f}(n) = \sum_{\substack{d|(m,n) \\ (d,N)=1}} d^{k-1} \hat{f}\left(\frac{mn}{d^2}\right).$$

L'existence de la base des formes primitives permet de développer une notion efficiente de fonction L en posant

$$L(f, s) = \sum_{n=1}^{+\infty} t_f(n) n^{-(k-1)/2} n^{-s}.$$

Un résultat de Deligne très difficile implique la majoration

$$|t_f(n)| \leq \#\{\, d \in \mathbb{N} \colon d \,|\, n \,\}\, n^{(k-1)/2}$$

de sorte que $L(f, s)$ converge pour $\mathrm{Re}\, s > 1$. On peut montrer qu'elle se prolonge en fonction entière. Ce prolongement, dû à Hecke, est une conséquence immédiate de la transformation de Mellin et de la modsularité. Ces fonctions admettent un développement en produit eulérien. On pose

$$\lambda_f(n) = t_f(n) n^{-(k-1)/2}.$$

À tout nombre premier p on associe le facteur $L(f_p, s)$ défini par

$$L(f_p, s) = \begin{cases} \left(1 - \lambda_f(p)p^{-s} + p^{-2s}\right)^{-1} & \text{si } (p, N) = 1, \\ \left(1 - \lambda_f(p)p^{-s}\right)^{-1} & \text{si } p \,|\, N. \end{cases}$$

On a

$$L(f, s) = \prod_{p \in \mathcal{P}} L(f_p, s).$$

Posant

$$L(f_\infty, s) = 2(2\pi)^{-(k-1)/2}(2\pi)^{-s}\Gamma\left(s + \frac{k-1}{2}\right)$$

on obtient l'équation fonctionnelle

$$L(f_\infty, s)L(f, s) = i^k \varepsilon_f(N) N^{-s+1/2} L(f_\infty, 1-s)L(f, 1-s) \tag{5.4}$$

où $\varepsilon_f(N)$ est un réel de norme 1. En particulier, si $i^k\varepsilon_f(N) = -1$, on a $L(f, \frac{1}{2}) = 0$. On définit donc les deux ensembles

$$\mathrm{H}_k^+(N) = \{f \in \mathrm{H}_k^*(N) \colon i^k \varepsilon_f(N) = 1\}$$

et

$$\mathrm{H}_k^-(N) = \{f \in \mathrm{H}_k^*(N) \colon i^k \varepsilon_f(N) = -1\}.$$

De même que pour la fonction ζ de Riemann, on conjecture que les points d'annulation de $L(f, s)$ à l'intérieur de la bande critique $\operatorname{Re} s \in [0, 1]$ sont situés sur l'axe critique $\operatorname{Re} s = \frac{1}{2}$. Cette conjecture s'appelle l'hypothèse de Riemann pour $L(f, s)$. Si $f \in \mathrm{H}_k^-(N)$, la suite des points d'annulation de $L(f, s)$ est

$$\left\{\frac{1}{2}\right\} \bigcup_{j>0} \left\{\frac{1}{2} + i\gamma_{f,j}, \frac{1}{2} - i\overline{\gamma_{f,j}}\right\}$$

avec $\operatorname{Re} \gamma_{f,j} \geq 0$ pour tout $j > 0$. Si $f \in \mathrm{H}_k^+(N)$, la suite des points d'annulation de $L(f, s)$ est

$$\bigcup_{j>0} \left\{\frac{1}{2} + i\gamma_{f,j}, \frac{1}{2} - i\overline{\gamma_{f,j}}\right\}$$

avec $\operatorname{Re} \gamma_{f,j} \geq 0$ pour tout $j > 0$. Dans les deux cas, il s'agit de multiensembles : les zéros sont répétés avec multiplicité.

On donne maintenant une généralisation des fonctions L de formes primitives, suivant la présentation de Cogdell & Michel [12]. Pour simplifier l'exposition, on se contente d'exposer la situation en poids 2. On note St la représentation standard de SU(2),

$$\begin{aligned} \mathrm{St} \colon \mathrm{SU}(2) &\longrightarrow \mathrm{GL}(\mathbb{C}^2) \\ M &\longmapsto \left| \begin{array}{ccc} \mathbb{C}^2 & \to & \mathbb{C}^2 \\ x & \mapsto & Mx. \end{array} \right. \end{aligned}$$

Si ρ est une représentation de dimension finie de SU(2), on pose

$$D(X, \rho, g) = \det[I - X\rho(g)]^{-1}$$

pour chaque $g \in \mathrm{SU}(2)$. Soit $f \in \mathrm{H}_2^*(N)$, avec N sans facteur carré. La majoration de Deligne implique que, pour chaque premier p ne divisant pas le niveau, il existe $\theta_{f,p} \in [0, \pi]$ tel que

$$\lambda_f(p) = \chi_{\mathrm{St}}[g(\theta_{f,p})]$$

où χ_{St} est le caractère de la représentation standard et

$$g(\theta) := \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

(autrement dit, $\lambda_f(p) = 2\cos\theta_{f,p}$). Pour tout entier $m \geq 0$, la fonction L de puissance symétrique m^{e} de f est

$$L(\mathrm{Sym}^m f, s) := \prod_{p \in \mathcal{P}} L_p(\mathrm{Sym}^m f, s) \tag{5.5}$$

avec

$$L_p(\mathrm{Sym}^m f, s) := D[p^{-s}, \mathrm{Sym}^m, g(\theta_{f,p})]$$

si p est premier à N et

$$L_p(\mathrm{Sym}^m f, s) := [1 - \lambda_f(p^m)p^{-s}]^{-1}$$

sinon. On a noté Sym^m la composée de la représentation puissance symétrique m^{e} de GL(2) avec la représentation standard SU(2). Ainsi,

$$\mathrm{Sym}^m(g(\theta_{f,p})) = \mathrm{Diag}\left(e^{i(m-2\ell)\theta}\right)_{0 \leq \ell \leq m}.$$

Si $m = 1$, on a $L(\mathrm{Sym}^1 f, s) = L(f, s)$. Il existe sur les fonctions L de puissances symétriques m^{e} une conjecture concernant l'équation fonctionnelle qu'on énonce maintenant (toujours dans le cas où f est de poids 2, voir [12, Hypothesis $\mathrm{Sym}^k(f)$]).

Conjecture 5. *Soit N un entier sans facteur carré et $f \in \mathrm{H}_2^*(N)$. Il existe une représentation unitaire irréductible automorphe cuspidale de $\mathrm{GL}_m(\mathbb{Q})$ dont les facteurs locaux sont les mêmes que ceux de $L(\mathrm{Sym}^m f, s)$. Posons*

$$L_\infty(\mathrm{Sym}^m f, s) := \begin{cases} \pi^{-s/2}\Gamma\left(\frac{s}{2}\right) 2^u \prod_{j=1}^u (2\pi)^{-s-j}\Gamma(s+j) \\ \qquad\qquad \text{si } m = 2u \text{ avec } u \text{ pair,} \\ \pi^{-(s+1)/2}\Gamma\left(\frac{s+1}{2}\right) 2^u \prod_{j=1}^u (2\pi)^{-s-j}\Gamma(s+j) \\ \qquad\qquad \text{si } m = 2u \text{ avec } u \text{ impair,} \\ 2^{u+1} \prod_{j=0}^u (2\pi)^{-s-j-1/2}\Gamma\left(s+j+\frac{1}{2}\right) \quad \text{si } m = 2u+1. \end{cases}$$

Alors, il existe $\varepsilon(\mathrm{Sym}^m f) \in \{-1, 1\}$ *tel que*

$$N^{ms/2} L_\infty(\mathrm{Sym}^m f, s) L(\mathrm{Sym}^m f, s) = \varepsilon(\mathrm{Sym}^m f) N^{m(1-s)/2} L_\infty(\mathrm{Sym}^m f, 1-s) L(\mathrm{Sym}^m f, 1-s).$$

Cette conjecture a été démontrée pour $m \in \{1, 2, 3, 4\}$ (voir [27], [55], [56], [54]).

Remarque 5.15. On montre [12, Proposition 3.6] que sous l'hypothèse que le conjecture précédente est vérifiée, le signe de l'équation fonctionnelle est

$$\varepsilon(\mathrm{Sym}^m f) = \varepsilon_\infty \mu(q) [\sqrt{q} \lambda_f(q)]^{m^2} = \pm 1$$

avec

$$\varepsilon_\infty = \begin{cases} -1 & \text{si } m \equiv 1 \text{ ou } 3 \pmod 8, \\ 1 & \text{sinon} \end{cases}$$

et $\mu(q)$ vaut -1 si q a un nombre impair de diviseurs premiers et 1 sinon.

5.5 Familles de fonctions L de formes modulaires et matrices aléatoires

5.5.1 Petits zéros. Le théorème suivant, dû à Iwaniec, Luo & Sarnak [47] compte les « petits zéros » des fonctions L de formes modulaires, c'est-à-dire ceux compris dans un intervalle autour de $\frac{1}{2}$ de longueur l'espacement moyen entre zéros. Dans le cas où $f \in \mathrm{H}_k^-(N)$, on ne compte pas le zéro évident $\frac{1}{2}$ (précisément, si $\frac{1}{2}$ annule $L(f, s)$ à l'ordre n, on compte $\frac{1}{2}$ seulement $n-1$ fois). Ce théorème implique que les « petits zéros » des fonctions L de formes de $\mathrm{H}_k^-(N)$ se comportent asymptotiquement (en N) comme les petits angles propres des matrices de $\mathrm{SO}(2N+1)$ (pour lesquelles 0 est angle propre évident) et que les « petits zéros » des fonctions L de formes de $\mathrm{H}_k^+(N)$ se comportent asymptotiquement (en N) comme les petits angles propres des matrices de $\mathrm{SO}(2N)$ (comparer avec le théorème 3.7).

Théorème 5.16. *En admettant « suffisamment » d'hypothèses de Riemann. Soit* ϕ *une fonction de localisation telle que* $\mathrm{Supp}\,\hat{\phi} \subset]-2, 2[$. *Lorsque* N *parcourt la suite des entiers sans facteurs carrés, on a*

$$\lim_{N\to+\infty} \frac{1}{\#\mathrm{H}_k^-(N)} \sum_{f\in\mathrm{H}_k^-(N)} \sum_{j\in\mathbb{Z}^*} \phi\left[\frac{\log(k^2N)}{2\pi}\gamma_{f,j}\right] = \int_{\mathbb{R}} \phi(t) D[\mathrm{SO}^-](t)\,\mathrm{d}t,$$

$$\lim_{N\to+\infty} \frac{1}{\#\mathrm{H}_k^+(N)} \sum_{f\in\mathrm{H}_k^+(N)} \sum_{j\in\mathbb{Z}^*} \phi\left[\frac{\log(k^2N)}{2\pi}\gamma_{f,j}\right] = \int_{\mathbb{R}} \phi(t) D[\mathrm{SO}^+](t)\,\mathrm{d}t.$$

Remarque 5.17. Le théorème 5.16 utilise les hypothèses de Riemann pour :

(1) les fonctions L de caractères de Dirichlet ;

(2) les fonctions L de formes primitives ;

(3) les fonctions L des carrés symétriques des formes primitives.

La troisième est essentiellement esthétique : elle permet d'énoncer les résultats avec la moyenne habituelle

$$\mathrm{Moy}(X) = \frac{1}{\#\,\mathrm{H}_k^*(N)} \sum_{f \in \mathrm{H}_k^*(N)} X(f)$$

plutôt qu'avec la moyenne harmonique

$$\mathrm{Moy}^{\mathrm{h}}(X) = \sum_{f \in \mathrm{H}_k^*(N)} \frac{(k-2)!}{(4\pi)^{k-1}(f,f)} X(f)$$

qui apparaît plus naturellement dans cette théorie. La deuxième hypothèse est technique, pour s'en passer, il suffirait de considérer des fonctions ϕ sur $\mathbb{C}$ plutôt que sur $\mathbb{R}$. En revanche, la première hypothèse est fondamentale : sans elle, on ne pourrait utiliser que des fonctions ϕ telles que $\mathrm{Supp}\,\hat{\phi} \subset]-1, 1[$. Or, pour distinguer les deux répartitions, il faut dépasser $]-1, 1[$ puisque

$$\int_{\mathbb{R}} \phi D[\mathrm{SO}^-] = \int_{\mathbb{R}} \hat{\phi} \widehat{D[\mathrm{SO}^-]} \quad \text{et} \quad \int_{\mathbb{R}} \phi D[\mathrm{SO}^+] = \int_{\mathbb{R}} \hat{\phi} \widehat{D[\mathrm{SO}^+]}$$

et

$$\widehat{D[\mathrm{SO}^-]}\big|_{]-1,1[} = \widehat{D[\mathrm{SO}^+]}\big|_{]-1,1[}.$$

On donne une conséquence de ce théorème, due à Iwaniec, Luo & Sarnak. On conjecture que l'ordre de l'annulation de $L(f, 1/2)$ (ou au moins sa parité) est gouverné par le signe de l'équation fonctionnelle de $L(f, s)$. Autrement dit, on conjecture qu'asymptotiquement, l'ordre d'annulation de $L(f, s)$ en $1/2$ est 0 si $f \in \mathrm{H}_k^+(N)$ et 1 si $f \in \mathrm{H}_k^-(N)$. Iwaniec, Luo & Sarnak déduisent du théorème 5.16 les deux inégalités suivantes :

$$\limsup_{N \to +\infty} \frac{1}{\#\,\mathrm{H}_k^+(N)} \#\left\{ f \in \mathrm{H}_k^+(N) \colon L\left(\frac{1}{2}, f\right) = 0 \right\} < 0,4323 \tag{5.6}$$

et

$$\limsup_{N \to +\infty} \frac{1}{\#\,\mathrm{H}_k^-(N)} \#\left\{ f \in \mathrm{H}_k^-(N) \colon L'\left(\frac{1}{2}, f\right) = 0 \right\} < 0,0573. \tag{5.7}$$

On donne leur preuve pour $\mathrm{H}_k^+(N)$ en notant

$$D[\phi](f) = \sum_{j \in \mathbb{Z}^*} \phi\left[\frac{\log(k^2 N)}{2\pi} \gamma_{f,j}\right].$$

Pour $m \geq 0$, on considère

$$P_m^+(N) = \frac{1}{\#\,\mathrm{H}_k^+(N)} \#\left\{ f \in \mathrm{H}_k^+(N) \colon \operatorname*{ord}_{s=1/2} L(s,f) = m \right\}.$$

Grâce à l'équation fonctionnelle, $P_m^+(N)$ est nul si m est impair. Ainsi,

$$\sum_{m \in 2\mathbb{N}} P_m^+(N) = 1. \tag{5.8}$$

D'autre part, pour tout $\varepsilon > 0$, il existe N_ε tel que si $N \geq N_\varepsilon$, alors

$$\frac{1}{\#\,\mathrm{H}_k^+(N)} \sum_{f \in \mathrm{H}_k^+(N)} D[\phi](f) < \int_{-\infty}^{+\infty} \phi(x) D[\mathrm{SO}^+](x)\,\mathrm{d}x + \varepsilon. \tag{5.9}$$

Si $f \in \mathrm{H}_k^+(N)$ et si $L(f,s)$ s'annule à l'ordre m en $1/2$, en séparant $1/2$ compté m fois, on a

$$D[\phi](f) = m\phi(0) + \sum_{\gamma_f \neq 0} \phi\left[\frac{\log(k^2 N)}{2\pi} \gamma_f \right].$$

Le choix de ϕ positive telle que $\phi(0) = 1$ conduit à

$$D[\phi](f) \geq m.$$

Mais,

$$\begin{aligned} \frac{1}{\#\,\mathrm{H}_k^+(N)} \sum_{f \in \mathrm{H}_k^+(N)} D[\phi](f) &= \sum_{m \in 2\mathbb{N}} \frac{1}{\#\,\mathrm{H}_k^+(N)} \sum_{\substack{f \in \mathrm{H}_k^+(N) \\ \mathrm{ord}_{s=1/2} L(f,s) = m}} D[\phi](f) \\ &\geq \sum_{m \in 2\mathbb{N}} m P_m^+(N). \end{aligned} \tag{5.10}$$

De (5.9) et (5.10), on déduit

$$\sum_{m \in 2\mathbb{N}} m P_m^+(N) < \int_{-\infty}^{+\infty} \phi(x) D[\mathrm{SO}^+](x)\,\mathrm{d}x + \varepsilon. \tag{5.11}$$

En retirant (5.8) de (5.11), on obtient

$$2 \sum_{m \in 2\mathbb{N}} P_m^+ - \sum_{m \in 2\mathbb{N}} m P_m^+ > 2 - \int_{-\infty}^{+\infty} \phi(x) D[\mathrm{SO}^+](x)\,\mathrm{d}x - \varepsilon.$$

Mais,

$$\begin{aligned} 2 \sum_{m \in 2\mathbb{N}} P_m^+ - \sum_{m \in 2\mathbb{N}} m P_m^+ &= 2P_0^+(N) + 2 \sum_{m > 0} (1-m) P_{2m}^+(N) \\ &\leq 2P_0^+(N) \end{aligned}$$

de sorte que

$$P_0^+(N) > 1 - \frac{1}{2}\int_{-\infty}^{+\infty} \phi(x) D[\mathrm{SO}^+](x)\,\mathrm{d}x - \frac{\varepsilon}{2}.$$

En minimisant

$$\int_{-\infty}^{+\infty} \phi(x) D[\mathrm{SO}^+](x)\,\mathrm{d}x$$

sur les fonctions ϕ positives, paires vérifiant $\phi(0) = 1$ et $\mathrm{Supp}\,\hat{\phi} \subset]-2, 2[$, on obtient

$$\liminf_{N\to+\infty} \frac{1}{\#\,\mathrm{H}_k^+(N)} \#\left\{ f \in \mathrm{H}_k^+(N) \colon L\left(\frac{1}{2}, f\right) \neq 0 \right\} \geq \alpha^+$$

avec

$$\alpha^+ = \frac{13 - \mathrm{cotan}(1/4)}{16} = 1 - 0.43226\ldots.$$

Les meilleurs résultats inconditionnels sont obtenus par Iwaniec & Sarnak pour (5.6) [48]. Ils obtiennent $1/2 + \varepsilon$ comme majorant. Pour (5.7), les meilleurs résultats inconditionnels sont essentiellement obtenus par Kowalski & Michel [58]. Ils obtiennent $1/8 + \varepsilon$ [46, Theorem 26.2].

5.5.2 Quantième zéro. Le théorème d'Iwaniec, Luo & Sarnak, compte-tenu de la modélisation des fonctions L par les polynômes caractéristiques permet de conjecturer le comportement du quantième zéro des fonctions de formes modulaires. Avec les notations du théorème 3.6, on peut faire la conjecture suivante.

Conjecture 6. *Soit $j \geq 1$, un entier et $k \geq 2$ un entier pair. Soit h une fonction continue à support compact. Lorsque N parcourt les entiers sans facteur carré,*

$$\lim_{N\to+\infty} \frac{1}{\#\,\mathrm{H}_k^+(N)} \sum_{f\in \mathrm{H}_k^+(N)} h\left(\frac{\log(k^2N)}{2\pi}\gamma_{f,j}\right) = \int_{\mathbb{R}} h(x) v_{k,+}(x)\,\mathrm{d}x$$

et

$$\lim_{N\to+\infty} \frac{1}{\#\,\mathrm{H}_k^-(N)} \sum_{f\in \mathrm{H}_k^-(N)} h\left(\frac{\log(k^2N)}{2\pi}\gamma_{f,j}\right) = \int_{\mathbb{R}} h(x) v_{k,-}(x)\,\mathrm{d}x.$$

On rappelle que les fonctions $v_{k,\pm}$ ont été définies au théorème 3.6.

Katz remarque [50, Introduction] que cette conjecture permet de montrer que

$$\lim_{N\to+\infty} \frac{1}{\#\,\mathrm{H}_k^+(N)} \#\left\{ f \in \mathrm{H}_k^+(N) \colon L\left(\frac{1}{2}, f\right) = 0 \right\} = 0 \tag{5.12}$$

et

$$\lim_{N\to+\infty} \frac{1}{\#\,\mathrm{H}_k^-(N)} \#\left\{ f \in \mathrm{H}_k^-(N) \colon L'\left(\frac{1}{2}, f\right) = 0 \right\} = 0 \tag{5.13}$$

(voir page 213). On montre (5.12), l'équation (5.13) se montrant de façon semblable. Pour tout $t \in]0,1[$, on construit une fonction $h_t : \mathbb{R} \to [0,1]$ dont le support est inclus dans $[-t,t]$ et telle que $h_t(0) = 1$. Puisque h_t est supérieure à la fonction de Dirac en 0, on a

$$\frac{1}{\#\mathrm{H}_k^+(N)} \sum_{f \in \mathrm{H}_k^+(N)} h_t\left[\frac{\log(k^2N)}{2\pi}\gamma_{f,1}\right] \geq \frac{1}{\#\mathrm{H}_k^+(N)} \#\left\{f \in \mathrm{H}_k^+(N) : L\left(\frac{1}{2}, f\right) = 0\right\}.$$

Lorsque N puis t tendent vers $+\infty$, on obtient

$$\lim_{N\to+\infty} \frac{1}{\#\mathrm{H}_k^+(N)} \#\left\{f \in \mathrm{H}_k^+(N) : L\left(\frac{1}{2}, f\right) = 0\right\} = 0.$$

L'intérêt des équations (5.12) et (5.13) dépasse le cadre de la théorie analytique des nombres grâce à la conjecture de Birch & Swinnerton-Dyer [80]. Soit E une courbe elliptique sur $\mathbb{Q}$: c'est une courbe algébrique sur $\mathbb{Q}$, de genre 1 et munie d'un point O. On sait munir une telle courbe d'une loi de groupe ayant O pour élément neutre [57]. D'après un théorème de Mordell, le groupe $E(\mathbb{Q})$ des points de E à coordonnées rationelles est de type fini. Il est donc isomorphe au produit de son sous-groupe de torsion (qui est bien connu grâce a un théorème de Mazur) et de r copies de $\mathbb{Z}$. L'entier r s'appelle le *rang* de E. Grâce aux travaux de Breuil, Conrad, Diamond & Taylor, de Taylor & Wiles et de Wiles, on sait associer (de façon assez explicite) à E une forme modulaire f de poids 2. La conjecture de Birch & Swinnerton-Dyer (dans sa version faible) prévoit l'égalité de r avec l'ordre d'annulation de $L(f,s)$ en $1/2$.

5.5.3 Moments des valeurs centrales. Le théorème 5.16, compte-tenu de la modélisation des fonctions L par les polynômes caractéristiques permet de conjecturer les moments des valeurs centrales de $L(f,s)$. On peut faire la conjecture suivante.

Conjecture 7. *Soit $k \geq 2$ un entier pair et $\ell > 0$ un entier. Alors,*

$$\lim_{N\to+\infty} \frac{1}{\#\mathrm{H}_k^+(N)} \sum_{f \in \mathrm{H}_k^+(N)} L\left(f, \frac{1}{2}\right)^\ell = a_{+,\ell} g_{\mathrm{SO}^+}(\ell) \left[\frac{\log(k^2N)}{2\pi}\right]^{\ell(\ell-1)/2}$$

$$\lim_{N\to+\infty} \frac{1}{\#\mathrm{H}_k^-(N)} \sum_{f \in \mathrm{H}_k^-(N)} L'\left(f, \frac{1}{2}\right)^\ell = a_{-,\ell} g_{\mathrm{SO}^-}(\ell) \left[\frac{\log(k^2N)}{2\pi}\right]^{\ell(\ell+1)/2}.$$

La fonction g_{SO^+} a été définie au théorème 3.9 et on a

$$g_{\mathrm{SO}^-}(\ell) = \frac{1}{2} g_{\mathrm{SO}^+}(\ell+1).$$

Une formulation générale de la valeurs des coefficients arithmétiques $a_{\pm,\ell}$ est assez délicate. On renvoie à [14, Conjecture 4.5.2] pour une expression générale (et une conjecture plus générale).

Fixons k pair. On déduit de la conjecture des moments, que pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ telle que

$$\left|L\left(f, \frac{1}{2}\right)\right| \leq C_\varepsilon N^\varepsilon \tag{5.14}$$

pour tout N et toute $f \in \mathrm{H}_k^*(N)$. Cette inégalité est la conjecture de Lindelöf au point central pour les fonctions L de formes primitives, relativement au niveau. Notons aussi que, la conjecture des moments implique l'existence, pour tout $n > 0$ de $c_n > 0$ tel que

$$\max_{f \in \mathrm{H}_k^+(N)} L\left(f, \frac{1}{2}\right) \geq c_n [\log(k^2 N)]^n$$

pour tout N.

Une majoration inconditionelle de $L(f, 1/2)$ est atteignable par le théorème de Phrägmèn-Lindelöf (dans une version due à Rademacher [67, Theorem 2]) grâce à l'équation fonctionnelle (5.4). On obtient alors

$$L\left(f, \frac{1}{2}\right) \ll_\varepsilon N^{1/4+\varepsilon}. \tag{5.15}$$

Obtenir une borne inconditionnelle améliorant l'exposant $1/4$ de la majoration (5.15) s'appelle *briser la convexité*. C'est un vaste et délicat sujet [62, Lecture 4]. Dans le cadre que nous étudions, le meilleur résultat actuel est dû à Duke, Friedlander & Iwaniec [20, Theorem 3]

$$L\left(f, \frac{1}{2}\right) \ll_\varepsilon N^{1/4-1/192+\varepsilon}.$$

Les conséquences de bris de convexité dépassent le simple record d'amélioration d'exposant. Pour un tour d'horizon des différentes conséquences (par exemple en *Chaos arithmétique*) on renvoie à [62, Lecture 5].

6 Détermination du type de symétrie

L'objet de cette partie est d'éclairer, par la théorie des représentations, l'origine des groupes de symétries trouvés précédemment. Cet éclairage est dû à Cogdell & Michel [12].

Soit G un groupe fini ou compact dont on note id la représentation identité (triviale) de dimension 1. Soit ρ une représentation irréductible de dimension finie de G. Pour $|X| < 1$, on pose

$$D(X, \rho, g) = \det(I - X\rho(g))^{-1}$$

et, pour tout $z \in \mathbb{C}$, on développe

$$D(X, \rho, g)^z = \sum_{\alpha=0}^{+\infty} \lambda_\rho^{z,\alpha}(g) X^\alpha.$$

Si $\mathrm{d}g$ est la mesure de Haar sur G, on définit

$$D^z(X,\rho)=\int_G D(X,\rho,g)^z\,\mathrm{d}g=\sum_{\alpha=0}^{+\infty}\lambda_\rho^{z,\alpha}X^\alpha$$

où

$$\lambda_\rho^{z,\alpha}=\int_G\lambda_\rho^{z,\alpha}(g)\,\mathrm{d}g.$$

De

$$D(X,\rho,g)^z=\exp\Big[z\log\Big(\sum_{\alpha=0}^{+\infty}X^\alpha\operatorname{tr}\operatorname{Sym}^\alpha\rho(g)\Big)\Big]$$

on déduit

$$\lambda_\rho^{z,0}(g)=1,\quad \lambda_\rho^{z,1}(g)=z\operatorname{tr}\rho(g)$$

et

$$\lambda_\rho^{z,2}(g)=z\operatorname{tr}\operatorname{Sym}^2\rho(g)+\frac{z(z-1)}{2}(\operatorname{tr}\rho(g))^2.$$

Ainsi,

$$\lambda_\rho^{z,0}=1,\quad \lambda_\rho^{z,1}=0$$

et

$$\begin{aligned}\lambda_\rho^{z,2}&=z\int_G\operatorname{tr}\operatorname{Sym}^2\rho(g)\,\mathrm{d}g+\frac{z(z-1)}{2}\int_G(\operatorname{tr}\rho(g))^2\,\mathrm{d}g\\&=\frac{z^2}{2}\int_G(\operatorname{tr}\rho(g))^2\,\mathrm{d}g+\frac{z}{2}\int_G\operatorname{tr}\rho(g^2)\,\mathrm{d}g.\end{aligned}$$

On note FrSc l'indicateur de Frobenius-Schur défini par

$$\operatorname{FrSc}(\rho)=\int_G\operatorname{tr}\operatorname{Sym}^2\rho(g)-\operatorname{tr}\wedge^2\rho(g)\,\mathrm{d}g.$$

Cet indicateur ne prend que trois valeurs selon le choix nécessaire suivant :

$$\operatorname{FrSc}(\rho)=\begin{cases}0 & \text{si } \rho \text{ n'est pas autoduale,}\\ 1 & \text{si id apparaît une fois dans la décomposition,}\\ & \quad\text{en représentations irréductibles de } \operatorname{Sym}^2\rho,\\ -1 & \text{si id apparaît une fois dans la décomposition,}\\ & \quad\text{en représentations irréductibles de } \wedge^2\rho.\end{cases}$$

On a

$$\operatorname{FrSc}(\rho)=\int_G\operatorname{tr}\rho(g^2)\,\mathrm{d}g\quad\text{et}\quad\operatorname{FrSc}(\rho)^2=\int_G[\operatorname{tr}\rho(g)]^2\,\mathrm{d}g.$$

Cogdell & Michel démontrent [12, équation (2.30)]

$$D^z(X,\rho)=1+\lambda_\rho^{z,2}X^2+O_{|z|}(X^3)\tag{6.1}$$

uniformément pour $|X| < 1$.

Pour $r \in \mathbb{R}^+$, on considère alors

$$L_X^r\left(\frac{1}{2}, \rho\right) = \prod_{p \le X} D^r(p^{-1/2}, \rho).$$

Grâce à (6.1), on a

$$\begin{aligned}\log L_X^r\left(\frac{1}{2}, \rho\right) &= \sum_{p \le X} \frac{\lambda_\rho^{r,2}}{p} + O\left(\frac{1}{p^{3/2}}\right) \\ &= \lambda_\rho^{r,2} \log\log X + c_1 + O\left(\frac{1}{\log X}\right)\end{aligned}$$

pour une constante $c_1 \in \mathbb{R}^+$. Ainsi,

$$L_X^r\left(\frac{1}{2}, \rho\right) = e^{c_1 + O(1/\log X)} (\log X)^{\lambda_\rho^{r,2}} \tag{6.2}$$

et la puissance du logarithme ne peut prendre que l'une des trois valeurs

$$0, \ \frac{r(r-1)}{2} \text{ et } \frac{r(r+1)}{2}.$$

Cogdell & Michel démontrent [12, Théorème 1.5] que

$$\begin{aligned}\lim_{q \to +\infty} \frac{1}{\# \mathrm{H}_2^*(q)} \sum_{f \in \mathrm{H}_2^*(q)} L(\mathrm{Sym}^k f, 1)^r &= \prod_{p \in \mathcal{P}} \int_{\mathrm{SU}(2)} D(p^{-1}, \mathrm{Sym}^k, g)^r \, \mathrm{d}g \\ &= \prod_{p \in \mathcal{P}} D^r(p^{-1}, \mathrm{Sym}^k)\end{aligned}$$

pour tout $r \in \mathbb{R}^+$ (en fait $r \in \mathbb{C}$) dès que les hypothèses naturelles d'automorphie et d'absence de zéro de Siegel sont vérifiées. Il semble dès lors raisonnable d'espérer une équivalence asymptotique (lorsque q tend vers $+\infty$) entre

$$\frac{1}{\# \mathrm{H}_2^*(q)} \sum_{f \in \mathrm{H}_2^*(q)} L\left(\frac{1}{2}, \mathrm{Sym}^k f\right)^r$$

et

$$\prod_{p \le c(r) q^k} D^r(p^{-1/2}, \mathrm{Sym}^k) = L_{c(r)q^k}^r\left(\frac{1}{2}, \mathrm{Sym}^k\right)$$

pour une certaine constante (dépendant de r) $c(r) > 0$. Grâce à (6.2), on a

$$L_{c(r)q^k}^r\left(\frac{1}{2}, \mathrm{Sym}^k\right) \sim C(r) (\log q^k)^{Q_k(r)}$$

avec

$$Q_k(r) = \begin{cases} \frac{r(r+1)}{2} & \text{si } k \text{ est impair,} \\ \frac{r(r-1)}{2} & \text{sinon.} \end{cases}$$

Il n'y a donc que deux types de comportements possibles dans ce cas, correspondant aux deux types de symétrie trouvées précédemment.

Pour la famille

$$\bigcup_{q\in\mathcal{P}\setminus\{2\}} \bigcup_{\chi\in\widehat{(\mathbb{Z}/q\mathbb{Z})^\times}} \{L(\chi,s)\}$$

on a $G=(\mathbb{Z}/q\mathbb{Z})^\times$ et la représentation associée est la représentation standard puisque pour tout χ, il existe $a_\chi\in G$ tel que

$$\chi(g^y)=\exp\left(\frac{2i\pi y a_\chi}{\varphi(q)}\right)$$

pour tout y, g étant un générateur de G. L'indicateur de Frobenius-Schur est alors

$$\frac{1}{\varphi(q)}\sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times}\exp\left(\frac{2i\pi}{\varphi(q)}a\right)=0.$$

Finalement, le fait que l'indicateur de Frobenius-Schur ne prenne que trois valeurs rejoint, et explique, le fait qu'on ne trouve que trois types de comportements pour les moments de valeurs centrales de fonctions L.

Remerciements. L'auteur remercie chaleureusement Emmanuel Kowalski, Louise Nyssen, Guillaume Ricotta et Jie Wu pour leurs commentaires.

Références

[1] V. S. Adamchik, Symbolic and numeric computations of the Barnes function, *Comput. Phys. Comm.* 157 (2004), no. 3, 181–190.

[2] R. C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes. II, *Proc. London Math. Soc.* (3) 83 (2001), no. 3, 532–562.

[3] E. W. Barnes, The Theory of the G-function, *Quart. J. Pure Appl. Mathematics* 31 (1899), 264–314.

[4] E. L. Basor and P. J. Forrester, Formulas for the evaluation of Toeplitz determinants with rational generating functions, *Math. Nachr.* 170 (1994), 5–18.

[5] E. B. Bogomolny and J. P. Keating, Random matrix theory and the Riemann zeros. I. Three- and four-point correlations, *Nonlinearity* 8 (1995), no. 6, 1115–1131.

[6] ——, Random matrix theory and the Riemann zeros. II. n-point correlations, *Nonlinearity* 9 (1996), no. 4, 911–935.

[7] E. Bombieri, Problems of the Millennium: The Riemann Hypothesis, Clay Mathematical Institute, 2000, `http://www.claymath.org/millennium/Riemann_Hypothesis/Official_Problem_Description.pdf`.

[8] J.-B. Bost, Le théorème des nombres premiers et la transformation de Fourier, *La fonction zêta*, pp. 1–35, Ed. Éc. Polytech., Palaiseau, 2003.

[9] E. Brézin and S. Hikami, Characteristic polynomials of random matrices, *Comm. Math. Phys.* 214 (2000), no. 1, 111–135.

[10] A. Y. Cheer and D. A. Goldston, Simple zeros of the Riemann zeta-function, *Proc. Amer. Math. Soc.* 118 (1993), no. 2, 365–372.

[11] J. W. Cogdell, L-functions and Converse Theorems for GL(n), *Automorphic Forms and Applications* (Park City, UT, 2002), IAS/Park City Math. Ser. 12, Amer. Math. Soc., Providence, RI, à paraître.

[12] J. Cogdell and P. Michel, On the complex moments of symmetric power L-functions at $s = 1$, *Internat. Math. Res. Notices* 2004 (2004), no. 31, 1561–1617.

[13] J. B. Conrey, More than two fifths of the zeros of the Riemann zeta function are on the critical line, *J. Reine Angew. Math.* 399 (1989), 1–26.

[14] J. B Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, Integral moments of L-functions, Technical Report 2002-6, American Institute of Mathematics, `arXiv:math.NT/0206018`.

[15] ——, Autocorrelation of random matrix polynomials, *Comm. Math. Phys.* 237 (2003), no. 3, 365–395.

[16] J. B. Conrey and A. Ghosh, A conjecture for the sixth power moment of the Riemann zeta-function, *Internat. Math. Res. Notices* 1998 (1998), no. 15, 775–780.

[17] J. B. Conrey, A. Ghosh, and S. M. Gonek, Simple zeros of the Riemann zeta-function, *Proc. London Math. Soc.* (3) 76 (1998), no. 3, 497–522.

[18] J. B. Conrey and S. M. Gonek, High moments of the Riemann zeta-function, *Duke Math. J.* 107 (2001), no. 3, 577–604.

[19] J. Dufresnoy and Ch. Pisot, Sur la relation fonctionnelle $f(x+1) - f(x) = \varphi(x)$, *Bull. Soc. Math. Belg.* 15 (1963), 259–270.

[20] W. Duke, J. B. Friedlander, and H. Iwaniec, Bounds for automorphic L-functions. II, *Invent. Math.* 115 (1994), no. 2, 219–239.

[21] ——, Erratum: "Bounds for automorphic L-functions. II" [Invent. Math. **115** (1994), no. 2, 219–239], *Invent. Math.* 140 (2000), no. 1, 227–242.

[22] F. J. Dyson, Statistical theory of the energy levels of complex systems. I, *J. Math. Phys.* 3 (1962), 140–156.

[23] ——, Statistical theory of the energy levels of complex systems. II, *J. Math. Phys.* 3 (1962), 157–165.

[24] ——, Statistical theory of the energy levels of complex systems. III, *J. Math. Phys.* 3 (1962), 166–175.

[25] P. X. Gallagher, Pair correlation of zeros of the zeta function, *J. Reine Angew. Math.* 362 (1985), 72–86.

[26] M. Gaudin, Sur la loi limite de l'espacement des valeurs propres d'une matrice aléatoire, *Nuclear Phys.* 25 (1961), 447–458.

[27] S. Gelbart and H. Jacquet, A relation between automorphic representations of GL(2) and GL(3), *Ann. Sci. École Norm. Sup.* (4) 11 (1978), no. 4, 471–542.

[28] J. W. L. Glaisher, On the Constant Which Occurs in the Formula for $1^1 2^2 3^3 \cdots n^n$, *Messenger of Mathematics* 24 (1894), 1–16.

[29] D. Goldfeld, J. Hoffstein, and D. Lieman, An effective zero-free region, *Ann. of Math.* (2) 140 (1994), no. 1, 177–181.

[30] D. A. Goldston, On the function $S(T)$ in the theory of the Riemann zeta-function, *J. Number Theory* 27 (1987), no. 2, 149–177.

[31] ——, Notes on pair correlation of zeros and prime numbers, *Recent perspectives in random matrix theory and number theory*, London Math. Soc. Lecture Note Ser. 322, Cambridge University Press, Cambridge 2005, pp. 79–110.

[32] D. A. Goldston and S. M. Gonek, Mean value theorems for long Dirichlet polynomials and tails of Dirichlet series, *Acta Arith.* 84 (1998), no. 2, 155–192.

[33] D. A. Goldston and H. L. Montgomery, Pair correlation of zeros and primes in short intervals, *Analytic number theory and Diophantine problems* (Stillwater, OK, 1984), Progr. Math. 70, Birkhäuser Verlag, Boston, MA 1987, pp. 183–203.

[34] Xavier Gourdon, *The* 10^{13} *first zeros of the Riemann Zeta function and zeros computation at very large height* October 24th 2004, `http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeros1e13-1e24.pdf`.

[35] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 6th ed., Academic Press Inc., San Diego, CA 2000, translated from the Russian; translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger.

[36] D. R. Heath-Brown and D. A. Goldston, A note on the differences between consecutive primes, *Math. Ann.* 266 (1984), no. 3, 317–320.

[37] J. Hoffstein and P. Lockhart, Coefficients of Maass forms and the Siegel zero, *Ann. of Math.* (2) 140 (1994), no. 1, 161–181 (with an appendix by Dorian Goldfeld, Hoffstein and Daniel Lieman).

[38] J. Hoffstein and D. Ramakrishnan, Siegel zeros and cusp forms, *Internat. Math. Res. Notices* 1995 (1995), no. 6, 279–308.

[39] G. Hoheisel, Primzahlprobleme in der Analysis, *Sitzungsber. Preuss. Akad. Wiss. Berlin* 1930 (1930), 580–588.

[40] C. P. Hughes, Random matrix theory and discrete moments of the Riemann zeta function, *J. Phys. A* 36 (2003), no. 12, 2907–2917 Random matrix theory.

[41] C. P. Hughes and Z. Rudnick, Linear statistics of low-lying zeros of L-functions, *Quart. J. Math. Oxford* 54 (2003), no. 3, 309–333.

[42] ——, Mock-Gaussian behaviour for linear statistics of classical compact groups, *J. Phys. A* 36 (2003), no. 12, 2919–2932.

[43] A. Ivić, *The Riemann zeta-function*, Wiley-Intersci. Publ., John Wiley & Sons Inc., New York 1985.

[44] ——, On some results concerning the Riemann hypothesis, *Analytic number theory* (Kyoto, 1996), London Math. Soc. Lecture Note Ser. 247, Cambridge University Press, Cambridge 1997, pp. 139–167.

[45] ——, On some reasons for doubting the Riemann hypothesis, 2003, `arXiv:math.NT/0311162`.

[46] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, RI 2004.

[47] H. Iwaniec, W. Luo, and P. Sarnak, Low lying zeros of families of L-functions, *Inst. Hautes Études Sci. Publ. Math.* (2000), no. 91, 55–131 (2001).

[48] H. Iwaniec and P. Sarnak, The non-vanishing of central values of automorphic L-functions and Landau-Siegel zeros, *Israel J. Math.* 120 (2000), 155–177.

[49] D. Joyner, *Distribution theorems of L-functions*, Pitman Res. Notes Math. Ser. 142, Longman Scientific & Technical, Harlow 1986.

[50] N. M. Katz, *Twisted L-functions and monodromy*, Ann. of Math. Stud. 150, Princeton University Press, Princeton, NJ 2002.

[51] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, Amer. Math. Soc. Colloq. Publ. 45, Amer. Math. Soc., Providence, RI 1999.

[52] J. P. Keating and N. C. Snaith, Random matrix theory and $\zeta(1/2 + it)$, *Comm. Math. Phys.* 214 (2000), no. 1, 57–89.

[53] ——, Random matrix theory and L-functions at $s = 1/2$, *Comm. Math. Phys.* 214 (2000), no. 1, 91–110.

[54] H. H. Kim, Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2, *J. Amer. Math. Soc.* 16 (2003), no. 1, 139–183 (with appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak).

[55] H. H. Kim and F. Shahidi, Cuspidality of symmetric powers with applications, *Duke Math. J.* 112 (2002), no. 1, 177–197.

[56] ——, Functorial products for $\mathrm{GL}_2 \times \mathrm{GL}_3$ and the symmetric cube for GL_2, *Ann. of Math.* (2) 155 (2002), no. 3, 837–893 (with an appendix by Colin J. Bushnell and Guy Henniart).

[57] Anthony W. Knapp, *Elliptic curves*, Math. Notes 40, Princeton University Press, Princeton, NJ 1992.

[58] E. Kowalski and P. Michel, A lower bound for the rank of $J_0(q)$, *Acta Arith.* 94 (2000), no. 4, 303–343.

[59] H.-Q. Liu and J. Wu, Numbers with a large prime factor, *Acta Arith.* 89 (1999), no. 2, 163–187.

[60] F. Martin and Emmanuel Royer, Formes modulaires et périodes, *Formes modulaires et transcendance* (*Colloque Jeunes*) (Marseille–Luminy, 2003), Sémin. Congr. 12, Soc. Math. France, Paris 2005, pp. 1–117.

[61] M. L. Mehta, Power series for level spacing functions of random matrix ensembles, *Z. Phys. B* 86 (1992), no. 2, 285–290.

[62] P. Michel, Analytic Number Theory and Families of L-functions, *Automorphic Forms and Applications* (Park City, UT, 2002), IAS/Park City Math. Ser. 12, Amer. Math. Soc., Providence, RI, à paraître.

[63] T. Miyake, *Modular forms*, Springer-Verlag, Berlin 1989.

[64] H. L. Montgomery, The pair correlation of zeros of the zeta function, *Analytic number theory* (Proc. Sympos. Pure Math., vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), Amer. Math. Soc., Providence, R.I., 1973, pp. 181–193.

[65] ——, Distribution of the zeros of the Riemann zeta function, *Proc. Internat. Congr. Mathematicians* (Vancouver, B. C., 1974), vol. 1, Canad. Math. Congress, Montreal, Que. 1975, pp. 379–381.

[66] A. Odlyzko, *Zeros number* $10^{22} + 1$ *through* $10^{22} + 10^4$ *of the Riemann zeta function*, `http://www.dtc.umn.edu/~odlyzko/zeta_tables/zeros5`.

[67] H. Rademacher, On the Phragmén-Lindelöf theorem and some applications, *Math. Z.* 72 (1959/1960), 192–204.

[68] K. Ramachandra, A note on numbers with a large prime factor, *J. London Math. Soc.* (2) 1 (1969), 303–306.

[69] D. Ramakrishnan and S. Wang, On the exceptional zeros of Rankin-Selberg L-functions, *Compositio Math.* 135 (2003), no. 2, 211–244.

[70] O. Ramaré and Y. Saouter, Short effective intervals containing primes, *J. Number Theory* 98 (2003), no. 1, 10–33.

[71] Z. Rudnick and P. Sarnak, Zeros of principal L-functions and random matrix theory, *Duke Math. J.* 81 (1996), no. 2, 269–322.

[72] M. Rubinstein, *Evidence for a Spectral Interpretation of the Zeros of L-Functions*, PhD Thesis, Princeton University, June 1998, `http://www.math.uwaterloo.ca/~mrubinst/thesis/thesis.html`.

[73] P. Sarnak, Problems of the Millennium: The Riemann Hypothesis (2004), Clay Mathematical Institute, 2004, `http://www.claymath.org/millennium/Riemann_Hypothesis/Sarnak_RH.pdf`.

[74] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, 2nd ed., Cours Spécialisés 1, Société Mathématique de France, Paris 1995.

[75] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., The Clarendon Press Oxford University Press, New York 1986, edited and with a preface by D. R. Heath-Brown.

[76] I. Vardi, Determinants of Laplacians and multiple gamma functions, *SIAM J. Math. Anal.* 19 (1988), no. 2, 493–507.

[77] M.-F. Vignéras, L'équation fonctionnelle de la fonction zêta de Selberg du groupe modulaire PSL(2, $\mathbf{Z}$), *Journées Arithmétiques de Luminy* (Colloq. Internat. CNRS, Centre Univ. Luminy, Luminy 1978), Astérisque 61, Soc. Math. France, Paris 1979, pp. 235–249.

[78] A. Voros, Spectral functions, special functions and the Selberg zeta function, *Comm. Math. Phys.* 110 (1987), no. 3, 439–465.

[79] H. Weyl, *The classical groups*, Princeton Landmarks Math., Princeton University Press, Princeton, NJ, 1997.

[80] A. Wiles, Problems of the Millennium: The Birch and Swinnerton-Dyer Conjecture, Clay Mathematical Institute, 2000, `http://www.claymath.org/millennium/Birch_and_Swinnerton-Dyer_Conjecture/BSD.pdf`.

Some recent applications of Kloostermania

Philippe Michel

Institut de Mathématique et de Modélisation de Montpellier
UMR CNRS 5149, CC051
Univ. Montpellier II, Place E. Bataillon,
FR-34095 Montpellier cdx, france
email:michel@math.univ-montp2.fr

1 Introduction

Given $a, b, c \geqslant 1$ such that $(ab, c) = 1$, Kloosterman sums are a special kind of algebraic exponential sums given by the following expression

$$\mathrm{Kl}(a, b; c) = \sum_{\substack{x(c)\\(x,c)=1}} e\Big(\frac{ax + b\bar{x}}{c}\Big), \quad x\bar{x} \equiv 1(c),\ e(*) = \exp(2\pi i *).$$

These sums were introduced by Kloosterman in 1926 [25] on the occasion of the so-called *Kloosterman refinement* in the circle method to give an asymptotic expression of the number of representations, $r_{abcd}(n)$, of a large integer n by a diagonal[1] quaternary definite quadratic form

$$ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2 = n.$$

He also provided a non-trivial bound for these sums, which enabled him to show that the Hasse principle holds for such quadratic forms. This was the first indication that Kloosterman sums have to do with modular forms since $r_{abcd}(n)$ is the n-th Fourier coefficient of a theta series of weight two. A more direct connection between Kloosterman sums and modular forms came with Petersson's formula which expresses Fourier coefficients of modular forms in terms of Kloosterman sums.

The word *Kloostermania* of the title was invented by M. Huxley in the 1980s to highlight a series of striking developments that took place at that time in analytic number theory and which were based on the modular nature of Kloosterman sums. The starting point was Kuznetsov's extension of Petersson's formula to the full automorphic spectrum for the modular group $\mathrm{SL}_2(\mathbb{Z})$. His formula enabled him to use spectral theory of the modular surface to solve essentially Linnik's conjecture on the existence of cancellation of sums of Kloosterman sums over the modulus [26]. But Kloostermania

[1] Later Kloosterman extended this result to general definite quaternary quadratic forms.

really lifted-off with the landmark paper of Deshouillers/Iwaniec [7] who generalized Kuznetsov's formula to arbitrary congruence subgroups. These formulae were then used to derive bounds of sums of exponential sums and provided a powerful new tool in analytic number theory. Amongst the many applications that followed from them, arguably the most striking are the works of Bombieri/Frieldander/Iwaniec and Fouvry who obtained improvements over the Bombieri/Vinogradov theorem which go far beyond the possibilities of the Generalized Riemann Hypothesis [1], [2], [16].

Besides their modular nature, Kloosterman sums, as a special type of algebraic exponential sums, also enjoy rather deep algebraic/geometric properties (the simplest being Weil's bound), which follows from the work of Deligne and Katz in ℓ-adic cohomology. In this paper, we review both of these aspects of Kloosterman sums and present several recent applications which make use of their arithmetic/geometrical and spectral properties.

Acknowledgments. The present survey follows essentially a series of lectures given in Strasbourg on the occasion of the "75-ièmes Rencontres entre Physiciens Théoriciens et Mathématiciens" and in Kyoto on the occasion of the RIMS workshop "Automorphic forms and automorphic L-functions". I would like to thank the organizers for their kind invitation. I also thank E. Royer, who provided the figures 1 and 2 given below.

2 Kloosterman sums from the algebraic viewpoint

Let us write once again the definition of Kloosterman sums: given three integers $a, b, c \geqslant 1$ such that $(ab, c) = 1$,

$$\mathrm{Kl}(a, b; c) = \sum_{\substack{x(c)\\(x,c)=1}} e\Big(\frac{ax + b\bar{x}}{c}\Big), \quad x\bar{x} \equiv 1(c),\ e(*) = \exp(2\pi i *).$$

Such sums satisfy various elementary properties:

$$\mathrm{Kl}(a, b; c) = \mathrm{Kl}(1, ab; c),\ \overline{\mathrm{Kl}(a, b; c)} = \mathrm{Kl}(-a, -b; c) = \mathrm{Kl}(a, b; c),$$

i.e., $\mathrm{Kl}(a, b; c) \in \mathbb{R}$.

There is another simple property which follows for the Chinese Remainder Theorem, namely the

Twisted multiplicativity property. If c_1, c_2 are coprime integers and $c = c_1c_2$, one has

$$\mathrm{Kl}(a, b; c) = \mathrm{Kl}(a\bar{c}_2, b\bar{c}_2; c_1)\,\mathrm{Kl}(a\bar{c}_1, b\bar{c}_1; c_2).$$

In particular, twisted multiplicativity reduces the problem of estimating Kloosterman sums to the case of a prime power modulus, and the only non-elementary case is that of a prime modulus.

In the prime modulus case, the first non-trivial bound for Kloosterman sums was given by Kloosterman himself. By computing (using purely elementary methods) the fourth moment of Kloosterman sums, he obtained

$$\sum_{a\neq 0 \pmod p} |\operatorname{Kl}(a,b;p)|^4 \leqslant 16p^3$$

so that

$$|\operatorname{Kl}(a,b;p)| \leqslant 2p^{3/4}.$$

This bound was already sufficient to solve the problem of representing an integer by a quaternary quadratic form.

In the 1950s, as a consequence of his resolution of the Riemann Hypothesis for curves over functions fields, A. Weil [33] established a stronger (probably optimal) bound

$$|\operatorname{Kl}(a,b;p)| \leqslant 2p^{1/2};$$

more precisely (when $p > 2$)

$$\operatorname{Kl}(a,b;p) = \alpha_{p,ab} + \beta_{p,ab}$$

is the sum of two algebraic integers such that $|\alpha_{p,ab}| = |\beta_{p,ab}| = \sqrt{p}$ and $\alpha_p\beta_p = p$.

It is then natural to raise the question of the optimality of Weil's bound. For this one defines the *angle*, $\theta_{p,ab} \in [0,\pi[$, of the Kloosterman sum $\operatorname{Kl}(a,b;p)$ by

$$\operatorname{Kl}(a,b;p) = 2p^{1/2}\cos(\theta_{p,ab}).$$

More generally, for c a squarefree integer and $(ab,c) = 1$, the *angle* $\theta_{c,ab} \in [0,\pi[$, of $\operatorname{Kl}(a,b;c)$ is defined by

$$\operatorname{Kl}(a,b;c) := 2^{\omega(c)}c^{1/2}\cos(\theta_{c,ab}).$$

The twisted multiplicativity is then expressed by the formula

$$\cos(\theta_{c_1c_2,ab}) = \cos(\theta_{c_1,\bar{c}_2^2ab})\cos(\theta_{c_2,\bar{c}_1^2ab}).$$

3 Horizontal and vertical distribution laws

The angle of a Kloosterman sum $\theta_{c,a}$ depends on two parameters: the *modulus* c (which in our case will be squarefree with a fixed number of prime factors) and the *argument* a (which is an integer coprime with c). Following the terminology introduced by N. Katz, we call

- *horizontal* a law which describes the asymptotic distribution of the family of angles

$$\{\theta_{c,a}\}_{\substack{1\leqslant a\leqslant c\\(a,c)=1}}$$

for $c \to +\infty$ amongst the squarefree integers with a fixed number of prime factors,

- *vertical* a law which describes the asymptotic distribution of the family of angles

$$\{\theta_{c,a}\}_{\substack{1\leqslant c\leqslant C\\(a,c)=1}}$$

for a fixed non-zero integer a, $C \to +\infty$ and c ranging amongst the squarefree integers less than C, coprime with a and with a fixed number of prime factors.

Of the two types of possible laws above, horizontal laws are by far the most mysterious.

4 The vertical Sato/Tate law (VST)

In the vertical direction, the situation is completely understood thanks to the work of Deligne and Katz [4], [5], [22].

VST . *When $p \to +\infty$, the angles $\{\theta_{p,a}\}_{1\leqslant a\leqslant p-1}$ become equidistributed relatively to the Sato/Tate measure on $[0,\pi]$,*

$$d\mu_{\mathrm{ST}}(\theta) = \frac{2}{\pi}\sin^2(\theta)\,d\theta,$$

i.e., for any $\theta \in [0,\pi]$

$$\frac{|\{\theta_{p,a},\ 1\leqslant a\leqslant p-1,\ 0\leqslant\theta_{p,a}\leqslant\theta\}|}{p-1} \longrightarrow \mu_{\mathrm{ST}}([0,\theta]) = \frac{2}{\pi}\int_0^\theta \sin^2(t)\,dt, \quad p\to+\infty.$$

The proof of Katz's vertical Sato/Tate law is a combination of three key ingredients.

(1) Deligne's Equidistribution Theorem for Frobenius conjugacy classes (a consequence of his fundamental theorem on weights).

(2) Katz's construction of the *Kloosterman sheaf*: For $p > 2$ and $\ell \neq p$, there exists an ℓ-adic sheaf $\mathcal{K}l$ such that the following holds:

- $\mathcal{K}l$ has rank 2, is lisse on $\mathbb{G}_{m,\mathbb{F}_p} = \mathbb{P}^1_{\mathbb{F}_p} - \{0,\infty\}$, irreducible with trivial determinant: $\mathcal{K}l$ "is" a 2-dimensional irreducible representation

$$\mathcal{K}l\colon \pi_1^{\mathrm{arith}}(\mathbb{G}_m) \to \mathrm{SL}_2(E_\lambda)$$

where E_λ is a finite extension of $\mathbb{Q}_\ell$;

- $\mathcal{K}l$ is pure of weight 0 and for $a \in \mathbb{G}_m(\mathbb{F}_p) = \mathbb{F}_p^\times$

$$\mathrm{tr}(\mathrm{Frob}_a|\,\mathcal{K}l) = \frac{\mathrm{Kl}(1,a;\,p)}{\sqrt{p}}.$$

(3) Katz computed the ramification at 0 and ∞ of $\mathcal{K}l$, enabling him to show that the geometric monodromy group of $\mathcal{K}l$ is as big as possible.

- $\mathcal{K}l$ has unipotent ramification at 0 and is totally wild at ∞ with Swan conductor equal to 1 (in particular this is independent of p);
- $\mathcal{K}l(\pi_1^{\mathrm{geom}}(\mathbb{G}_m)) = \mathcal{K}l(\pi_1^{\mathrm{arith}}(\mathbb{G}_m)) = \mathrm{SL}_2$.

4.1 Proof of the vertical Sato/Tate law

- One embeds $\overline{\mathbb{Q}}_\ell$ into $\mathbb{C}$ and one chooses K a maximal compact subgroup of $\mathrm{SL}_2(\mathbb{C})$ ($K = \mathrm{SU}(2)$ say). $\mathcal{K}l$ being pure of weight 0 with image contained in $\mathrm{SL}_2(\mathbb{C})$, the Frobenius conjugacy classes $\{\mathcal{K}l(\mathrm{Frob}_a)\}_{a\in\mathbb{F}_p^\times}$ define conjugacy classes into $K^\natural$. The latter is identified with $[0,\pi]$ and the direct image of the Haar measure is μ_{ST}.
- By the Weyl equidistribution criterion and the Peter/Weyl theorem (and the unitary trick) it is then sufficient to show that for any non-trivial irreducible representation of SL_2 (Sym_k say) the corresponding Weyl sum is small, that is,

$$\frac{1}{p-1}\sum_{a=1}^{p-1}\mathrm{tr}(\mathrm{Frob}_a \,|\, \mathrm{Sym}^k\,\mathcal{K}l) = \frac{1}{p-1}\sum_{a=1}^{p-1}\mathrm{Sym}_k(\theta_{p,a}) \to 0,$$

  as $p \to +\infty$ (here $\mathrm{Sym}_k(\theta) = \frac{\sin((k+1)\theta)}{\sin(\theta)}$).
- By the Lefschetz trace formula, the latter sum equals

$$\frac{1}{p-1}\big[\,\mathrm{tr}(\mathrm{Frob}_p \,|\, H_c^0(\mathbb{G}_m|\,\mathrm{Sym}^k\,\mathcal{K}l)) - \mathrm{tr}(\mathrm{Frob}_p \,|\, H_c^1(\mathbb{G}_m|\,\mathrm{Sym}^k\,\mathcal{K}l)) + \mathrm{tr}(\mathrm{Frob}_p \,|\, H_c^2(\mathbb{G}_m|\,\mathrm{Sym}^k\,\mathcal{K}l))\big].$$

- Now, by Katz's determination of the geometric monodromy group of $\mathcal{K}l$, $\mathrm{Sym}^k\,\mathcal{K}l$ is geometrically irreducible, hence $H_c^0 = H_c^2 = 0$, and by Deligne's theorem

$$|\mathrm{tr}(\mathrm{Frob}_p \,|\, H_c^1(\mathbb{G}_m|\,\mathrm{Sym}^k\,\mathcal{K}l))| \leqslant \dim H_c^1(\mathbb{G}_m|\,\mathrm{Sym}^k\,\mathcal{K}l)\,p^{1/2}.$$

- Finally, by the Grothendieck/Ogg/Shafarevitch formula, $\dim H_c^1(\mathbb{G}_m|\,\mathrm{Sym}^k\,\mathcal{K}l)$ can be estimated only in terms of the ramification of $\mathrm{Sym}^k\,\mathcal{K}l$ and shown to be bounded in terms of $\dim \mathrm{Sym}^k = k+1$ but independently of p.

4.2 A variant of VST for composite moduli

When the modulus $c = p_1 p_2 \dots p_k$ is composite with say $k \geqslant 1$ prime factors, an analog of the vertical Sato/Tate law quickly follows from the twisted multiplicativity,

$$\cos(\theta_{c,a}) = \frac{\mathrm{Kl}(1,a;c)}{2^{\omega(c)}\sqrt{c}} = \prod_{p|c}\cos(\theta_{p,a\overline{c/p}^2}),$$

and from the Chinese Remainder Theorem:

For $k \geqslant 1$, let $\mu_{\mathrm{ST}}^{(k)}$ denote the measure on $[0, \pi]$ provided by the direct image of $\mu_{\mathrm{ST}}^{\otimes k}$ of the map $[0, \pi]^k \to [0, \pi]$ given by

$$(\theta_1, \dots, \theta_k) \to \arccos(\cos(\theta_1) \dots \cos(\theta_k)).$$

VST(k). *As $c \to +\infty$ amongst the squarefree integers having k prime factors none of which is small (for example $p \mid c \Rightarrow p \geqslant c^{1/2k}$) the angles $\{\theta_{c,a}\}_{\substack{1 \leqslant a \leqslant c \\ (a,c)=1}}$ are equidistributed on $[0, \pi]$ relatively to the measure $\mu_{\mathrm{ST}}^{(k)}$.*

5 The horizontal Sato/Tate conjecture

The *horizontal* analog of the vertical Sato/Tate law was conjectured by Katz (before his proof of the VST) as a close analog of the Sato/Tate conjecture for elliptic curves.

Conjecture HST. *Given $a \geqslant 1$; as $P \to +\infty$, the angles $\{\theta_{p,a}\}_{\substack{p \leqslant P \\ (a,p)=1}}$ become equidistributed relatively to the Sato/Tate measure on $[0, \pi]$,*

$$d\mu_{\mathrm{ST}}(\theta) = \frac{2}{\pi} \sin^2(\theta) d\theta,$$

i.e., for any $\theta \in [0, \pi]$

$$\frac{|\{\theta_{p,a},\ p \leqslant P,\ (a,p)=1,\ 0 \leqslant \theta_{p,a} \leqslant \theta\}|}{|\{p \leqslant P\}|} \to \mu_{\mathrm{ST}}([0, \theta]) = \frac{2}{\pi} \int_0^\theta \sin^2(t) dt,\ P \to +\infty.$$

It is remarkable how little we know about this conjecture. One still does not know the answer to the following simple questions:

Question 1. Are there

(1) infinitely many primes p for which $\mathrm{Kl}(1, 1; p) > 0$?

(2) infinitely many primes p for which $\mathrm{Kl}(1, 1; p) < 0$?

Question 2. Is there an $\varepsilon > 0$ for which

(1) there are infinitely many primes p for which $|\mathrm{Kl}(1, 1; p)| \leqslant (1 - \varepsilon) p^{1/2}$?

(2) there are infinitely many primes p for which $|\mathrm{Kl}(1, 1; p)| \geqslant \varepsilon p^{1/2}$?

(3) or even, there are infinitely many primes such that $|\mathrm{Kl}(1, 1; p)| \geqslant \varepsilon$?

One could answer such questions if one had some non-trivial analytic information on the Euler products

$$L(\mathrm{Kl}_1, s) = \prod_{p>2} \Big(1 - \frac{\mathrm{Kl}(1, 1; p)}{p^s} + \frac{p}{p^{2s}}\Big)^{-1} = \prod_{p>2} \Big(1 - \frac{\alpha_{p,1}}{p^s}\Big)^{-1} \Big(1 - \frac{\beta_{p,1}}{p^s}\Big)^{-1}$$

and

$$L(\mathrm{Sym}^2, \mathrm{Kl}_1, s) = \prod_{p>2} \Big(1 - \frac{\alpha_{p,1}^2}{p^s}\Big)^{-1} \Big(1 - \frac{\beta_{p,1}^2}{p^s}\Big)^{-1} \Big(1 - \frac{\alpha_{p,1}\beta_{p,1}}{p^s}\Big)^{-1}.$$

Apparently, at that moment, no one has a clue on how to get some control of these Euler products (like analytic continuation in a non-obvious region). Of course this would be the case if $L(\mathrm{Kl}_1, s + 1/2)$ were the Hecke L-function of an automorphic form (probably a weight 0 Maass form of level a multiple of 2)…

…but in fact some numerical computations of A. Booker show that it is very unlikely to be the case [3].

Still, there are some reasons to believe in the horizontal Sato/Tate conjecture. The first one is the validity of Katz's vertical Sato/Tate law. Another reason is that numerical computations of Kloosterman sums show a very good agreement with HST.

(1) Figure 1 represents the cumulative distribution functions of the measures induced from the vertical and horizontal families of points

$$\{\cos(\theta_{p,a})\}_{1\leqslant a\leqslant p},\ \{\cos(\theta_{p,1})\}_{p\in[P,2P]}$$

and the Sato/Tate cumulative distribution function $x \to \frac{2}{\pi}\int_{-1}^{x}\sqrt{1-t^2}dt$ for $p \sim P \sim 25000$.

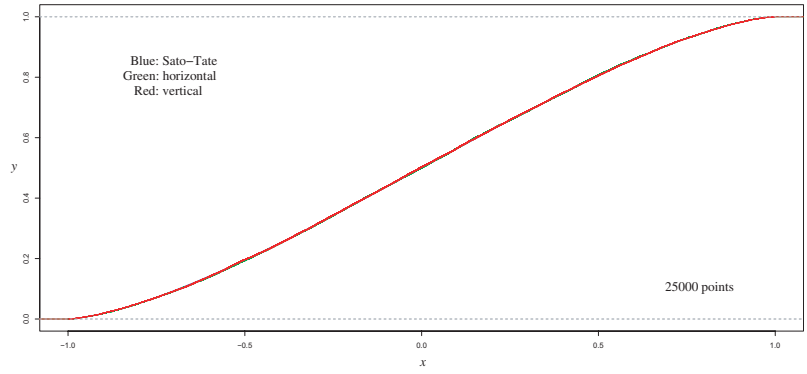


Figure 1. Graphs of the cumulative distribution functions (CDF) of HST, VST and Sato/Tate for 25000 points.

(2) Figure 2 represents the differences between the cumulative distribution functions of the measures induced from the vertical and horizontal families of points

$$\{\cos(\theta_{p,a})\}_{1\leqslant a\leqslant p-1} \text{ and } \{\cos(\theta_{p,1})\}_{p\in[P,2P]}$$

with the Sato/Tate cumulative distribution function for $p \sim P \sim 25000$. One should note that these differences are bounded by $0.005 \leqslant 1/\sqrt{25000}$.

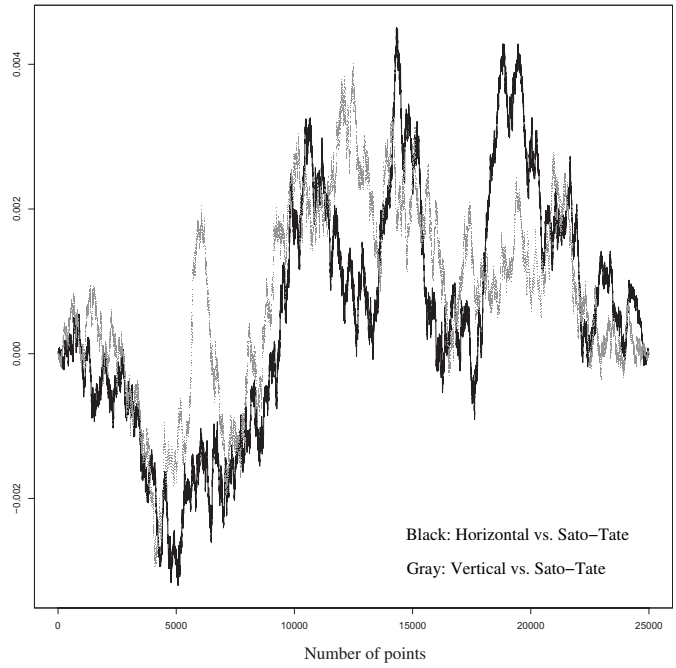


Figure 2. Graphs of the differences between the CDF of HST and VST with the Sato/Tate CDF for 25000 points.

But probably the best reason to believe in HST is the fact that for other exponential sums the horizontal Sato/Tate Law was proven.

(1) Heath-Brown/Patterson [20] established the (uniform) equidistribution in $[0, 2\pi]$ of the angles of cubic Gauss sums

$$G((\frac{-}{\pi})_3) = \sqrt{N_{K/\mathbb{Q}}(\pi)}e^{i\theta_\pi}$$

(associated to the cubic residue symbol of $K = \mathbb{Q}(\sqrt{-3})$) for (split) prime moduli π.

(2) Duke/Friedlander/Iwaniec [13] established the equidistribution of angles of Salié sums

$$S(1, 1; p) = \sum_{\substack{x(p)\\(x,p)=1}} \left(\frac{x}{p}\right) e\left(\frac{x+\bar{x}}{p}\right) =: 2\sqrt{p}\cos(\theta^S_{p,1})$$

for the uniform measure on $[0, \pi]$. In the analogy with the classical Sato/Tate conjecture for elliptic curves, the case of Salié sum would correspond to the case of CM elliptic curves (but the proof of equidistribution of Salié sums is much harder).

The proofs of both cases above make use of

- Sieve techniques (to detect prime moduli amongst arbitrary moduli),
- the analytic theory of automorphic forms (to show that the Weyl sums corresponding to these equidistribution problems are small).

6 Horizontal Sato/Tate for composite moduli

Interestingly, some variants of the vertical Sato/Tate law can be used to provide results in the *horizontal* direction. For this one has to mollify the problem by allowing composite moduli.

Let $c = p_1 \dots p_k$ be squarefree with a fixed number, $k \geqslant 2$, of primes factors; by twisted multiplicativity the angles of the Kloosterman sum $\mathrm{Kl}(1, 1; c)$ satisfy

$$\cos(\theta_{c,1}) = \frac{\mathrm{Kl}(1, 1; c)}{2^k \sqrt{c}} = \prod_{p|c} \cos(\theta_{p,\overline{(c/p)}^2}).$$

Then, since we do expect the primes to behave independently, it is reasonable to make the following

Conjecture HST(k). *As $C \to +\infty$, the family of angles $\{\theta_{c,1}\}_{c \leqslant C}$ are equidistributed on $[0, \pi]$ relatively to the Sato/Tate measure of order k, $\mu_{ST}^{(k)}$, where the moduli c range over the squarefree integers $\leqslant C$, with k prime factors none of which is small (for example $p|c \Rightarrow p \geqslant c^{1/2k}$).*

Observe that conjecture HST(k) is not implied by conjecture HST and at the present time seems as intractable as the original one. On the other hand, one can ask the same basic questions about the size and the existence of sign changes amongst Kloosterman sums with composite moduli. As we shall see, due to the extra flexibility allowed by multiple prime factors in the moduli, both questions can be answered affirmatively if k is sufficiently large.

6.1 Three variants of the vertical Sato/Tate law

Theorem 6.1. *There exist infinitely many pairs of distinct primes (p, q) such that $|\mathrm{Kl}(1, 1; pq)| \geqslant \frac{4}{25}\sqrt{pq}$. More precisely, for X large enough,*

$$\left|\left\{(p, q),\ p \neq q,\ p, q \leqslant P,\ |\mathrm{Kl}(1, 1; pq)| \geqslant \tfrac{2}{25}\sqrt{pq}\right\}\right| \gg P^2/\log^2 P.$$

In particular (take $P = X^{1/2}$)

$$\sum_{\substack{c \leqslant X \\ \omega(c)=2}} \mu^2(c) \frac{|\mathrm{Kl}(1, 1; c)|}{\sqrt{c}} \gg \frac{X}{\log^2 X}$$

(*here the implied constant can be evaluated exactly*).

In particular, this first result shows that Weil's bound is indeed optimal in the horizontal aspect for moduli with 2 prime factors [28]! The next result shows that at the expense of allowing an extra prime factor in the moduli, one can improve the second inequality by a factor $\log X$ [17].

Theorem 6.2.

$$\sum_{\substack{c\leqslant X\\ \omega(c)=3}} \mu^2(c)\frac{|\mathrm{Kl}(1,1;c)|}{\sqrt{c}} \geqslant 0.078\frac{X}{\log X}.$$

Finally we state a mean square estimate on Kloosterman sums with arbitrary moduli, gaining an extra $\exp((\log\log X)^{5/17})$ factor [18].

Theorem 6.3. *For $X \to +\infty$ one has*

$$X\frac{\exp((\log\log X)^{5/17})}{\log X} \ll \sum_{c\leqslant X}\frac{|\mathrm{Kl}(1,1;c)|^2}{c} \ll X(\log\log X)^3.$$

Observe that the last upper bound is also non-trivial (the trivial upper bound being $\ll X(\log X)^3$); on the other hand, a natural probabilistic model for Kloosterman sums predicts that

$$\sum_{c\leqslant X}\frac{|\mathrm{Kl}(1,1;c)|^2}{c} \simeq bX$$

for some constant $b > 0$ that can be explicitly computed.

The proofs of these three results are based on common principles.

For some $k \geqslant 2$, let $\mathcal{C}_k$ be the set of squarefree integers c less than X with k prime factors. We want to show that for many (i.e. a positive proportion of) $c \in \mathcal{C}_k$ the Kloosterman sum is large.

Step 1. For some decomposition $k = k_1 + k_2$ with $k_1, k_2 \geqslant 1$, pick $\Omega_1 \subset \mathcal{C}_{k_1}$ and $\Omega_2 \subset \mathcal{C}_{k_2}$, two (big enough) subsets of squarefree integers with k_1 and k_2 prime factors respectively such that

$$\Omega_1.\Omega_2 := \{c_1.c_2,\ c_1 \in \Omega_1,\ c_2 \in \Omega_2\} \subset \mathcal{C}_k.$$

Here big enough means in particular that

$$|\Omega_1.\Omega_2| \gg |\mathcal{C}_k| \simeq c(k)\frac{X}{\log X}, \quad c(k) > 0.$$

Step 2. Prove that, for $X \to +\infty$, the two families of angles of Kloosterman sums

$$\left\{\theta_{c_1,\bar{c}_2^2},\ (c_1,c_2) \in \Omega_1 \times \Omega_2\right\},$$
$$\left\{\theta_{c_2,\bar{c}_1^2},\ (c_1,c_2) \in \Omega_1 \times \Omega_2\right\}$$

become equidistributed on $[0, \pi]$ relatively to the respective measures $\mu_{\mathrm{ST}}^{(k_1)}$ and $\mu_{\mathrm{ST}}^{(k_2)}$. In particular, this implies that there is $\alpha > 0$ and $\delta > 0$ such that

$$\left|\left\{(c_1, c_2) \in \Omega_1 \times \Omega_2 \text{ such that } |\cos\theta_{c_1,\bar{c}_2^2}| \geqslant \alpha\right\}\right| \geqslant (1/2 + \delta)|\Omega_1.\Omega_2|,$$
$$\left|\left\{(c_1, c_2) \in \Omega_1 \times \Omega_2 \text{ such that } |\cos\theta_{c_2,\bar{c}_1^2}| \geqslant \alpha\right\}\right| \geqslant (1/2 + \delta)|\Omega_1.\Omega_2|.$$

Step 3. Use the following simple probability argument

Lemma. *Let* (Ω, μ) *be a probability space and let* $\Omega_1, \Omega_2 \subset \Omega$ *such that* $\mu(\Omega_1) + \mu(\Omega_2) > 1$. *Then*

$$\mu(\Omega_1 \cap \Omega_2) \geqslant \mu(\Omega_1) + \mu(\Omega_2) - 1 > 0.$$

along with the twisted multiplicativity property

$$\cos\theta_{c_1,\bar{c}_2^2} \cos\theta_{c_2,\bar{c}_1^2} = \cos\theta_{c_1c_2,1} = \frac{\mathrm{Kl}(1, 1; c_1c_2)}{2^{k_1+k_2}\sqrt{c_1c_2}}$$

to conclude that

$$\left|\left\{c_1c_2 \in \Omega_1.\Omega_2 \text{ such that } \tfrac{|\mathrm{Kl}(1,1;c_1c_2)|}{2^{k_1+k_2}\sqrt{c_1c_2}} \geqslant \alpha^2\right\}\right| \geqslant 2\delta\left|\Omega_1.\Omega_2\right| \gg \left|\mathcal{C}_k\right| \gg \frac{X}{\log X}.$$

6.2 Some variants of the vertical Sato/Tate laws

Clearly performing Step 2, i.e., the equidistribution of the angles

$$\left\{\theta_{c_1,\bar{c}_2^2},\ (c_1, c_2) \in \Omega_1 \times \Omega_2\right\},$$

amounts to proving some variant of Katz's vertical Sato/Tate laws (possibly extended to composite moduli). In these variants the set of arguments of the angle of Kloosterman sums considered is not merely the set of integers contained in the interval $[1, c_1]$ but a more general set of integers. By standard techniques from analytic number theory it is indeed possible to pass to such smaller sets.

In practice the set of moduli Ω_1 and of arguments Ω_2 are contained in two intervals $[1, X_1]$ and $[1, X_2]$ respectively, where X_1, X_2 are two positive powers of X with $X_1X_2 = X$ and one has

$$|\Omega_1| \gg X_1/\log X_1,\ |\Omega_2| \gg X_2/\log X_2.$$

The easiest case is when the set of argument is somewhat greater than the set of moduli, say

$$X_2 \gg X_1 \log^{100} X.$$

Here one uses large sieve techniques (like the Barban/Davenport/Halberstam inequality) to show that on average over the moduli $c_1 \in \Omega_1$ the set of arguments $\bar{c}_2^2 \pmod{c_1}$ with $c_2 \in \Omega_2$ are very well distributed amongst the congruence classes of the c_1. Since the congruence classes are essentially well covered one can now invoke the vertical Sato/Tate law VST or its variant VST(k_1) to finish the proof of equidistribution.

A harder case is when $X_2 \leqslant X_1$; in that case not all congruence classes mod c_1 are covered and one has to prove that the vertical Sato/Tate law still holds for restricted subsets.

For example assuming that $c_1 = p$ is a prime, by Weyl's criterion, one needs to prove that for any $k \geqslant 1$

$$\frac{1}{|\Omega_2|} \sum_{c_2 \in \Omega_2} \mathrm{Sym}_k(\theta_{p,\bar{c}_2^2}) \to 0,$$

which by a slight abuse of notation is rewritten as

$$\frac{1}{|\Omega_2|} \sum_{c_2 \in \Omega_2} \mathrm{tr}(\mathrm{Frob}_{\bar{c}_2^2} \mid \mathrm{Sym}_k \mathcal{K}l) \to 0.$$

Here we have identified $\mathbb{F}_p^\times$ with the interval of integers $[1, p-1]$ (which is odd from the view point of algebraic geometry but perfectly natural from the viewpoint of analytic number theory). Of course, in order to prove such estimate we cannot apply directly the methods from algebraic geometry, but after several (more or less complicated) transformations this becomes possible.

The simplest example of a VST over a restricted subset is the case of a short interval.

Proposition. *Given any $\varepsilon > 0$, as $p \to +\infty$, the set of Kloosterman angles $\{\theta_{p,a}\}_{1 \leqslant a \leqslant p^{1/2+\varepsilon}}$ becomes equidistributed for the Sato/Tate measure.*

To obtain this result, one express the characteristic function of the integers in the interval $[1, p^{1/2+\varepsilon}]$ in terms of the additive characters mod p. Then the Polya/Vinogradov completion method reduces the estimate of the corresponding Weyl sums to prove that

$$\sum_{a \in \mathbb{F}_p^\times} \mathrm{tr}(\mathrm{Frob}_a \mid \mathrm{Sym}_k \mathcal{K}l)\psi(a) \ll_k p^{1/2}$$

uniformly for ψ ranging over the additive characters of $\mathbb{F}_p$. In particular such bounds lead to the consideration of another family of sheaves, the twisted sheaves

$$\mathrm{Sym}^k \mathcal{K}l \otimes \mathcal{L}_\psi$$

where $\mathcal{L}_\psi$ ranges over the rank one sheaves on $\mathbb{A}^1_{\mathbb{F}_p}$ associated with the characters ψ of $(\mathbb{F}_p, +)$.

The new (easy) algebro-geometric input here is the (simple) fact that a geometrically irreducible sheaf of rank > 1 (i.e. $\mathrm{Sym}^k \mathcal{K}l$) remains geometrically irreducible if it is twisted by any sheaf of rank 1 (i.e. $\mathcal{L}_\psi$).

Another example is that of the set of primes less than p.

Theorem. *As $p \to +\infty$, the set of Kloosterman angles $\{\theta_{p,q}\}_{1 \leqslant q \leqslant p-1, \ q \text{ prime}}$ becomes equidistributed for the Sato/Tate measure.*

The second variant is more involved and requires more sophisticated transformations coming from sieve methods. After these transformations are performed one needs to use the previous variant and another form of the VST.

Proposition. *For any $\varepsilon > 0$, as $p \to +\infty$, and for any $b \in \mathbb{F}_p - \{0, 1\}$, the set of pairs Kloosterman angles $\{(\theta_{p,a}, \theta_{p,ba})\}_{1 \leqslant a \leqslant p^{1/2+\varepsilon}}$ becomes equidistributed on $[0, \pi] \times [0, \pi]$ for the product of Sato/Tate measures.*

To prove this proposition, the new sheave to be considered is the Rankin/Selberg sheaf

$$\mathrm{Sym}^k \mathcal{K}l \otimes \mathrm{Sym}^k [b]^* \mathcal{K}l$$

where $b \in \mathbb{F}_p^\times - \{1\}$ and $[b]\colon x \to bx$ denote the (non-trivial) translation on $\mathbb{G}_m$. Finally the main geometrical result needed is an independence statement for Kloosterman sheaves.

Proposition. *If $b \neq 1$ the geometric monodromy group of $\mathcal{K}l \oplus [b]^* \mathcal{K}l$ is as big as possible, i.e., equals $\mathrm{SL}_2 \times \mathrm{SL}_2$.*

The latter proposition follows from the Goursat/Kolchin/Ribet criterion which is verified either by using the Rankin/Selberg method or by comparing the monodromies at ∞ of $\mathcal{K}l$ and $[b]^* \mathcal{K}l$.

More elaborated transformation can be used to obtain other geometric statements. For example, at some point one uses the following proposition

Proposition. *Let $p > 2$, $m_1, m_2 \in \mathbb{F}_p^\times$, and let $U = \mathbb{G}_{m,\mathbb{F}_p} - \{m_1, m_2\}$ be an open subset. Consider the four morphisms $f_1, f_2, f_3, f_4\colon U \to \mathbb{G}_{m,\mathbb{F}_p}$ given by*

$$f_1(T) = \big(m_1(m_1 - T)\big)^{-2}, \quad f_2(T) = \big(T(m_1 - T)\big)^{-2}$$
$$f_3(T) = \big(m_2(m_2 - T)\big)^{-2}, \quad f_4(T) = \big(T(m_2 - T)\big)^{-2}.$$

Then the geometric monodromy group of the Sheaf

$$f_1^* \mathcal{K}l \oplus f_2^* \mathcal{K}l \oplus f_3^* \mathcal{K}l \oplus f_4^* \mathcal{K}l$$

is as big as possible, i.e., equals

$$\mathrm{SL}_2 \times \mathrm{SL}_2 \times \mathrm{SL}_2 \times \mathrm{SL}_2 .$$

Almost all what has been said so far can be generalized to wide classes of families of algebraic exponential sums (essentially families for which the geometric monodromy group has been computed – mostly by Katz). Such example include higher dimensional Kloosterman sums

$$\mathrm{Kl}_n(a; p) = \sum_{\substack{x_1 x_2 \ldots x_n = a \\ x_i \in \mathbb{F}_p^\times}} e\Big(\frac{x_1 + x_2 + \cdots + x_n}{p}\Big),$$

more general hypergeometric sums or exponential sums obtained by geometric Fourier transforms

$$S_{x^3+x}(a;\,p) := \sum_{x\in\mathbb{F}_p} e\Big(a\frac{x^3+x}{p}\Big)$$

or

$$S_f(a;\,p) = \sum_{x\in\mathbb{F}_p} e\Big(a\frac{f(x)}{p}\Big)$$

where $f(X)\in\mathbb{Z}[X]$ is an irreducible polynomial of degree n with maximal Galois group (i.e. $\simeq \Sigma_n$).

7 Spectral theory of Kloosterman sums

For $q\geqslant 1$ et $k\geqslant 2$ even, let $S_k(q)$ be the space of holomorphic cusp forms of weight k and level q: $f:\mathbb{H}\to\mathbb{C}$,

$$f\Big(\frac{az+b}{cz+d}\Big) = (cz+d)^k f(z),\quad \begin{pmatrix} a & b\\ c & d\end{pmatrix}\in\Gamma_0(q)\subset \mathrm{SL}_2(\mathbb{Z}),\ c\equiv 0(q).$$

The Petersson inner product is given by

$$\langle f,\,f\rangle := \int_{\Gamma_0(q)\backslash\mathbb{H}} y^k|f(z)|^2\frac{dxdy}{y^2} < +\infty.$$

Let

$$f(z+1) = f(z) = \sum_{n\geqslant 1}\rho_f(n)n^{\frac{k-1}{2}}e(nz)$$

be the Fourier expansion of $f\in S_k(q)$. For $\mathcal{B}_k(q)$ an orthonormal basis of $S_k(q)$, Peterson's formula is

$$\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\sum_{f\in\mathcal{B}_k(q)}\overline{\rho_f}(m)\rho_f(n) = \delta_{m,n} - 2\pi i^k\sum_{c\equiv 0(q)}\frac{\mathrm{Kl}(m,n;c)}{c}J_{k-1}\Big(\frac{4\pi\sqrt{mn}}{c}\Big).$$

In particular, for $\sigma > 1/2$ Weil's bound for Kloosterman sums implies that

$$\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\sum_{f\in\mathcal{B}_k(q)}\overline{\rho_f}(m)\rho_f(n) = \delta_{m,n} + O\Big((m,n,q)^{1/2}\Big(\frac{\sqrt{mn}}{q}\Big)^{\sigma}\Big),$$

hence taking $m=n$ above and $\sigma = 1/2+\varepsilon$ one gets

$$|\rho_f(n)| \ll_f n^{1/4+\varepsilon},$$

which is a non-trivial bound (the trivial bound being $|\rho_f(n)|\ll_f n^{1/2+\varepsilon}$). Of course much better bounds are now available thanks to Deligne, but it is nevertheless interesting to note that any non-trivial estimate for Kloosterman sums is already sufficient

to bound the Fourier coefficients non-trivially; for instance Kloosterman's original bound gives $|\rho_f(n)| \ll_f n^{3/8+\varepsilon}$.

For us the most useful version of this formula is Kuznetsov's generalization of the Petersson formula [26]. Consider the L^2-space of functions on the (punctured) modular curve $\Gamma_0(q)\backslash\mathbb{H}$ equipped with the Petersson inner product,

$$\langle f, g\rangle = \int_{\Gamma_0(q)\backslash\mathbf{H}} \bar{f}(z)g(z)\frac{dxdy}{y^2}.$$

Then $L^2(\Gamma_0(q)\backslash\mathbb{H})$ is decomposed spectrally by eigenfunction of the hyperbolic Laplace operator (Maass forms, which we may also choose to be Hecke-eigenforms)

$$\Delta = -y^2\left(\frac{\partial^2}{\partial^2 x} + \frac{\partial^2}{\partial^2 y}\right),$$

$$L^2(\Gamma_0(q)\backslash\mathbb{H}) = \mathbb{C} \oplus \bigoplus_{j\geqslant 1} \mathbb{C}.u_j \oplus \bigoplus_{\mathfrak{a}} \frac{1}{4\pi}\int_{\mathbb{R}} \mathbb{C}.E_{\mathfrak{a}}(\,.\,, 1/2+it)\,dt.$$

If u_j is such a Maass (cusp) form satisfying say

$$\Delta.u_j = \left(\frac{1}{4} + t_j^2\right)u_j,$$

its Fourier expansion is given by

$$u_j(z) = \sum_{n\neq 0} \rho_j(n) W_{0,it_j}(4\pi|n|y)e(nx).$$

The Petersson/Kuznetsov formula relates sums of Kloosterman sums to Fourier coefficients of modular forms.

Let $\varphi \in C_c^\infty((0,\infty))$ be a test function and let $m, n \geqslant 1$. Then one has

$$\begin{aligned}\frac{1}{4}\sum_{c\equiv 0\,(q)} \frac{\mathrm{Kl}(m,n;c)}{c}\varphi\left(\frac{4\pi\sqrt{mn}}{c}\right) &= \sum_{k\equiv 0\,(2)} \Gamma(k-1)\tilde{\varphi}(k-1) \sum_{f\in\mathcal{B}_k(q)} \overline{\rho_f}(m)\rho_f(n)\\ &+ \sum_{j\geqslant 1} \frac{\hat{\varphi}(t_j)}{\cosh(\pi t_j)}\overline{\rho_j}(m)\rho_j(n)\\ &+ \frac{1}{4\pi}\sum_{\mathfrak{a}}\int_{-\infty}^{+\infty} \frac{\hat{\varphi}(t)}{\cosh(\pi t)}\overline{\rho_{\mathfrak{a}}}(m,t)\rho_{\mathfrak{a}}(n,t)\,dt.\end{aligned}$$

Here $\tilde{\varphi}$ and $\hat{\varphi}$ denote some Bessel transforms of φ. There are variants of this formula for forms of weight one and forms of half-integral weight and all these formulae are better understood in the context of the relative trace formula. Note that in this latter context the ultimate goal is to use the connection

$$\begin{gathered}\text{Spectral Data (= Fourier coefficients of automorphic forms)}\\ \Longleftrightarrow \text{ Geometric data (= sums of Kloosterman sums)}\end{gathered}$$

to compare spectral data associated to very different groups. In analytic number theory (so far) this connection is used, yet in both directions, but only at the level of a single group. This is what analytic number theorists call (after Martin Huxley) *Kloostermania*. In the sequel we will describe some recent applications of Kloostermania and in particular make some links with the first part of this survey.

The most obvious application of Kloosterman sums is as above to provide estimates for Fourier coefficients of automorphic forms (and more generally) to give estimates for the spectral parameters of the associated automorphic representations. As above, Weil's bound can be used along with Kuznetsov's formula to prove that for a Hecke eigenform $g(z)$

- $|\lambda_g(p^\alpha)| \leqslant 2p^{\alpha/4}$, and
- $|\Im m t_g| \leqslant 1/4$ if f is a Maass form with Laplace eigenvalue

$$\lambda_g = (1/2 + it_g)(1/2 - it_g).$$

The last bound is due to Selberg and was also obtained by Gelbart/Jacquet as a consequence of the existence of the adjoint square lift. Of course due to the recent progress on the functoriality conjecture by Kim and Shahidi [24], [23] one can do much better: with $1/4$ replaced by $7/64$ (Kim/Sarnak).

However Kuznetsov's formula is powerful to control *linear combinations* of Fourier coefficients (rather than the individual ones), i.e., sums of the form

$$\sum_{|t_j|\leqslant T} \frac{1}{\cosh(\pi t_j)} \Big| \sum_{n\leqslant N} a_n \rho_j(n)\Big|^2.$$

Such sums are expressed in terms of Kloosterman sums

$$\sum_c \sum_{m,n} \overline{a_m} a_n \frac{\mathrm{Kl}(m,n;c)}{c} \varphi_T\Big(\frac{\sqrt{mn}}{c}\Big),$$

and depending on the case, one may either use Weil's bound or use the shape of Kloosterman sums to factor that expression further as

$$\sum_c \frac{1}{c} \sum_{\substack{a(c)\\(a,c)=1}} \Big(\sum_m \overline{a_m} e\Big(\frac{am}{c}\Big)\Big)\Big(\sum_n a_n e\Big(\frac{\overline{a}n}{c}\Big)\Big).$$

Combining this with the classical large sieve inequality one gets a general large sieve inequality for a linear combination of Fourier coefficients of modular forms (which for some applications is stronger than the Ramanujan/Petersson conjecture)

$$\sum_{|t_j|\leqslant T} \frac{1}{\cosh(\pi t_j)} \Big| \sum_{n\leqslant N} a_n \rho_j(n)\Big|^2 \ll \Big(T^2 + \frac{N}{q}\Big) \sum_{n\leqslant N} |a_n|^2.$$

8 Dimension of the space of modular forms of weight 1

Let $q > 1$ et $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$ an odd Dirichlet character. We denote by $S_1(q, \chi)$ the space of holomorphic forms of weight 1 level q and nebentypus χ, i.e., of the $f(z)$ such that

$$f\left(\frac{az+b}{cz+d}\right) = \chi(d)(cz+d)f(z), \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q).$$

Contrary to the case of forms of weight $k \geqslant 2$, there is no way to compute explicitly the dimension of this space.

In fact, there should be only very few of such forms. Serre conjectured that $\dim S_1(q, \chi) \ll_\varepsilon q^{1/2+\varepsilon}$. In this direction, the trace formula would give (at best) the upper bound $\dim S_1(q, \chi) \ll q/\log q$, but the first really non-trivial result is due to W. Duke [9] who proved that

$$\dim S_1(q, \chi) \ll_\varepsilon q^{1-1/12+\varepsilon}.$$

Duke's method was geometric in nature, in [30] the spectral theory of automorphic forms has been used to give a further improvement.

Theorem 8.1. *For any $\varepsilon > 0$ one has*

$$\dim S_1(q, \chi) \ll_\varepsilon q^{1-1/7+\varepsilon}.$$

In fact, since χ is of order divisible by 60, one even has

$$\dim \bigoplus_{\substack{\chi \bmod q \\ \chi \text{ odd}}} S_1(q, \chi) \ll_\varepsilon q^{1-1/7+\varepsilon}.$$

One of the key points of the proof, already used by Duke (and proved by Deligne/ Serre [6]), is that any such Hecke-eigenform, f say, is associated with a complex two-dimensional complex Galois representation ρ_f such that the Hecke-eigenvalues of f at a prime $p \nmid q$ equal

$$\lambda_f(p) = \operatorname{tr}(\rho_f \operatorname{Frob}_p), \quad \chi(p) = \det(\rho_f \operatorname{Frob}_p).$$

The outcome is that (since the (finite) images of 2-dimensional complex Galois representations are classified) there exist complex numbers a_2, a_8, a_{12} of modulus at most 1 (depending only on χ) such that for any $f \in S_1(q, \chi)$ one has the linear relation

$$a_{12}\rho_f(p^{12}) + a_8\rho_f(p^8) + a_2\rho_f(p^2) = \rho_f(1).$$

Hence

$$\sum_{f \in \mathcal{B}_1(q,\chi)} |\rho_f(1)|^2 |\{p \leqslant N,\ (p,q) = 1\}|^2 = \sum_{f \in \mathcal{B}_1(q,\chi)} \Big| \sum_{p \leqslant N} a_{12}\rho_f(p^{12}) + a_8\rho_f(p^8) + a_2\rho_f(p^2) \Big|^2.$$

There is of course no Petersson trace formula for $S_1(q,\chi)$ enabling to estimate the above sum, but we observe that if f belong to $S_1(q,\chi)$ then $y^{1/2}f(z)$ is a Maass form of weight 1 and eigenvalue $1/4$ for the Laplace operator

$$\Delta_1 = -y^2\Big(\frac{\partial^2}{\partial^2 x}+\frac{\partial^2}{\partial^2 y}\Big)+iy\frac{\partial}{\partial x}$$

of weight 1. Hence, we may embed the basis $\mathcal{B}_1(q,\chi)$ in an orthonormal basis of Maass form of weight 1. Then, by positivity, one can bound the sum above by the corresponding full spectral sum (including all Maass form and Eisenstein series). This is a priori wasteful but not too much thanks to Kuznetsov's formula and Weil's bound for the corresponding Kloosterman sums.

This way we obtain

$$\sum_{f\in\mathcal{B}_1(q,\chi)}|\rho_f(1)|^2\Big(\frac{N}{\log N}\Big)^2 \ll \frac{N}{\log N}+\frac{N^6}{q}\Big(\frac{N}{\log N}\Big)^2.$$

We conclude by choosing N optimally ($N=q^{1/7}$) and by observing that

$$\sum_{f\in\mathcal{B}_1(q,\chi)}|\rho_f(1)|^2 \gg \dim S_1(q,\chi)q^{\varepsilon-1}.$$

Note that in the above proof above we have used the linear relation (valid a priori only for Maass forms of Galois type)

$$a_{12}\rho_f(p^{12})+a_8\rho_f(p^8)+a_2\rho_f(p^2)=\rho_f(1)$$

to (coarsely) detect (and isolate) automorphic Maass forms of weight 1 of Galois type from all other Maass forms. Thus the above estimate can be interpreted as a very coarse and very primitive manifestation of Langlands "beyond endoscopy" program.

9 The shifted convolution problem

Given f and g two modular forms, the *shifted convolution problem* (SCP) consist in evaluating the sum

$$S(f,g;h)=\sum_{m\pm n=h}\rho_f(m)\rho_g(n)F\Big(\frac{m}{M},\frac{n}{N}\Big)$$

where $h\neq 0$, $F(x,y)$ is a test function and M,N are parameters going to $+\infty$.

Note that for $h=0$ the sum becomes

$$\sum_m \rho_f(m)\rho_g(m)F\Big(\frac{m}{M},\frac{m}{N}\Big)$$

and so is closely related to the Rankin/Selberg L-function of the pair (f,g).

Observe that if one takes for f and g the theta series associated to two definite positive binary quadratic forms $Q(x, y)$, $R(x, y)$, and if one chooses $\pm = +$ and for F the constant function 1, the SCP becomes the problem of evaluating the number of the representations of h by the quaternary quadratic form $Q(x, y) + R(x', y')$, which was Kloosterman's original motivation.

Kloosterman original way for solving this instance of the SCP was to use the circle method. Since

$$\int_{[0,1]} e(h\alpha)d\alpha = \delta_{h=0},$$

one has

$$S(f, g; h) = \int_{[0,1]} S(f, g; \alpha)e(-h\alpha)d\alpha$$

where

$$S(f, g; \alpha) = \sum_m \rho_f(m)e(m\alpha) \sum_n \rho_g(n)e(\pm n\alpha)F\left(\frac{m}{M}, \frac{n}{N}\right).$$

The essence of the circle method consists in replacing (possibly up to a good error term) the above integral over [0, 1] by a discrete weighted average of the function $S(f, g; \alpha)e(-h\alpha)$ over points α ranging over a subset of the set of rational numbers in [0, 1], with denominator bounded by some parameter C, $\{\frac{a}{c},\ (a, c) = 1,\ c \leqslant C\} \subset [0, 1]$. In the specific case of the SCP, one can then use the modular properties of f and g to evaluate the corresponding sums $S(f, g; \frac{a}{c})$, which makes eventually Kloosterman sums appear. In Kloosterman's original context this approximation step is called Kloosterman's refinement, but nowadays more flexible treatments are available (the δ-symbol method of Friedlander/Iwaniec or Jutila's method of overleaping intervals [11], [21])

The outcome is a sum of the following shape:

$$\frac{1}{L}\sum_c w\left(\frac{c}{C}\right) \sum_{\substack{a(c)\\(a,c)=1}} e\left(h\frac{a}{c}\right) \sum_m \rho_f(m)e\left(m\frac{a}{c}\right) \sum_n \rho_g(n)e\left(n\frac{a}{c}\right)G(m, n, h; c)$$

where w and G are compactly supported test functions and L is the total weight of the summation

$$L = \sum_c w\left(\frac{c}{C}\right)\varphi(c).$$

Now the modularity of f and g is exploited by means of the Voronoi summation formula which has the roughly shape

$$\sum_m \rho_f(m)e\left(m\frac{a}{c}\right)G(m) = \frac{1}{c}\sum_m \rho_f(m)e\left(-m\frac{\bar{a}}{c}\right)\widehat{G}(m),$$

where G is a test function and $\widehat{G}$ is a Bessel transform of G (which depends of the infinity type of g).

Applying such formula to f and g, and summing over $a \pmod c$ Kloosterman sums appear since the sum becomes

$$\frac{1}{L}\sum_c w\Big(\frac{c}{C}\Big)\sum_m \rho_f(m)\sum_n \rho_g(n)\frac{\mathrm{Kl}(m\pm n,h;c)}{c}\widehat{G}(m,n,h;c),$$

with possibly an extra principal main term (if f and g are not cuspidal). Now applying Weil's bound one deduces a non-trivial upper bound for (the error term of) the shifted convolution sum. This is essentially the path Kloosterman followed to evaluate the number of representations of an integer by a quaternary quadratic form.

9.1 Back to spectral theory

Nowadays one can get estimates which are stronger than the ones provided by Weil's bound. Indeed pursuing the transformation further, one can apply Kuznetsov's trace formula, but *backwards* (from Kloosterman sum to Fourier coefficients of automorphic forms) to get

$$\frac{1}{L}\sum_j \frac{1}{\cosh(\pi t_j)}\overline{\sigma}_j(h)\Big(\sum_{|h'|\leqslant H'} a_{f,g}(h')\rho_j(h')\Big)\widetilde{G}(h',h;t_j)$$
$$+\text{ holomorphic contribution} + \text{Eisenstein contribution}$$

where $a_{f,g}(h')$ is of the shape

$$a_{f,g}(h') = \sum_{\substack{m\pm n=h'\\ |m|,|n|\leqslant H'}} \rho_f(m)\rho_g(n).$$

Note that the coefficient $a_{f,g}(h')$ is again a shifted convolution sum; however, in practice the range of the variable H' is quite different from the initial range (shorter) and moreover the $a_{f,g}(h')$ need only to be bounded on average over h' (for instance by using the Large Sieve Inequality for Fourier coefficients of automorphic forms), so the argument is not circular. At this point we may take advantage of the fact that much better bounds are available for the Fourier coefficients $\rho_j(h)$ than the ones provided through Weil's bound:

$$|\rho_j(h')| \ll_\varepsilon |\rho_j(1)|n^{7/64+\varepsilon}.$$

There is a further advantage of this approach since on several occasions the SCP occurs with a extra averaging over the h variable. One may need to evaluate

$$\sum_{h\leqslant H}\chi(h)S(f,g;h)$$

where $\chi(h)$ are oscillating complex numbers. On these occasions, even the bound provided by the Ramanujan/Petersson conjecture for individual h may not be sufficient to bound the averaged shifted convolution sum properly. However going back to the

previous expression our averaged sum becomes

$$\frac{1}{L}\sum_j \frac{1}{\cosh(\pi t_j)}\Big(\sum_h \chi(h)\overline{\sigma}_j(h)\Big)\Big(\sum_{h'} a_{f,g}(h')\rho_j(h')\Big)\widetilde{G}(h',h;t_j)$$
$$+ \text{ holomorphic contribution} + \text{Eisenstein contribution}.$$

Then one can hope to get extra cancellation by estimating non-trivially the sum

$$\sum_h \chi(h)\overline{\sigma}_j(h)$$

for each j. Interesting examples of such coefficients are given by Dirichlet characters or Fourier coefficients of modular forms.

10 Application to the subconvexity problem

One of the most important recent applications of the SCP is the resolution of the subconvexity problem for GL_2-automorphic L-functions. Here we state the problem only for the conductor aspect.

Let π an automorphic representation of $\mathrm{GL}_n(\mathbb{A}_{\mathbb{Q}})$ of conductor q. We say that an L-function associated to π satisfies the *convexity bound* if

$$L(\pi, s) \ll_{\varepsilon,\pi_\infty,s} q^{1/4+\varepsilon}$$

for $\mathrm{Re}\, s = 1/2$. The subconvexity problem (ScP) consists in replacing the exponent $1/4$ by one strictly smaller. This has many applications, in particular to Linnik's equidistribution problems on modular or Shimura curves [8], [29].

The SCP has been the key for the resolution of many instances of the ScP (notably through the work of Duke/Friedlander/Iwaniec [10], [12], [14], [15]). Recently, in joint work with G. Harcos, we could solve the problem for Rankin/Selberg L-functions in great generality [19].

Theorem 10.1. *Let g be a fixed cusp form and let f be a cusp form of level q (say). Then, if the product of the nebentypus $\chi_f\chi_g$ is not trivial, one has*

$$L(f\otimes g, s) \ll q^{1/2-1/2700}.$$

This result is obtained by solving an averaged version of the SCP. An application of this is the equidistribution of short orbits of Heegner points of Shimura curves associated to indefinite quaternion algebras over $\mathbb{Q}$.

Observe that in the modern approaches to the SCP Kloosterman sums act merely as a catalyst: either Kuznetsov's trace formula or some transformations make them appear, but a backwards application of the Kuznetsov formula make them disappear afterwards. This suggests that Kloosterman sums could be avoided for some problems. This is indeed the case in the ScP. Recently A. Venkatesh [32] found a different

approach to the ScP which builds only on the realization of the L-function in terms of a (square of a) period. In particular his approach avoids entirely the use of Fourier coefficients and Kloosterman sums. In fact, his method, which is based on the ergodic properties of Hecke operators, works smoothly over an arbitrary number field. Amongst other cases, Venkatesh could solve the ScP for

- the standard L-function of a $\mathrm{GL}_2(\mathbb{A}_F)$-automorphic representation with trivial nebentypus (a large conductor),
- the Rankin/Selberg L-function of a pair of $\mathrm{GL}_2(\mathbb{A}_F)$-automorphic representations with trivial nebentypus, one being fixed, the other having large conductor.

More recently, Venkatesh and I were able to remove the assumptions on the nebentypus above. The proof differs from Venkatesh's proof in several aspects (in particular, it does not use – directly – the ergodicity of Hecke operators). In fact the principal extra ingredient was inspired by the resolution of the ScP for Rankin/Selberg L-functions over $\mathbb{Q}$ discussed above [31]. A consequence of these results is the resolution of the ScP for modular Artin L-functions over a number field, and the equidistribution of Galois orbits of special points on quaternionic Shimura varieties over totally real fields (which is meaningful in the context of the Andre/Oort conjectures).

11 Algebraic and modular aspect of Kloosterman sums combined

We conclude this survey with two further applications which combine both the modular and the algebraic aspects of Kloosterman sums discussed so far.

11.1 Number variance on the modular surface and the error term in Weyl's law

The main motivation of Selberg for developing the trace formula was to prove the existence of cusp forms. In particular for the full modular curve he obtained the Weyl law: for $T \geqslant 1$,

$$N(T) := |\{j \geqslant 1,\ |t_j| \leqslant T\}| = \mathrm{MainTerm}(T) + S(T);$$

here MainTerm is well understood and asymptotic to

$$\mathrm{MainTerm} \simeq \frac{\mathrm{Vol}\big(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}\big)}{4\pi} T^2.$$

On the other hand $S(T)$ is rather small error term ($= O(\frac{T}{\log T})$), but its asymptotic properties are not so well understood. Selberg also established a lower bound for the

variance of this error term showing that it is often not too small:

$$\int_T^{2T} |S(t)|^2\, dt \gg T^2/\log^2 T.$$

Until recently there has been no progress at all concerning the evaluation of this variance. However last year X. Li and P. Sarnak [27], by using the Petersson/Kuznetsov formula instead of Selberg's trace formula and the lower bound (along with other analytic techniques)

$$X\frac{\exp((\log\log X)^{5/17})}{\log X} \ll \sum_{c\leqslant X}\frac{|\mathrm{Kl}(1,1;c)|^2}{c},$$

were able to make the first (modest but meaningful) improvement over Selberg's lower bound, namely:

$$\int_T^{2T} |S(t)|^2\, dt \gg \frac{T^2}{\log^2 T}\exp((\log\log T)^{5/17}).$$

In fact, they provided further refinements of this estimate. These refinements can be interpreted in term of Quantum Chaos. While it is expected that the local scaled spacing distributions of the Laplace eigenvalues of a generic hyperbolic surface should be Gaussian, it is expected that for arithmetic surfaces this distribution should be Poissonian. The results of Li and Sarnak then show that the distribution in the case of the modular surface is definitely not Gaussian and give the first hint for a Poissonian behavior. As we have seen, this is the consequence of the arithmetical structure of the Kloosterman sums associated to congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

11.2 On the sign of Kloosterman sums

For this concluding application we return to the question raised in Section 5 on the existence of sign changes of Kloosterman sums. Here we shall use Kuznetsov's formula in the opposite direction (from automorphic forms to Kloosterman sums). Indeed by using his formula (and Roelcke's lower bound on the first eigenvalue λ_1 of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$), Kuznetsov proved that

$$\sum_{c\leqslant X}\frac{\mathrm{Kl}(1,1;c)}{c^{1/2}} \ll X^{1/2+1/6+\varepsilon}, \tag{1}$$

which is clearly non trivial by comparison with the bound provided by Weil's estimate. He gave the first non-trivial result towards Linnik's conjecture (that the sum above should be $\ll X^{1/2+\varepsilon}$); however until recently (as was remarked by Serre) it was not evident whether this estimate accounted for the existence of sign changes amongst the $\mathrm{Kl}(1,1;c)$ or for an extraordinary uniform smallness of Kloosterman sums. However thanks to the lower bounds of Section 6.1 one has

Proposition. *There are infinitely many integers c such that*

$$\mathrm{Kl}(1,1;c) > 0 \quad (\textit{resp.}\ \mathrm{Kl}(1,1;c) < 0).$$

In fact the number of such integers which are less than X is at least $\gg X/\log X$.

The next step consists in limiting the number of allowed prime factors of the moduli c above and thus to prove a horizontal type of result for Kloosterman sums with almost prime moduli. The following theorem resolves the first basic question concerning HST(k) [17]:

Theorem 11.1. *There exist infinitely many squarefree c (a positive proportion in fact) having at most 23 prime factors such that* $\mathrm{Kl}(1,1;c) > 0$ (*resp.* $\mathrm{Kl}(1,1;c) < 0$).

Sketch of Proof. Refining the lower bound of Theorem 6.1 one can prove that ($u_0 = 1/23.9$)

$$\sum_{\substack{c\leqslant X\\ p|c\Rightarrow p\geqslant X^{1/u_0}}} \mu^2(2c)\frac{|\mathrm{Kl}(1,1;c)|}{\sqrt{c}} \geqslant 0.166\frac{X}{\log X}.$$

Thus it is sufficient to show that

$$\Big|\sum_{\substack{c\leqslant X\\ p|c\Rightarrow p\geqslant X^{1/u_0}}} \mu^2(2c)\frac{\mathrm{Kl}(1,1;c)}{\sqrt{c}}\Big| \leqslant 0.1659999\frac{X}{\log X}. \tag{2}$$

This kind of estimate follows from sieve methods; here we use a variant of Selberg's upper bound sieve. Recall that the input in Selberg's sieve is a non-negative arithmetic function $(a_c)_{c\leqslant X}$, say, and that when it works, the sieve provides bounds of the shape

$$\sum_{\substack{c\leqslant X\\ p|c\Rightarrow p\geqslant X^{1/u}}} a_c \leqslant C(u)\frac{X}{\log X},$$

with $C(u) > 0$ a decreasing function of u.

In the present case (to force positivity) we need to sieve the two sequences

$$a_c^{\pm} = 2^{\omega(c)} \pm \frac{\mathrm{Kl}(1,1;c)}{\sqrt{c}}.$$

Moreover a necessary condition for the sieve to work is to control such sequences well in arithmetic progressions to large moduli. This provides good bounds for the sums

$$\sum_{c\equiv 0(q)} \frac{\mathrm{Kl}(1,1;c)}{\sqrt{c}}.$$

Such bounds can be obtained by means of Kuznetsov's formula for $\Gamma_0(q)$ and with the help of the large sieve inequality (for Maass forms) together with the

Luo/Rudnick/Sarnak lower bound for λ_1 (any bound strictly better that Selberg's $\lambda_1 > 3/16$ would be sufficient). This way we have good control for q up to size $X^{1/2-\varepsilon}$ (this is an analog of the Bombieri/Vinogradov theorem):

Proposition. *For any $\varepsilon > 0$, $Q \leqslant X^{1/2-\varepsilon}$ and any $B > 0$, one has*

$$\sum_{q \leqslant Q} \Big| \sum_{\substack{c \leqslant X \\ c \equiv 0(q)}} \frac{\mathrm{Kl}(1,1;c)}{\sqrt{c}} \Big| \ll_{\varepsilon, B} \frac{X}{\log^B X}.$$

Then, applying to this situation (a variant of) the sieve of Selberg, one deduces that

$$\sum_{\substack{c \leqslant X \\ p|c \Rightarrow p \geqslant X^{1/u_0}}} \mu^2(2c) \Big(2^{\omega}(c) \pm \frac{\mathrm{Kl}(1,1;c)}{\sqrt{c}} \Big) \leqslant \mathrm{MT}(u_0) \frac{X}{\log X} + 0.1659999 \frac{X}{\log X} \quad (3)$$

where $\mathrm{MT}(u_0)\frac{X}{\log X}$ is a main term which can be proven to be equal to

$$\mathrm{MT}(u_0) \frac{X}{\log X} = \sum_{\substack{c \leqslant X \\ p|c \Rightarrow p \geqslant X^{1/u_0}}} \mu^2(2c) 2^{\omega}(c) + O\Big(\frac{X}{\log^2 X} \Big).$$

Subtracting this contribution from (3) we obtain the desired estimate (2).

References

[1] E. Bombieri, J. B. Friedlander, and H. Iwaniec, Primes in arithmetic progressions to large moduli, *Acta Math.* 156 (1986), no. 3-4, 203–251.

[2] ——, Primes in arithmetic progressions to large moduli. III, *J. Amer. Math. Soc.* 2 (1989), no. 2, 215–224.

[3] A. R. Booker, A test for identifying Fourier coefficients of automorphic forms and application to Kloosterman sums, *Experiment. Math.* 9 (2000), no. 4, 571–581.

[4] P. Deligne, La conjecture de Weil. I, *Inst. Hautes Études Sci. Publ. Math.* (1974), no. 43, 273–307.

[5] ——, La conjecture de Weil. II, *Inst. Hautes Études Sci. Publ. Math.* (1980), no. 52, 137–252.

[6] P. Deligne and J.-P. Serre, Formes modulaires de poids 1, *Ann. Sci. École Norm. Sup.* (4) 7 (1974), 507–530 (1975).

[7] J.-M. Deshouillers and H. Iwaniec, Kloosterman sums and Fourier coefficients of cusp forms, *Invent. Math.* 70 (1982/83), no. 2, 219–288.

[8] W. Duke, Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.* 92 (1988), 73–90.

[9] ——, The dimension of the space of cusp forms of weight one, *Internat. Math. Res. Notices* (1995), no. 2, 99–109 (electronic).

[10] W. Duke, J. B. Friedlander, and H. Iwaniec, Bounds for automorphic L-functions, *Invent. Math.* 112 (1993), no. 1, 1–8.

[11] ——, A quadratic divisor problem, *Invent. Math.* 115 (1994), no. 2, 209–217.

[12] ——, Bounds for automorphic L-functions. II, *Invent. Math.* 115 (1994), no. 2, 219–239.

[13] ——, Equidistribution of roots of a quadratic congruence to prime moduli, *Ann. of Math.* (2) 141 (1995), no. 2, 423–441.

[14] ——, Bounds for automorphic L-functions. III, *Invent. Math.* 143 (2001), no. 2, 221–248.

[15] ——, The subconvexity problem for Artin L-functions, *Invent. Math.* 149 (2002), no. 3, 489–577.

[16] É. Fouvry, Autour du théorème de Bombieri-Vinogradov, *Acta Math.* 152 (1984), no. 3-4, 219–244.

[17] E. Fouvry and P. Michel, Sur le changement de signe des sommes de Kloosterman, *Annals of Math.*, to appear.

[18] ——, Sommes de modules de sommes d'exponentielles, *Pacific J. Math.* 209 (2003), no. 2, 261–288.

[19] G. Harcos and P. Michel, The subconvexity problem for Rankin–Selberg L-functions and equidistribution of Heegner points. II, *Invent. Math.* 163 (2006), no. 3, 581–655.

[20] D. R. Heath-Brown and S. J. Patterson, The distribution of Kummer sums at prime arguments, *J. Reine Angew. Math.* 310 (1979), 111–130.

[21] M. Jutila, A variant of the circle method, *Sieve methods, exponential sums, and their applications in number theory* (Cardiff, 1995), London Math. Soc. Lecture Note Ser. 237, Cambridge Univ. Press, Cambridge 1997, pp. 245–254.

[22] N. M. Katz, *Gauss sums, Kloosterman sums, and monodromy groups*, Ann. of Math. Stud. 116, Princeton University Press, Princeton, NJ 1988.

[23] H. H. Kim, Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2, *J. Amer. Math. Soc.* 16 (2003), no. 1, 139–183, with appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak.

[24] H. H. Kim and F. Shahidi, Functorial products for $\mathrm{GL}_2 \times \mathrm{GL}_3$ and the symmetric cube for GL_2, *Ann. of Math.* (2) 155 (2002), no. 3, 837–893, with an appendix by Colin J. Bushnell and Guy Henniart.

[25] H. Kloosterman, On the representations of numbers in the forms $ax^2 + by^2 + cz^2 + dt^2$, *Acta Math.* 49 (1926), 407–464.

[26] N. V. Kuznecov, The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums, *Mat. Sb.* (*N.S.*) 111 (153) (1980), no. 3, 334–383, 479 (Russian), English transl. in Math. USSR Sb. 39 (1981), 299–342.

[27] X. Li and P. Sarnak, On the number variance for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, *GAFA*, to appear.

[28] P. Michel, Autour de la conjecture de Sato-Tate pour les sommes de Kloosterman. I, *Invent. Math.* 121 (1995), no. 1, 61–78.

[29] ——, The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points, *Ann. of Math.* (2) 160 (2004), no. 1, 185–236.

[30] P. Michel and A. Venkatesh, On the dimension of the space of cusp forms associated to 2-dimensional complex Galois representations, *Int. Math. Res. Not.* (2002), no. 38, 2021–2027.

[31] ——, Periods, subconvexity and equidistribution, in preparation.

[32] A. Venkatesh, Sparse equidistribution problems, period bounds, and subconvexity, preprint 2005.

[33] A. Weil, On some exponential sums, *Proc. Nat. Acad. Sci. U. S. A.* 34 (1948), 204–207.

Introduction à la correspondance de Langlands locale

Ariane Mézard

Équipe d'arithmétique et de géométrie algébrique,
Bât. 425, Mathématique, Université Paris-Sud
91405 Orsay Cedex, France
email : ariane.mezard@math.u-psud.fr

Abstract. This paper is an introduction to the local Langlands correspondence in dimension 2. We define some representations of $\mathrm{GL}_2(\mathbb{Q}_p)$ and some Galois representations with specified properties. We present strategies to connect them. We conclude with a discussion on the expected correspondence for p-adic representations.

1 Introduction

L'objet de cet exposé introductif est de donner un aperçu de la correspondance de Langlands locale classique (ou ℓ-adique) et de présenter brièvement les objets mathématiques et les stratégies mis en œuvre pour les décrire. Dans le cas non classique (ou p-adique) nous présentons des résultats récents qui suggèrent une nouvelle correspondance de Langlands.

La correspondance de Langlands est une correspondance, autrement dit une bijection entre les classes d'isomorphismes de

$$\left\{\begin{array}{c}\text{certaines}\\ \text{représentations}\\ \text{de type « automorphe »}\end{array}\right\} \longleftrightarrow \left\{\begin{array}{c}\text{certaines}\\ \text{représentations}\\ \text{galoisiennes}\end{array}\right\}.$$

Nous commençons par définir ces deux notions (§2, §3) et nous introduisons une troisième notion (§2) qui est utile pour établir la correspondance dans le cas p-adique

$$\left\{\begin{array}{c}G\text{-espaces}\\ \text{de Banach}\\ \text{unitaires}\end{array}\right\}.$$

Puis nous évoquons la nature du lien qui les relie (§4). Nous nous limitons ici au cas des représentations de dimension deux car nous n'avons encore aucune idée dans le cas p-adique en dimension supérieure.

Les notes présentées ici reprennent et complètent les exposés présentés par C. Breuil et l'auteur lors des 70-ièmes rencontres entre physiciens théoriciens et mathématiciens. L'auteur remercie L. Nyssen et M.-F. Vignéras pour des remarques sur cet article.

2 Du corps des nombres rationnels aux représentations de $\mathrm{GL}_2(\mathbb{Q}_p)$

La structure du corps $\mathbb{Q}$ des nombres rationnels est bien connue :

$$x \in \mathbb{Q},\ x = 0 \text{ ou } x = \Pi_{p_i \text{ premier}} p_i^{\alpha_{p_i}}$$

où les $\alpha_{p_i} \in \mathbb{Z}$ sont presque tous nuls. Nous complètons $\mathbb{Q}$ de deux façons différentes :

- complétion topologique,
- complétion algébrique (voir §3).

Pour p un nombre premier, nous avons une valeur absolue sur $\mathbb{Q}$

$$|0|_p = 0 \text{ ou } |x|_p = p^{-\alpha_p}.$$

Cette valeur absolue définit une topologie sur $\mathbb{Q}$. Le complété de $\mathbb{Q}$ par rapport à cette valeur absolue est noté $\mathbb{Q}_p$. La valeur absolue $|\cdot|_p$ s'étend à $\mathbb{Q}_p$ et fait de $\mathbb{Q}_p$ un espace topologique. Le groupe $\mathrm{GL}_2(\mathbb{Q}_p)$ des matrices inversibles de dimension deux à valeurs dans $\mathbb{Q}_p$ est donc aussi un espace topologique. L'objet du paragraphe 2 est de définir des représentations de $\mathrm{GL}_2(\mathbb{Q}_p)$ sur des $\mathbb{Q}_\ell$-espaces vectoriels. Quand $\ell \neq p$, ces représentations sont dites ℓ-adiques et sont définies dans le paragraphe 2.1. Quand $\ell = p$, les représentations sont dites p-adiques et sont définies au paragraphe 2.2.

2.1 Représentations ℓ-adiques

Soit ℓ un nombre premier distinct de p. Soit E une extension finie de $\mathbb{Q}_\ell$. Une *représentation ℓ-adique* de $G = \mathrm{GL}_2(\mathbb{Q}_p)$ est un homomorphisme de groupes

$$\pi :\ \mathrm{GL}_2(\mathbb{Q}_p) \to \mathrm{Aut}_E\, V$$

de $\mathrm{GL}_2(\mathbb{Q}_p)$ dans les automorphismes E-linéaires de V où V est un espace vectoriel sur E (éventuellement de dimension infinie).

La représentation ℓ-adique est dite *lisse* si tous les vecteurs $v \in V$ sont lisses : le stabilisateur $\{g \in \mathrm{GL}_2(\mathbb{Q}_p), \pi(g) \cdot v = v\}$ est ouvert.

Elle est dite *entière*, s'il existe une norme invariante sur V :

$$\|\pi(g) \cdot v\| = \|v\| \quad \text{pour tout } g \in \mathrm{GL}_2(\mathbb{Q}_p) \text{ et } v \in V.$$

Elle est dite *admissible* si elle est lisse et si pour tout sous-groupe ouvert compact H de G, V^H est de dimension finie.

Remarquons qu'une représentation irréductible entière et lisse est admissible. Le premier objet auquel nous nous intéressons est alors

$$\left\{ \begin{array}{c} \text{représentations} \\ \text{irréductibles,} \\ \text{entières, lisses} \\ \text{de } \mathrm{GL}_2(\mathbb{Q}_p) \text{ sur } \overline{\mathbb{Q}}_\ell \end{array} \right\}.$$

En fait ce sont des représentations sur des extensions finies E de $\mathbb{Q}_\ell$, où le corps E est variable. Les extensions finies E sont toutes incluses dans une clôture algébrique $\overline{\mathbb{Q}}_\ell$ de $\mathbb{Q}_\ell$.

Nous pouvons à présent définir le troisième objet (voir §1) auquel nous nous intéressons. Un *G-espace de Banach unitaire* Π sur E est un espace de Banach muni d'une action E-linéaire de G telle que

$$G \to \Pi, \quad g \mapsto g \cdot v$$

soit continue pour tout $v \in \Pi$ et telle que la topologie sur Π soit donnée par une norme invariante. Le complété d'une représentation lisse entière de G sur E par rapport à une norme invariante donne un G-espace de Banach unitaire sur E. Le sous-espace des vecteurs lisses d'un G-espace de Banach unitaire est une représentation lisse entière de G. De cette façon, Vignéras a obtenu ([20])

Théorème 2.1 (Vignéras). *Les représentations entières lisses absolument irréductibles de* $\mathrm{GL}_2(\mathbb{Q}_p)$ *sur E sont en correspondance avec les* $\mathrm{GL}_2(\mathbb{Q}_p)$*-espaces de Banach unitaires topologiquement absolument irréductibles sur E.*

Remarque 2.2. Dans les années 40, le physicien Bargman ([1]) a classifié toutes les représentations admissibles de $\mathrm{SL}_2(\mathbb{R})$ (donc de $\mathrm{GL}_2(\mathbb{R})$) sur les espaces de Banach sur $\mathbb{C}$ (pas forcement unitaires). Suite à ces travaux, R. P. Langlands a établi la correspondance pour le corps $\mathbb{R}$ (dans les années 70) non seulement pour $\mathrm{GL}_n(\mathbb{R})$ mais aussi pour n'importe quel groupe de Lie réductif ([16]).

Remarque 2.3. Vignéras a également donné une description complète des $\mathrm{GL}_2(\mathbb{Q}_p)$-espaces de Banach unitaires dans le cas ℓ-adique ([20]).

Remarque 2.4. Le théorème 2.1 implique que toutes les normes invariantes d'une représentation entière lisse induisent la même topologie.

2.2 Représentations p-adiques

Nous supposons à présent que, dans les définitions précédentes (§2.1), $\ell = p$. L'analogue du théorème de Vignéras n'est plus vrai. Nous ne savons même plus déterminer l'ensemble des $\mathrm{GL}_2(\mathbb{Q}_p)$-espaces de Banach unitaires qui est probablement très gros. Nous mettons alors d'autres hypothèses sur les représentations de $\mathrm{GL}_2(\mathbb{Q}_p)$: nous

remplaçons « entière lisse » par « entière localement analytique et fortement admissible » (la notion de « fortement admissible » a été définie par Schneider et Teitelbaum ([19])). Pour définir les représentations localement analytiques, on remplace les vecteurs lisses par les vecteurs localement analytiques : $v \in \Pi$ est *localement analytique* si pour tout sous-groupe ouvert compact $H \subset \mathrm{GL}_2(\mathbb{Q}_p)$ suffisamment petit, l'action de H sur v est analytique.

Soit $\mathcal{O}_E$ l'anneau des entiers de E, soit B un espace de Banach unitaire et B^0 sa boule unité. Un espace de Banach unitaire B est *admissible* (au sens de Schneider et Teitelbaum) si pour tout sous-groupe ouvert compact $H \subset \mathrm{GL}_2(\mathbb{Q}_p)$, $\mathrm{Hom}_{\mathcal{O}_E}((B/B^0)^H, E/\mathcal{O}_E)$ est un $\mathcal{O}_E$-module de type fini ([18]). On peut alors définir une application

$$\left\{\begin{array}{c}\mathrm{GL}_2(\mathbb{Q}_p)\text{-espaces de Banach}\\ \text{unitaires, admissibles,}\\ \text{topologiquement absolument}\\ \text{irréductibles}\end{array}\right\} \longrightarrow \left\{\begin{array}{c}\text{représentations}\\ \text{absolument irréductibles,}\\ \text{entières, localement analytiques,}\\ \text{fortement admissibles}\\ \text{de } \mathrm{GL}_2(\mathbb{Q}_p) \text{ sur } \overline{\mathbb{Q}}_p\end{array}\right\}$$

$$\Pi \longmapsto \Pi^{\text{loc. an.}}.$$

Mais nous n'avons pas d'équivalence entre les représentations de $\mathrm{GL}_2(\mathbb{Q}_p)$ localement analytiques irréductibles fortement admissibles et les $\mathrm{GL}_2(\mathbb{Q}_p)$ espaces de Banach unitaires admissibles. Dans certains cas, il y a même une infinité de normes invariantes. Il s'agirait donc de préciser encore les définitions pour obtenir une bijection comme au paragraphe 2.1 entre de « bons espaces de Banach p-adiques » (notamment associés à des représentations galoisiennes) et de « bonnes représentations » localement analytiques.

3 Du corps des nombres rationnels aux représentations galoisiennes

Une autre façon d'étendre $\mathbb{Q}$ est de prendre la complétion algébrique $\overline{\mathbb{Q}} = \bigcup \mathbb{Q}(\alpha)$ où α décrit les racines de polynômes à coefficients dans $\mathbb{Q}$. Le corps $\overline{\mathbb{Q}}$ est très mystérieux. Pour l'étudier, nous étudions le groupe $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ des $\mathbb{Q}$-automorphismes de $\overline{\mathbb{Q}}$ que nous allons comparer à des groupes plus petits grâce aux représentations galoisiennes. Une *représentation galoisienne* est un homomorphisme continu :

$$\rho\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(E)$$

où E est, par exemple, une extension finie de $\mathbb{Q}_\ell$. Pour tout nombre premier p, en restreignant ρ à $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$, nous obtenons

$$\rho_p\colon \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \to \mathrm{GL}_2(E)$$

Nous imposons une condition technique sur ρ_p : ρ_p doit être « ϕ-semi-simple ».[1] Pour $\ell \neq p$, nous savons décrire toutes les représentations ρ_p (car nous contrôlons la ℓ-partie de $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$). Mais pour $\ell = p$, comme dans §2.2, il y a « trop » de représentations p-adiques. L'idée de Fontaine est de s'intéresser d'abord à celles qui apparaissent « naturellement » en géométrie : les représentations semi-stables.[2] Nous obtenons de telles représentations notamment

– en prenant la cohomologie étale d'une variété propre et lisse sur $\mathbb{Q}_p$ à réduction semi-stable ;

– en construisant la représentation p-adique associée à une forme modulaire pour $\Gamma_0(pN)$, $(p, N) = 1$ (voir §4).

Grâce à la théorie de Fontaine, nous contrôlons les représentations semi-stables. Elles sont, à isomorphisme près, paramétrées par un ou deux invariants : les poids de Hodge–Tate $(0, k-1)$ et l'invariant $\mathcal{L}$ qui est dans $\overline{\mathbb{Q}}_p$. Précisément,

Théorème 3.1 (Colmez–Fontaine). *Nous avons une équivalence de catégories*:

$$\left\{ \begin{array}{c} \textit{représentations semi-stables} \\ \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \to \mathrm{GL}_2(\mathbb{Q}_p) \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} (\varphi, N)\textit{-modules filtrés de dim } 2 \\ \textit{faiblement admissibles} \end{array} \right\}.$$

Les paramètres k, $\mathcal{L}$ déterminent la filtration des (φ, N)-modules faiblement admissibles. Il y a deux types de représentations semi-stables : les représentations cristallines et les représentations semi-stables, non cristallines. A une représentation semi-stable, non cristalline, irréductible est associé le (φ, N)-module D faiblement admissible suivant : D est un $\overline{\mathbb{Q}}_p$-espace vectoriel de dimension deux muni de deux opérateurs N (qui est nilpotent) et φ, et d'une filtration exhaustive, décroissante, séparée

$$\begin{cases} D = \overline{\mathbb{Q}}_p e_1 \oplus \overline{\mathbb{Q}}_p e_2, \\ N(e_1) = e_2,\ N(e_2) = 0, \\ \varphi(e_1) = \pi^k \mu e_1,\ \varphi(e_2) = \pi^{k-2} \mu e_2,\ \mu \in \overline{\mathbb{Z}}_p^*, \\ \mathrm{Fil}^i\, D = D,\ i \leq 0,\ \mathrm{Fil}^i\, D = 0,\ i \geq k, \\ \mathrm{Fil}^i\, D = \overline{\mathbb{Q}}_p(e_1 + \mathcal{L} e_2),\ 1 \leq i \leq k-1, \end{cases}$$

où π est un élément de $\overline{\mathbb{Z}}_p$ vérifiant $\pi^2 = p$.

[1] L'action du Frobenius sur le module de Fontaine associé à ρ_p est semi-simple.

[2] Fontaine a défini des anneaux de période B_{dR}, B_{st} munis de certaines structures dont une action de $G_{\mathbb{Q}_p} = \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. La représentation p-adique V est dite de de Rham si $\dim_{\mathbb{Q}_p}(B_{\mathrm{dR}} \otimes_{\mathbb{Q}_p} V)^{G_{\mathbb{Q}_p}} = \dim_{\mathbb{Q}_p} V$. La représentation p-adique V est dite semi-stable si $\dim_{\mathbb{Q}_p}(B_{\mathrm{st}} \otimes_{\mathbb{Q}_p} V)^{G_{\mathbb{Q}_p}} = \dim_{\mathbb{Q}_p} V$.

A une représentation cristalline irréductible est associé le (φ, N)-module D faiblement admissible suivant (ici le paramètre $\mathcal{L}$ n'intervient pas)

$$\begin{cases} D = \overline{\mathbb{Q}}_p e_1 \oplus \overline{\mathbb{Q}}_p e_2, \\ N(e_1) = N(e_2) = 0, \\ \varphi(e_1) = p^{k-1}\mu e_2,\ \varphi(e_2) = -e_1 + \nu e_2,\ \mu \in \overline{\mathbb{Z}}_p^*\nu \in \mathcal{M}_{\overline{\mathbb{Z}}_p}, \\ \mathrm{Fil}^i D = D, i \leq 0,\ \mathrm{Fil}^i D = 0,\ i \geq k, \\ \mathrm{Fil}^i D = \overline{\mathbb{Q}}_p e_1,\ 1 \leq i \leq k-1. \end{cases}$$

Expliquons maintenant le lien entre tous ces objets.

4 Correspondance de Langlands locale

4.1 Cas ℓ-adique

Une « correspondance de Langlands » est une bijection entre représentations galoisiennes et représentations de GL_n. La correspondance de Langlands locale ℓ-adique a été démontrée pour la dimension $n \geq 2$ par Harris et Taylor en 1998 ([13]). Henniart en a donné une preuve simplifiée en 2000 ([14]).[3] Elle établit la correspondance

$$\left\{ \begin{array}{c} \text{représentations admissibles,} \\ \text{irréductibles de } \mathrm{GL}_n(\mathbb{Q}_p) \text{ sur } \overline{\mathbb{Q}}_\ell \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \rho_\ell : W_p \to \mathrm{GL}_n(\overline{\mathbb{Q}}_\ell) \\ \text{irréductibles,} \\ \phi\text{-semi-simple} \end{array} \right\}$$

$$\pi_\ell \longleftrightarrow \rho_\ell$$

où W_p est le sous-groupe de Weil de $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. La démonstration réalise géométriquement la correspondance lorsque les représentations proviennent de formes modulaires. La méthode est globale au sens où nous traitons tous les nombres premiers en même temps en regardant

- les représentations de $\mathrm{GL}_2(\mathbb{A})$ où $\mathbb{A}$ est le produit restreint sous-anneau de $\mathbb{R} \times \Pi\mathbb{Q}_\ell$ (ces représentations globales sont dites automorphes),
- les représentations galoisiennes du groupe global $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Précisons un peu la stratégie de réalisation géométrique. Une *forme modulaire parabolique de poids* $k \geq 2$ sur $\Gamma_0(N) := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z}),\ c \equiv 0 \mod N\}$ est une fonction holomorphe

$$f : \boldsymbol{h} = \{z \in \mathbb{C}, \mathrm{Im}\, z > 0\} \to \mathbb{C}$$

[3] La correspondance de Langlands locale pour les corps locaux de caractéristique p a été démontrée par Laumon, Rapoport et Stuhler en 1993 ([17]). La correspondance de Langlands pour les corps de fonctions globaux a été démontré par Drinfeld en 1989 pour $n = 2$ ([9]) et par Lafforgue en 2002 pour $n > 2$ ([15]).

telle que

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z), \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$$

qui admet comme développement de Fourier une écriture $f(z) = \sum_{n\geq 1} a_n e^{2i\pi nz}$ (ce qui impose un certain nombre de conditions quand z tend vers $i\infty$ ou vers un élément de $\mathbb{Q}$). Supposons que k est pair. L'espace $S_k(N)$ de ces formes modulaires est un $\mathbb{C}$-espace vectoriel de dimension finie sur lequel agissent des opérateurs $\mathbb{C}$-linéaires T_q pour q premier appelés opérateurs de Hecke. Nous appelons forme *propre*, une forme modulaire vecteur propre de ces opérateurs. A une telle forme propre, nous pouvons associer une représentation ℓ-adique

$$\rho(f)\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_\ell)$$

et pour tout p premier, une représentation lisse irréductible $\pi_p(f)$ de $\mathrm{GL}_2(\mathbb{Q}_p)$ (sur $\overline{\mathbb{Q}}_\ell$) telle que si $p \neq \ell$, $\pi_p(f)$ peut se retrouver à partir de $\rho(f)_{|\,\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)}$ et vice-versa :

$$\pi_p(f) \longleftrightarrow \rho(f)|_{\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)}.$$

Cette correspondance se réalise géométriquement sur un groupe de cohomologie : par souci de simplification, nous supposons ici que $N = p^n$ est une puissance de p. Soit

$$Y(p^n)(\mathbb{C}) := \Gamma(p^n)\backslash \boldsymbol{h} \quad \text{où} \quad \Gamma(p^n) := \ker(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})).$$

Nous disposons d'une surjection

$$Y(p^{n+1})(\mathbb{C}) \to Y(p^n)(\mathbb{C})$$

et $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ agit sur $Y(p^n)(\mathbb{C})$ par $\overline{z} \mapsto \overline{qz}$ où z relève $\overline{z}$ dans $\boldsymbol{h}$ et $g \in \mathrm{SL}_2(\mathbb{Z})$ (cette action ne dépend pas des choix car $\Gamma(p^n)$ est distingué).

Considérons

$$\lim_{\to n} H^1_{\mathrm{sing}}(Y(p^n)(\mathbb{C}), \overline{\mathbb{Q}}_\ell)$$

où H^1_{sing} est la cohomologie singulière de l'espace topologique $\mathbb{Y}(p^n)(\mathbb{C})$ à coefficients dans $\overline{\mathbb{Q}}_\ell$. Nous pouvons montrer que cet espace est naturellement muni d'une action de $\mathrm{SL}_2(\mathbb{Q}_p)$. Par exemple l'action de $\mathrm{SL}_2(\mathbb{Z}_p)$ provient de l'action de $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ à chaque étage. Il est aussi muni d'une action continue de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ mais elle est beaucoup plus subtile : cela provient du fait que $Y(p^n)(\mathbb{C})$ correspond aux points complexes d'une variété algébrique définie sur $\mathbb{Q}$ et que nous pouvons dans ce cas remplacer $H^1_{\mathrm{sing}}(Y(p^n)(\mathbb{C}), \overline{\mathbb{Q}}_\ell)$ par la cohomologie étale ℓ-adique de Grothendieck de $Y(p^n)(\overline{\mathbb{Q}})$ sur laquelle il y a une action du groupe de Galois. En pratique, nous préfèrons une action de $\mathrm{GL}_2(\mathbb{Q}_p)$ à une action de $\mathrm{SL}_2(\mathbb{Q}_p)$ et nous remplaçons à chaque étage $Y(p^n)(\mathbb{C})$ par $|(\mathbb{Z}/p^n\mathbb{Z})^*|$ copies de $Y(p^n)(\mathbb{C})$ sur lesquelles nous faisons agir $(\mathbb{Z}/p^n\mathbb{Z})^*$ par permutation. Notons $\mathbb{Y}(p^n)(\mathbb{C})$ ce nouvel espace et cette fois, nous avons une action de $\mathrm{GL}_2(\mathbb{Q}_p) \times \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ sur

$$\lim_{\to n} H^1_{\mathrm{sing}}(\mathbb{Y}(p^n)(\mathbb{C}), \overline{\mathbb{Q}}_\ell).$$

Nous pouvons alors énoncer le théorème de réalisation géométrique de la correspondance de Langlands pour $k = 2$:

Théorème 4.1 (Deligne, Langlands, Carayol). *Nous avons un isomorphisme* $\mathrm{GL}_2(\mathbb{Q}_p) \times \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$*-équivariant*

$$\lim_{\to n} H^1_{\mathrm{sing}}(\mathbb{Y}(p^n)(\mathbb{C}), \overline{\mathbb{Q}}_\ell) \simeq \oplus' \pi_p(f) \otimes \rho(f) \oplus (\textit{reste})$$

où la somme $\oplus'$ *porte sur les formes modulaires paraboliques* $f \in S_2(p^*)$ *propres nouvelles et le terme* (*reste*) *est lié aux formes non paraboliques.*

La condition « forme nouvelle » dans le théorème évite de compter deux fois la même forme modulaire. Pour les poids $k > 2$, il faut remplacer les coefficients constants $\overline{\mathbb{Q}}_\ell$ par le système local issu du fibré

$$\Gamma(p^n)\backslash(\mathrm{Sym}^{k-2}\,\overline{\mathbb{Q}}_\ell^2 \times \boldsymbol{h}) \to \Gamma(p^n)\backslash\boldsymbol{h}$$

4.2 Cas *p*-adique

Dans le cas $\ell = p$, la théorie n'en est qu'à ses débuts. La correspondance globale précédente donne un candidat

$$\left\{ \begin{array}{c} \rho_p : \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p) \\ \text{irréductible, semi-stable,} \\ \phi\text{-semi-simple} \end{array} \right\} \longrightarrow \left\{ \begin{array}{c} \text{représentation admissible,} \\ \text{irréductible de } \mathrm{GL}_2(\mathbb{Q}_p) \text{ sur } \overline{\mathbb{Q}}_p \end{array} \right\}$$

$$\rho_p \longmapsto \pi_p.$$

D'après le théorème de Colmez–Fontaine, la donnée de ρ_p est équivalente à la donnée d'un (φ, N)-module filtré faiblement admissible. Nous nous apercevons que π_p correspond au (φ, N)-module sans sa filtration. L'application définie prédemment n'est donc pas inversible car nous avons « oublié » la filtration : pour retrouver ρ_p à partir de π_p, il manque les paramètres k, $\mathcal{L}$.

Nous disposons de la représentation algébrique $\mathrm{Sym}^{(k-2)}\,\overline{\mathbb{Q}}_p^2$ de $\mathrm{GL}_2(\mathbb{Q}_p)$ de plus haut poids $(0, k-2)$. Pour des raisons techniques, k est supposé pair et nous travaillerons plutôt avec $\underline{\mathrm{Sym}}^{k-2}\overline{\mathbb{Q}}_p^2 = \mathrm{Sym}^{k-2}\,\overline{\mathbb{Q}}_p^2 \otimes |\det|_p^{\frac{k-2}{2}}$ dont le caractère central est la puissance $(k-2)$-ième du caractère cyclotomique. Une première idée dûe à Breuil ([5]) pour se souvenir des poids de Hodge–Tate de ρ_p est donc de tensoriser

$$\pi_p \otimes \underline{\mathrm{Sym}}^{k-2}\overline{\mathbb{Q}}_p^2$$

(l'exposant $k-2$ provient du système local mentionné après le théorème 4.1 de réalisation géométrique de la correspondance). Pour se souvenir de l'invariant $\mathcal{L}$, nous introduisons de la topologie : l'idée de Breuil est d'associer à ρ_p un $\mathrm{GL}_2(\overline{\mathbb{Q}}_p)$-espace de Banach $B(k, \mathcal{L})$ dans le cas semi-stable non cristallin (resp. $B(k)$ dans le

cas cristallin) obtenu en complétant la représentation $\pi_p \otimes \underline{\mathrm{Sym}}^{k-2}\overline{\mathbb{Q}}_p^2$ par rapport à une norme qui dépend de $\mathcal{L}$. Idéalement, nous obtiendrions donc une correspondance

$$\left\{ \begin{array}{c} \rho_p\colon \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p) \\ \text{irréductible semi-stable} \\ \phi\text{-semi-simple} \end{array} \right\} \overset{?}{\longleftrightarrow} \left\{ \begin{array}{c} \mathrm{GL}_2(\mathbb{Q}_p)\text{-espace de Banach} \\ \text{admissible unitaire} \end{array} \right\}$$

$$\rho_p \overset{?}{\longleftrightarrow} B(k, \mathcal{L}) \text{ (resp. } B(k))$$

et cette correspondance admettrait une réalisation géométrique lorsque la représentation galoisienne provient d'une forme modulaire.

Nous commençons par indiquer la définition des espaces de Banach $B(k)$ et $B(k, \mathcal{L})$ ([4]). Pour cela, introduisons le développement de Mahler d'une fonction $f\colon \mathbb{Z}_p \to E$:

$$f(z) = \sum_{n=0}^{\infty} a_n(f)\binom{z}{n}, \quad \text{où } a_n(f) = \sum_{i=1}^{n} (-1)^i \binom{n}{i} f(n-i)$$

Soit $r \in \mathbb{R}_+$, une fonction $f\colon \mathbb{Q}_p \to E$ est dite de classe $\mathcal{C}^r$ si les coefficients de Mahler $a_n(f)$ de $f_{|\mathbb{Z}_p}$ satisfont

$$|a_n(f)|n^r \to_{n\to+\infty} 0.$$

Notons $B(k)$ le E-espace vectoriel des fonctions $\mathbb{Q}_p \to E$ de classe $\mathcal{C}^{\frac{k-2}{2}}$ tels que $x^{k-2}f(1/x)$ est aussi de classe $\mathcal{C}^{\frac{k-2}{2}}$. L'espace $B(k)$ est muni de l'action suivante de $\mathrm{GL}_2(\mathbb{Q}_p)$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot f = |ad-bc|^{\frac{k-2}{2}} (bx+d)^{k-2} f\left(\frac{ax+c}{bx+d}\right).$$

L'espace $B(k, \mathcal{L})$ est défini comme l'espace $B(k)$ modulo l'adhérence du sous-espace vectoriel engendré par les polynômes en z de degré $\leq k-2$ et les fonctions

$$\sum_{\text{finie}} (z-a_i)^{n_i} \log_{\mathcal{L}}(z-a_i), \quad \tfrac{k-2}{2} < n_i \leq k-2,\ a_i \in \mathbb{Q}_p,\ \deg \textstyle\sum (z-a_i)^{n_i} < \frac{k-2}{2}$$

($B(k, \mathcal{L})$ est toujours muni de l'action de $\mathrm{GL}_2(\mathbb{Q}_p)$). Expliquons rapidement quelle est l'origine de cette définition. En 2002, du côté galoisien pour k pair $4 \leq k \leq p-1$, Breuil et l'auteur ([6]) ont calculé les semi-simplifiées modulo p des représentations galoisiennes semi-stables non cristallines grâce à la théorie des modules de Breuil fortement divisibles. Ils ont obtenu différents cas suivant la position de la valuation de $\mathcal{L}$ par rapport à k. Les représentations dépendent de $a(\mathcal{L})$ et $b(\mathcal{L})$ où $a(\mathcal{L})$ et $b(\mathcal{L})$ sont donnés par les formules explicites suivantes

$$a(\mathcal{L}) := (-1)^{\frac{k}{2}} \left(-1 + \frac{k}{2}\left(\frac{k}{2}-1\right)\left(\mathcal{L} + 2H_{k/2-1}\right)\right),$$

$$b(\mathcal{L}) := (-1)^{\frac{k}{2}-\ell} \left(\frac{k}{2}-\ell\right) \binom{\frac{k}{2}-1-\ell}{-2\ell+1} \alpha \quad \text{si } \ell \in \left\{-\tfrac{k}{2}+2, -\tfrac{k}{2}+1, \ldots, -1\right\}$$

où $\ell := \mathrm{val}(\mathcal{L})$, $\alpha := \frac{\mathcal{L}}{p^\ell}$ et $H_0 := 0$, $H_n := \sum_{i=1}^{n} \frac{1}{i}$ désigne la série harmonique classique. De là est née la conviction de Breuil que ces formules ont une origine « automorphe » : nous devrions pouvoir les retrouver à partir d'une représentation p-adique de $\mathrm{GL}_2(\mathbb{Q}_p)$ réduite modulo p. C'est ainsi que Breuil a construit les candidats $B(k)$ et $B(k, \mathcal{L})$ ([4],[5]) et c'est sur ces espaces de fonctions que Berger, Breuil et Colmez ont mis en œuvre la démonstration de la conjecture de Breuil.

Conjecture (Breuil). *Si ρ_p est absolument irréductible alors $B(k, \mathcal{L})$ (resp. $B(k)$) est non nul, topologiquement irréductible, admissible, détermine ρ_p et ne dépend que de lui.*

Remarque 4.2. La conjecture de Breuil stipule qu'on peut retrouver la représentation ρ_p à partir de $B(k, \mathcal{L})$ (resp. $B(k)$). Dans les cas où cette conjecture est démontrée, les vecteurs localement algébriques de $B(k, \mathcal{L})$ (resp. $B(k)$) redonnent le (φ, N)- module associé à ρ_p sans la filtration, et le poids k

$$B(k, \mathcal{L})^{\text{loc. alg.}} \simeq \underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p$$

(la représentation π_p est lisse, la représentation $\underline{\mathrm{Sym}}^{k-2} E^2$ est algébrique). Par ailleurs, Breuil ([5]) établit que $B(k, \mathcal{L})$ est le complété de $\underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p$ par rapport à un $\mathcal{O}_E$-réseau stable par $\mathrm{GL}_2(\mathbb{Q}_p)$ dont la classe de commensurabiblité détermine la valeur de $\mathcal{L}$. L'espace de Banach $B(k)$ est le complété de $\underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p$ par rapport à un $\mathcal{O}_E$-réseau de type fini stable par $\mathrm{GL}_2(\mathbb{Q}_p)$.

Cette première version de la conjecture de Breuil a été démontrée en 2004. La stratégie initiée par Colmez ([8]) dans le cas semi-stable non cristallin et suivie par Berger et Breuil ([3]) dans le cas cristallin consiste à construire un modèle Borel-équivariant du Banach dual à $B(k, \mathcal{L})$ (resp. $B(k)$) grâce au (φ, Γ)-module associé à ρ_p. De ce modèle, ils déduisent les propriétés attendues.

Remarque 4.3. Si $4 \leq k \leq p-1$ les travaux de Breuil et l'auteur ([7]) redémontrent la correspondance p-adique et établissent sa compatibilité à la réduction modulo p : nous retrouvons la même dépendance par rapport à k, $\mathcal{L}$ du coté $\mathrm{GL}_2(\mathbb{Q}_p)$-espace de Banach que du côté galoisien.

Il s'agirait à présent d'étendre cette correspondance aux représentations associées à des formes modulaires et d'obtenir une réalisation géométrique de cette correspondance. Pour $n \in \{0, 1\}$, une représentation galoisienne associée à une forme modulaire nouvelle et propre de $S_k(\Gamma_0(p^n))$ est semi-stable. Mais pour $n \geq 2$, une telle représentation est seulement de de Rham. Nous recherchons donc une correspondance du type :

$$\left\{ \begin{array}{c} \rho_p : \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p) \\ \text{de Rham, } \phi\text{-semi-simple} \end{array} \right\} \overset{?}{\longleftrightarrow} \left\{ \begin{array}{c} \mathrm{GL}_2(\mathbb{Q}_p)\text{-espace de Banach} \\ \text{admissible unitaire} \end{array} \right\}$$

$$\rho_p \overset{?}{\longleftrightarrow} \hat{\pi}_p.$$

Breuil propose une réalisation géométrique de cette correspondance dans le cas des représentations galoisiennes associées à des formes modulaires ([4]). Supposons que la représentation $\rho_p(f)$ provient d'une forme modulaire parabolique f de $S_k(\Gamma_0(p^n))$ propre et nouvelle. Considérons $\widehat{H^1_{\text{sing}}}$, complété p-adique de

$$\lim_{\to n} H^1_{\text{sing}}(\mathbb{Y}(p^n)(\mathbb{C}), \mathbb{Z}_p)$$

qui est muni d'une action continue de $\mathrm{GL}_2(\mathbb{Q}_p) \times \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Il est unitaire, c'est-à-dire qu'il existe une norme invariante pour $\mathrm{GL}_2(\mathbb{Q}_p)$. Soit $\Pi_p(f)$ l'espace de Banach p-adique unitaire muni d'une action de $\mathrm{GL}_2(\mathbb{Q}_p)$

$$\Pi_p(f) = \mathrm{Hom}_{\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})}(\rho(f), \widehat{H^1_{\text{sing}}})$$

D'après [4],[10], on a toujours une injection $\mathrm{GL}_2(\mathbb{Q}_p)$-équivariante :

$$\underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p(f) \to \Pi_p(f)$$

En fait $\underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p(f)$ s'identifie aux vecteurs localement algébriques de $\Pi_p(f)$. Soit $\widehat{\pi_p(f)}$ l'adhérence de $\underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p(f)$ dans $\Pi_p(f)$. Ainsi $\widehat{\pi_p(f)}$ est la complétion de $\underline{\mathrm{Sym}}^{k-2} E^2 \otimes \pi_p(f)$ induite par la norme invariante sur $\widehat{H^1_{\text{sing}}}$.

Conjecture (Breuil [4]). *Si $\rho(f)_{|\,\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)} = \rho_p(f)$ est absolument irréductible alors $\widehat{\pi_p(f)}$ est non nul, topologiquement irréductible, admissible, détermine $\rho_p(f)$ et ne dépend que de lui.*

Remarque 4.4. Lorsque $\rho_p(f)$ est réductible, $\widehat{\pi_p(f)}$ ne suffit pas toujours à déterminer $\rho_p(f)$. Breuil propose alors un autre candidat : c'est un sous-espace de Banach topologiquement réductible de $\widehat{H^1_{\text{sing}}}$ qui détermine $\mathcal{L}(f)$ (l'invariant $\mathcal{L}$ associé à $\rho_p(f)$), qui ne dépend que de $\rho_p(f)$ et dont un sous-objet est la composante isotypique $\widehat{\pi_p(f)}$ associée à $\rho_p(f)$ ([4]).

Remarque 4.5. Lorsque $\rho_p(f)$ est absolument irreductible, Breuil conjecture également que $\Pi_p(f) = \widehat{\pi_p(f)}$.

La conjecture de Breuil est démontrée dans les cas suivants : le cas d'une forme f nouvelle propre de $S_k(\Gamma_0(p))$ (alors $\rho_p(f)$ est semi-stable non cristalline et $\pi_p(f)$ est dite série spéciale) est établi : en utilisant les symboles modulaires, Breuil montre qu'il existe un morphisme non nul unique de $B(k, \mathcal{L})$ dans $\widehat{\pi}_p(f)$ si et seulement si $\mathcal{L} = \mathcal{L}(f)$ ([4]). Autrement dit, l'espace de Banach $B(k, \mathcal{L})$ défini localement est celui induit par la cohomologie globale pour le « bon » paramètre $\mathcal{L}(f)$. Colmez montre que $B(k, \mathcal{L})$ est admissible et topologiquement irréductible pour tout $\mathcal{L}$ ([8]). Ces deux résultats induisent alors le fait que le morphisme $B(k, \mathcal{L}(f)) \to \widehat{\pi_p(f)}$ est un isomorphisme.

Le cas d'une forme f nouvelle propre de $S_k(\Gamma_0(1))$ (alors $\rho_p(f)$ est cristalline et $\pi_p(f)$ est dite série principale) a été traité par Berger et Breuil ([3]). Dans ce cas, il n'y a qu'une complétion unitaire locale $B(k)$ de $\underline{\mathrm{Sym}}^{k-2}E^2 \otimes \pi_p$ donc forcement induite par la cohomologie globale.

Le cas d'une forme f nouvelle propre de $S_k(\Gamma_0(p^n))$ avec $n \geq 2$ (alors $\pi_p(f)$ est dite supercuspidale) n'a pas encore été traité.

Références

[1] V. Bargman, Irreducible unitary Representations of the Lorentz group, *Ann. Math.* 48 (1947), 568–640.

[2] L. Berger and C. Breuil, Towards a p-adic Langlands programme, cours au C.M.S. de Hangzhou, Chine, 2004.

[3] —, Représentations cristallines irréductibles de $\mathrm{GL}_2(\mathbb{Q}_p)$, prépublication I.H.E.S., 2004.

[4] C. Breuil, Séries spéciales p-adiques et cohomologie étale complétée, prépublication I.H.E.S., 2003.

[5] —, Invariant $\mathcal{L}$ et série spéciale p-adique, *Ann. Scient. Éc. Norm. Sup.* 37 (2004), 559–610.

[6] C. Breuil and A. Mézard, Multiplicités modulaires et représentations de $\mathrm{GL}_2(\mathbb{Z}_p)$ et de $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ en $\ell = p$, *Duke Math. J.* 115 (2002), 205–310.

[7] —, Représentations semi-stables de $GL_2(\mathbb{Q}_p)$, demi-plan p-adique et réduction modulo p, en préparation.

[8] P. Colmez, Une correspondance de Langlands locale p-adique pour les représentations semi-stables de dimension 2, prépublication, 2004.

[9] V. Drinfeld, Cohomology of compactified modules of F-sheaves of rank 2, *J. Soviet Math.* 46 (1989), 1789–1821.

[10] M. Emerton, p-adic L-functions and unitary completions of representations of p-adic reductive groups, prépublication, 2004.

[11] J.-M. Fontaine, Représentations p-adiques des corps locaux. I, dans *The Grothendieck Festschrift*, vol. II, Progr. Math. 87, Birkhäuser Verlag, Boston 1990, 249–309.

[12] J.-M. Fontaine and P. Colmez, Construction des représentations p-adiques semi-stables, *Invent. Math.* 140 (2000), 1–43.

[13] M. Harris and R. Taylor, *On the geometry and cohomology of some simple Shimura varieties*, Ann. of Math. Studies 151, Princeton Univ. Press, Princeton 2001.

[14] G. Henniart, Une preuve simple des conjectures de Langlands locales pour GL_n sur un coprs p-adique, *Invent. Math.* 139, (2000), 439–455.

[15] L. Lafforgue, Chtoucas de Drinfeld et correspondance de Langlands, *Invent. Math.* 147 (2002), 1–241.

[16] R. P. Langlands, On the classification of irreducible representations of real algebraic groups, dans *Representations Theory and Harmonic Analysis on Semi-simple Groups*, P. Sally et D. Vogan eds., Math. Surveys Monogr. 31, Amer. Math. Soc., Providence, RI, 1989, 101–170.

[17] G. Laumon, M. Rapoport, and U. Stuhler, D-elliptiv sheaves and the Langlands correspondance, *Invent. Math.* 113 (1993), 217–338.

[18] P. Schneider and J. Teitelbaum, Banach space representations and Iwasawa theory, *Israel J. Math.* 127 (2002), 359–380.

[19] —, Locally analytic distributions and p-adic representations theory, with applications to GL_2, *J. Amer. Math. Soc.* 15 (2002), 443–468.

[20] M.-F. Vignéras, ℓ-adic Banach continuous representations of reductive p-adic groups when $\ell \neq p$, prépublication, 2005.