Astérisque

PETER SARNAK

Equidistribution and primes

Astérisque, tome 322 (2008), p. 225-240

<http://www.numdam.org/item?id=AST_2008__322__225_0>

Astérisque
322, 2008, p. 225–240

EQUIDISTRIBUTION AND PRIMES

by

Peter Sarnak

*To Jean Pierre Bourguignon**

Abstract. — We begin by reviewing various classical problems concerning the existence of primes or numbers with few prime factors as well as some of the key developments towards resolving these long standing questions. Then we put the theory in a natural and general geometric context of actions on affine n-space and indicate what can be established there. The methods used to develop a combinational sieve in this context involve automorphic forms, expander graphs and unexpectedly arithmetic combinatorics. Applications to classical problems such as the divisibility of the areas of Pythagorean triangles and of the curvatures of the circles in an integral Apollonian packing, are given.

***Résumé* (Équidistribution des nombres premiers).** — Nous commençons par l'examen de divers problèmes classiques concernant l'existence de nombres premiers ou de nombres avec peu de facteurs premiers, ainsi que quelques-uns des développéments clés vers la résolution de ces questions posées il y a bien longtemps. Ensuite, nous plaçons la théorie dans un contexte géométrique naturel et général d'actions sur le n-espace affine et nous indiquons ce qui peut être établi dans ce contexte. Les méthodes utilisées pour développer un crible combinatoire dans ce contexte impliquent les formes automorphes, les graphes d'expansion et, de manière inattendue, les combinatoires arithmétiques. Nous fournissons des applications aux problèmes classiques, tels que la divisibilité des aires des triangles pythagoriens et les courbures des cercles dans un paquetage apollonien entier.

I have chosen to talk on this topic because I believe it has a wide appeal and also there have been some interesting developments in recent years on some of these classical problems. The questions that we discuss are generalizations of the twin prime conjecture; that there are infinitely many primes p such that $p+2$ is also a prime. I

2000 Mathematics Subject Classification. — 11Axx, 20Gxx.
Key words and phrases. — Primes, sieves, affine orbits, saturation numbers, expanders and sum-product.

* This is an expanded version of the lecture that I had intended to give at the conference honoring Bourguignon on the occasion of his 60^{th} birthday and which I gave at the Pacific Institute of Mathematical Sciences in 2007.

am not sure who first asked this question but it is ancient and it is a question that occurs to anyone who looks, even superficially, at a list of the first few primes. Like Fermat's Last Theorem there appears to be nothing fundamental about this problem. We ask it simply out of curiosity. On the other hand the techniques, theories and generalizations that have been developed in order to understand such problems are perhaps more fundamental.

Dirichlet's Theorem. — In many ways this theorem is still the center piece of the subject. Like many landmark papers in mathematics, Dirichlet's paper proving the theorem below, initiated a number of fields: abelian groups and their characters, L-functions, class number formulae... The theorem asserts that an arithmetic progression $c, c+q, c+2q, c+3q, \ldots$ contains infinitely primes if and only if there is no obvious congruence obstruction. An obvious such obstruction would be say that c and q are both even or more generally that the greatest common divisor (c,q) of c and q is bigger than 1. Stated somewhat differently, let $L \neq 0$ be a subgroup of $\mathbb{Z}$, so $L = q\mathbb{Z}$ for some $q \geq 1$, and let $\mathcal{O} = c + L$ be the corresponding orbit of c under L, then $\mathcal{O}$ contains infinitely many primes iff $(c,q) = 1$ (strictly speaking this statement is slightly weaker since Dirichlet considers one-sided progressions and here and elsewhere we allow negative numbers and call $-p$ a prime if p is a positive prime).

Initial Generalizations. — There are at least two well known generalizations of Dirichlet's Theorem that have been investigated. The first is the generalization of his L-functions to ones associated with general automorphic forms on linear groups. This topic is one of the central themes of modern number theory but other than pointing out that these are used indirectly in proving some of the results mentioned below, I will not discuss them in this lecture. The second generalization is to consider other polynomials besides linear ones. Let $f \in \mathbb{Z}[x]$ be a polynomial with integer coefficients and let $\mathcal{O} = c + L$ as above. Does $f(\mathcal{O})$ contain infinitely many primes? For example if $\mathcal{O} = \mathbb{Z}$; is $f(x) = x^2 + 1$ a prime number for infinitely many x (a question going back at least to Euler). Is $f(x) = x(x+2)$ a product of two primes infinitely often? (this is a reformulation of the twin prime question). Neither of these questions have been answered and the answer to both is surely, yes. We will mention later what progress has been made towards them. In his interesting and provocative article "Logical Dreams" [**35**], Shelah puts forth the dream, that this question of Euler "cannot be decided". This is rather far fetched but for the more general questions about primes and saturation on very sparse orbits associated with tori that are discussed below, such a possibility should be taken seriously. We turn first in the next paragraph to several variables, that being the setting in which some problems of this type have been resolved.

Two Variables. — Let $\mathcal{O} = \mathbb{Z}^2$ and let f be a nonconstant polynomial in $\mathbb{Z}[x_1, x_2]$. If f is irreducible in $\mathbb{Q}[x_1, x_2]$ and the greatest common divisor of the numbers $f(x)$ with $x \in \mathcal{O}$ is 1, then it is conjectured that f takes on infinitely many prime values. In this higher dimensional setting we have found it more intrinsic and natural from many points of view to ask for more. That is the set of $x \in \mathcal{O}$ at which $f(x)$ is prime should not only produce an infinite set of primes for the values $f(x)$ but these (infinitely) many points should not satisfy any nontrivial algebraic relation. In the language of algebraic geometry, these points should be Zariski dense in the affine plane A^2. The Zariski topology on affine n-space A^n is gotten by declaring the closed sets to be the zero sets (over $\mathbb{C}$) of a system of polynomial equations. Thus a subset S of A^n is Zariski dense in A^n iff S is not contained in the zero set of a nontrivial polynomial $g(x_1, \ldots, x_n)$. In A^1 a set is the zero set of a nontrivial polynomial iff the set is finite. So the Zariski dense subsets of A^1 are simply the infinite sets. We denote the operation of taking the Zariski closure of a set in A^n by Zcl.

All the approaches to the conjecture that we are discussing involve giving lower bounds for the number of points in finite subsets of $\mathcal{O}$ at which $f(x)$ is prime. Usually one defines these sets by ordering by size of the numbers (so a big box in A^2) but in some variations of these problems that I discuss later quite different orderings are employed. A measure of the quality of the process is whether in the end the lower bound is strong enough to ensure the Zariski density of the points produced. As far as the conjecture that under the assumptions on f at the beginning of (4), the set of $x \in \mathcal{O}$ at which $f(x)$ is prime, is Zariski dense in A^2, the following is known:

(i) For f linear it follows from Dirichlet's theorem.

(ii) For f of degree two and f non-degenerate (in the sense of not reducing to a polynomial in one variable) it follows from Iwaniec [**23**]. His method uses the combinatorial sieve which we will discuss a bit further on, as well as the Bombieri-A. Vinogradov theorem which is a sharp quantitative version of Dirichlet's theorem (when counting primes p of size at most x and which are congruent to varying c modulo q, with q as large as $x^{1/2}$).

(iii) A striking breakthrough was made by Friedlander and Iwaniec [**10**]. It follows from their main result that the conjecture is true for $f(x_1, x_2) = x_1^2 + x_2^4$. They exploit the structure of this form in that it can be approached by examining primes $\alpha = a + b\sqrt{-1}$ in $\mathbb{Z}[\sqrt{-1}]$ with $b = z^2$. This was followed by work of Heath-Brown and the results in [**19**] imply that the conjecture is true for any homogeneous binary cubic form. They exploit a similar structure, in that such an $f(x_1, x_2)$ is of the form $N(x_1, x_2, 0)$ where $N(x_1, x_2, x_3)$ is the norm form of cubic extension of $\mathbb{Q}$, so that the problem is to produce prime ideals in the latter with one coordinate set to 0.

(iv) If $f(x_1, x_2)$ is reducible then we seek a Zariski dense set of points $x \in \mathbb{Z}^2$ at which $f(x)$ has as few as possible prime factors. For polynomials f of the special form $f(x) = f_1(x)f_2(x)\cdots f_t(x)$, with $f_j(x) = x_1 + g_j(x_2)$ where $g_j \in \mathbb{Z}[x]$ and $g_j(0) = 0$, it follows from the results in the recent paper of Tao and Ziegler [**36**] that the set of $x \in \mathbb{Z}^2$ at which $f(x)$ is a product of t primes, is Zariski dense in A^2. Equivalently the set of x at which $f_1(x), \ldots, f_t(x)$ are simultaneously prime, is dense. This impressive result is based on the breakthrough in Green and Tao [**16**] in particular their transference principle, which is a tool for replacing sets of positive density in the usual setting of Szémeredi type theorems with a set of positive density in the primes. The corresponding positive density theorem is that of Bergelson and Leibman [**2**]. Note that for these f_j's there is no local obstruction to $x_1 + g_j(x_2)$ being simultaneously prime since for a given $q \geq 1$ we can choose $x_1 \equiv 1(q)$ and $x_2 \equiv 0(q)$ $(g_j(0) = 0)$. Apparently this is a feature of these positive density Szemerédi type theorems in that they don't allow for congruence obstructions. (*) The above theorem with $g_j(x_2) = (j-1)x_2$, $j = 1, \ldots, t$ recovers the Green-Tao theorem, that the primes contain arbitrary long arithmetic progressions. From our point of view in paragraph (8) the amusing difference between the "existence of primes in an arithmetic progression" and that of an "arithmetic progression in the primes", will be minimized as they both fall under the same umbrella.

Hardy-Littlewood n-tuple Conjecture. — This is concerned with $\mathbb{Z}^n$ and subgroups L of $\mathbb{Z}^n$ acting by translations. If L is such a group denote by $r(L)$ its rank. We assume $L \neq 0$ so that $1 \leq r \leq n$ and also that for each j the coordinate function x_j restricted to L is not identically zero. For $c \in \mathbb{Z}^n$ and $\mathcal{O} = c + L$ the conjecture is about finding points in $\mathcal{O}$ all of whose coordinates are simultaneously prime. We state it as the following local to global conjecture:

HLC. — *If $\mathcal{O} = c + L$ as above then the set of $x = (x_1, \ldots, x_n) \in \mathcal{O}$ for which the x_j's are simultaneously prime, is Zariski dense in $Zcl(\mathcal{O})$ iff for each $q \geq 1$ there is an $x \in \mathcal{O}$ such that $x_1 x_2 \ldots x_n \in (\mathbb{Z}/q\mathbb{Z})^*$.*

Note that the condition on q, which is obviously necessary for the Zariski density, involves only finitely many q (for each given $\mathcal{O}$). Also to be more accurate, the conjecture in [**18**] concerns only the case of $r(L) = 1$ (which in fact implies the general case). In this case $Zcl(\mathcal{O})$ is a line and the conjecture asserts that there are infinitely many points in $x \in \mathcal{O}$ for which the n-tuples $(x_1, x_2, \ldots x_n)$ are all prime

(*) Though the paper "Intersective polynomials and polynomial Szemerédi theorem" by V. Bergelson, A. Leibman and E. Lesigne posted on ArXiv Oct/25/07, begins to address this issue.

iff there is no local obstruction. We observe that for any r the $Zcl(\mathcal{O})$ is simply a translate of a linear subspace.

The main breakthrough on the HLC as stated above is due to I. Vinogradov (1937) in his proof of his celebrated "ternary Goldbach theorem", that every sufficiently large positive odd number is a sum of 3 positive prime numbers. His approach was based on Hardy and Littlewood's circle method, a novel sieve and the technique of bilinear estimates, see Vaughan [**37**]. It can be used to prove HLC for a non-degenerate L in $\mathbb{Z}^3$ of rank at least 2. Special cases of HLC in higher dimensions are established by Balog in [**1**] and recently Green and Tao [**15**] made a striking advance. Their result implies HLC for $L \leq \mathbb{Z}^4$ and $r(L) \geq 2$ and L non-degenerate in a suitable sense. Their approach combines Vinogradov's methods with their transference principle. It makes crucial use of Gowers' techniques from his proof of Szemerédi's theorem, and it has close analogies with the ergodic theoretic proofs of Szémeredi's theorem due to Furstenberg and in particular the work of Host and Kra [**22**]. These ideas have potential to establish HLC for $L \leq \mathbb{Z}^n$ of rank at least two (and non-degenerate), which would be quite remarkable.

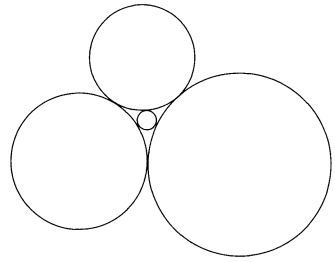
Pythagorean Triples. — We turn to examples of orbits $\mathcal{O}$ in $\mathbb{Z}^n$ of groups acting by matrix multiplication rather than by translations (i.e. addition). By a Pythagorean triple we mean a point $x \in \mathbb{Z}^3$ lying on the affine cone C given as $\{x : F(x) = x_1^2 + x_2^2 - x_3^2 = 0\}$ and for which $gcd(x_1, x_2, x_3) = 1$. We are allowing x_j to be negative though in this example we could stick to all $x_j > 0$, so that such triples correspond exactly to primitive integral right triangles. Let O_F denote the orthogonal group of F, that is the set of 3×3 matrices g for which $F(xg) = F(x)$ for all x. Let $O_F(\mathbb{Z})$ be the group of all such transformations with entries in $\mathbb{Z}$. Some elements of $O_F(\mathbb{Z})$ are

$$A_1 = \begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & -2 \\ 2 & 2 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 3 \end{bmatrix}, A_3 = \begin{bmatrix} -1 & -2 & -2 \\ 2 & 1 & 2 \\ 2 & 2 & 3 \end{bmatrix}.$$

In fact $O_F(\mathbb{Z})$ is generated by A_1, A_2 and A_3. It is a big group and one can show that the set of all Pythagorean triples P is the orbit of $(3,4,5)$ under $O_F(\mathbb{Z})$, i.e. $P = (3,4,5)O_F(\mathbb{Z})$. Following the lead of Dirichlet, let L be a subgroup of $O_F(\mathbb{Z})$ and let $\mathcal{O} = (3,4,5)L$ be the corresponding orbit of Pythagorean triples. The area $A(x) = x_1x_2/2$ of the corresponding triangle is in $\mathbb{Q}[x_1, x_2, x_3]$. We seek triangles in $\mathcal{O}$ for which the area has few prime factors. What is the minimal divisibility of the areas of a Zariski dense (in $Zcl(\mathcal{O})$, which for us will be equal to C) set of triples in $\mathcal{O}$? We return to this later on. As a side comment, a similar problem asks which numbers are the square free parts of the areas of Pythagorean triangles in P? This is the ancient

"congruent number problem" about which much has been written especially because of its connection to the question of the ranks of a certain family of elliptic curves. Heegner [**20**] using his precious method for producing rational points on elliptic curves shows that any prime $p \equiv 5$ or $7 \bmod 8$ is a congruent number. For a given such p the set of triangles realizing p is very sparse but never-the-less is Zariski dense in C. Via the same relation the congruent number problem is connected to automorphic L-functions through the Birch and Swinnerton-Dyer Conjecture (see [**38**]).

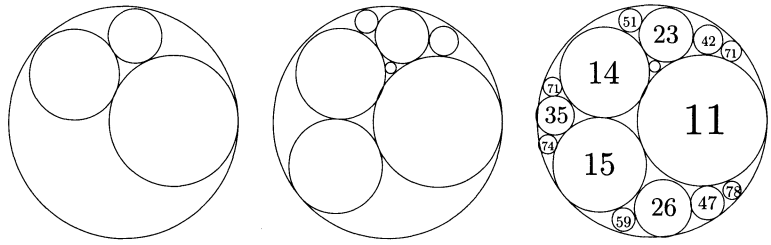
Integral Apollonian Packings. — As a final example before putting forth the general theory we discuss some Diophantine aspects of integral Apollonian packings. Descartes is well known among other things for his describing various geometric facts in terms of his Cartesian coordinates. One such example is the following relation between four mutually tangent circles:



If the radius of the j^{th} circle is R_j then its curvature a_j is equal to $1/R_j$, $j = 1, 2, 3, 4$. The relation is that

$$F(a_1, a_2, a_3, a_4) := 2\left(a_1^2 + a_2^2 + a_3^2 + a_4^2\right) - \left(a_1 + a_2 + a_3 + a_4\right)^2 = 0\,.$$

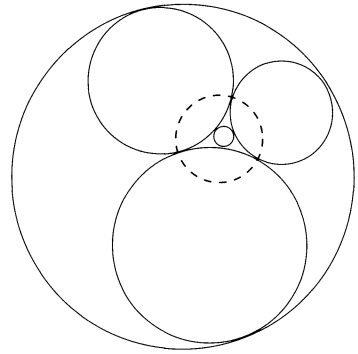
Consider now an Apollonian packing which is defined as follows; starting with 4 tangent circles of the first generation in Figure 2 (in this configuration the outer circle has all the other circles in its interior so by convention its curvature is $-1/R$ where R is its radius).



Génération 1 Génération 2 $a = (-6, 11, 14, 23)$

Now place a circle in each of the 4 lune regions in generation 1 so that these are tangent to the three circles that bound the lune. The placement is possible and is unique according to a theorem of Apollonius. At generation 2, there are now 12 new

lunes and we repeat the process ad-infinitum. The resulting packing by circles is called an Apollonian packing. The complement of all the open disks in the packing is a closed fractal set whose Hausdorff dimension δ is approximately 1.30. Boyd [**7**] has shown that if $N(T)$ is the number circles in the packing whose curvature is at most T, then $\log N(T)/\log T \to \delta$ as $T \longrightarrow \infty$. The interesting Diophantine fact is that if the initial 4 circles have integral curvatures then so do all the rest of the circles in the packing. This is apparent in the example in Figure 2 where the initial 4 circles have curvatures (-6, 11, 14, 23) and where the curvatures of each circle is displayed in the circle. It is customary in any lecture to offer at least one proof. Ours is the demonstration of this integrality of curvatures.



In this figure the inversion S in the dotted circle, which is the unique circle orthogonal to the inside circles, takes the outermost circle to the innermost one and fixes the other three. It takes the 4-tuple (a_1, a_2, a_3, a_4) representing the curvatures of the 4 outer circles to (a_1', a_2, a_3, a_4) where a_1' is the curvature of the inner most circle. From the Descartes relation it follows that a_1 and a_1' are roots of the same quadratic equation and a simple calculation yields that $a_1' = -a_1 + 2a_2 + 2a_3 + 2a_4$. This inversion is also the step which places the circle in the corresponding lune, that being a single step in the Apollonian packing. It follows that if the 4×4 integral involutions S_1, S_2, S_3, S_4 are given by

$$S_1 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}, S_2 = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}, S_3 = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}, S_4 = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

and A is the group generated by S_1, S_2, S_3, S_4 then the orbit $\mathcal{O} = (a_1, a_2, a_3, a_4)A$, corresponds precisely to the configurations of 4 mutually tangent circles in the packing. Hence if $a \in \mathbb{Z}^4$ and is primitive then so is every member of the orbit and in particular every curvature is an integer. The Diophantine properties of the numbers that are

curvatures of an integral packing are quite subtle and are investigated in [**13**]. The reason for the subtlety is that the Apollonian group A is clearly a subgroup of $O_F(\mathbb{Z})$ but it is of infinite index in the latter (corresponding to the fractal dimension $\delta = 1.30\ldots$). Still, the $Zcl(A)$ is all of O_F, which is important for our investigation below. From the point of view of our theme in this lecture the immediate question is whether there are infinitely many circles in an integral packing with curvature a prime number. Or on looking at Figure 2, are there infinitely many "twin primes" that is pairs of tangent circles with curvatures that are both prime?

Affine Orbits and Saturation. — There is a simple and uniform formulation of all the questions above which is as follows: Let L be a group of morphisms (that is polynomial maps) of affine n-space which preserves $\mathbb{Z}^n$. Let $c \in \mathbb{Z}^n$ and $\mathcal{O} = cL$ the corresponding orbit. If $f \in \mathbb{Q}[x_1, \ldots, x_n]$ for which $f(\mathcal{O})$ is integral and is infinite, we seek points $x \in \mathcal{O}$ at which $f(x)$ has few (or fewest) prime factors. We assume that f when restricted to $\mathcal{O}$ is primitive, that is $\gcd(f(\mathcal{O})) = 1$ (otherwise divide f by the gcd). The key definition is the *saturation number* $r_0(\mathcal{O}, f)$, of the pair $(\mathcal{O}, f)$, which is the least r such the set of $x \in \mathcal{O}$ for which $f(x)$ has at most r prime factors, is Zariski dense in $Zcl(\mathcal{O})$. This number is by no means easy to determine and it is far from clear that it is even finite. Knowing it however answers all our questions. For example the following are easy to check

(i) $r_0(c + q\mathbb{Z}, x) = 1$ is Dirichlet's Theorem.

(ii) $r_0(\mathbb{Z}, x(x+2)) = 2$ is the twin prime conjecture.

(iii) If $f \in \mathbb{Z}[x]$ and f factors into t irreducible factors over $\mathbb{Q}[x]$, then $r_0(\mathbb{Z}, f) = t$ is equivalent to Schinzel's hypothesis H [**34**] concerning simultaneous primality of t distinct irreducible integral polynomials in one variable.

(iv) Let $\mathcal{O} = c + L$ as in the HLC in (5). Then $r_0(\mathcal{O}, x_1 x_2 \ldots x_n) = n$ is equivalent to the HLC as stated in (5).

The fundamental general tool to study r_0 is the Brun combinatorial sieve. He used his ingenious invention to show that $r_0(\mathbb{Z}, x(x+2))$ is finite and his arguments can be easily extended to show that $r_0(\mathbb{Z}, f) < \infty$ for any $f \in \mathbb{Z}[x]$. In fact the combinatorial sieve in any of its axiomatic modern formulations can be used to show that $r_0(\mathcal{O}, f) < \infty$ for any orbit $\mathcal{O}$ of L which is a subgroup of $\mathbb{Z}^n$ acting by additive translations. As pointed out at the end of paragraph (2) above we insist on not restricting $f(x)$ to be positive when looking for primes or numbers with few prime factors. The reason is that in this several variable context the condition that $f(x) > 0$, $f \in \mathbb{Z}[x_1, \ldots, x_n]$ can encode the general diophantine equation (for example if $f(x) = 1 - g^2(x)$ then $f(x) > 0$ is equivalent to $g(x) = 0$). The work of Matiyasevich et al [**26**] on Hilbert's 10^{th} problem shows that given any recursively enumerable subset S of the positive integers, there is a $g \in \mathbb{Z}[x_1, \ldots, x_{10}]$ such that S is exactly the set of positive values

assumed by g. From this it is straight forward to construct an $f \in \mathbb{Z}[x_1, \ldots, x_{10}]$ such that for any $r < \infty$, $\{x \in \mathbb{Z}^{10} : f(x) > 0$ and $f(x)$ is a product of at most r primes$\}$ is not Zariski dense in $Zcl\{x \in \mathbb{Z}^{10} : f(x) > 0\}$. That is if we insist on positive values for f we may lose the basic finiteness of saturation property.

Returning to one variable the theory of the sieve has been developed and refined in far-reaching ways to give good bounds for r_0. For example:

$$r_0(\mathbb{Z}, x(x+2)) \leq 3 \qquad \text{(Chen 1973)}$$

$$r_0(\mathbb{Z}, x^2+1) \leq 2 \qquad \text{(Iwaniec 1978)}$$

$$r_0(\mathbb{Z}, f) \leq d+1, \quad \text{if } f \text{ is irreducible over } \mathbb{Q}[x] \text{ and has degree } d \text{ [17].}$$

The first two are especially striking as they come as close as possible to the twin prime and Euler problems, without solving them.

While there are interesting examples of groups L acting nonlinearly and morphically on A^n and preserving $\mathbb{Z}^n$, that come from the actions of mapping class groups on representation varieties [11], the understanding of anything about saturation numbers in such cases is very difficult and is at its infancy. For L acting linearly (as in paragraphs (6) and (7)) a theory can be developed.

An Affine Linear Sieve. — The classical setting is concerned with motions of n-space of the form $x \longrightarrow x + b$. In this affine linear setting we allow multiplication as well, that is transformations of the form $x \longrightarrow xa + b$ with $a \in GL_n(\mathbb{Z})$ and $b \in \mathbb{Z}^n$ (such as the orthogonal group examples (6) and (7)). Note that it is only for $n \geq 2$ that this group of motions is significantly larger than translations (since $GL_1(\mathbb{Z}) = \pm 1$). For the purpose of developing a Brun combinatorial sieve, apparently multiplication is quite a bit more difficult than addition. The basic problems for our pair $(\mathcal{O}, f)$ are

(i) Is $r_0(\mathcal{O}, f)$ finite?

(ii) If it is, then to give good upper bounds for $r_0(\mathcal{O}, f)$. Ideally these should be in terms of the degree of f and its factorization in the coordinate ring of $Zcl(\mathcal{O})$, as has been done in the setting of one variable [17].

(iii) To determine $r_0(\mathcal{O}, f)$ for some interesting pairs and to give an algorithm to predict its exact value in general, that is a generalized local to global conjecture for which HLC and Schinzel's Hypothesis H, are special cases.

When L is a group of affine linear transformations we now have a theory that comes close to answering these questions, there being the caveat of tori (see below) and some other nontrivial technical issues that still need to be resolved in general. In Bourgain-Gamburd-Sarnak ([4], [5]) the finiteness of $r_0(\mathcal{O}, f)$ is proven in many cases. The new tools needed to address these questions, as well as the general setup that we have been discussing are introduced in these papers. The proof given there

of the finiteness does not yield any feasible values for $r_0(\mathcal{O}, f)$. In [**28**] the problem is studied in the case that L is a congruence subgroup of the $\mathbb{Q}$ points of a semi-simple linear algebraic group defined over $\mathbb{Q}$, such as the group $O_F(\mathbb{Z})$ in paragraphs (6) and (7) above (an affine linear action can be linearized by doubling the number of variables). For such congruence L's we develop the combinatorial sieve using tools from the general theory of automorphic forms on such groups and in particular make use of the strong bounds towards the general Ramanujan Conjectures that are now known (see [**31**], [**9**]). With this we get, in this congruence subgroup case, effective bounds for $r_0(\mathcal{O}, f)$ which in many such cases are of the same quality as what is known in one variable.

There is a lacuna in this affine linear sieve theory coming from tori. As we mentioned, allowing multiplication as well as addition, is what makes the problem hard and in fact pure multiplication is simply too hard and even the finiteness is questionable in that case. Consider the example of $L = \left\{ \left[\begin{smallmatrix} 3 & 1 \\ -1 & 0 \end{smallmatrix} \right]^n, n \in \mathbb{Z} \right\} \leq \mathrm{SL}_2(\mathbb{Z})$. L is infinite cyclic, $Zcl(L)$ is a torus and if $\mathcal{O} = (1,0) \cdot L$ then $Zcl(\mathcal{O})$ is the hyperbola $\{(x_1, x_2) = x_1^2 - 3x_1x_2 + x_2^2 = 1\}$. The orbit consists of pairs (F_{2n}, F_{2n-2}) $n \in \mathbb{Z}$ where F_m is the m^{th} Fibonacci number. This kind of sequence is too sparse both from the analytic and algebraic points of view to do any kind of (finite) sieve. While it is conjectured that F_m is prime for infinitely many m, as was pointed out to me by Lagarias, standard heuristic probabilistic considerations suggest a very different behavior for F_{2n}. Indeed $F_{2n} = F_n \cdot L_n$ where L_n is the n^{th} Lucas number and assuming a probabilistic model for the number of prime factors of a large integer in terms of its size and that F_n and L_n are independent leads to F_{2n} having an unbounded number of prime factors as $n \to \infty$. A precise conjecture along these lines is put forth in [**8**] (see Conjecture 5.1). In our language this asserts that if $\mathcal{O}$ is as above and $f(x_1, x_2) = x_1$ then $r_0(\mathcal{O}, f) = \infty$. It would be very interesting to produce an example of a pair $(\mathcal{O}, f)$ for which one can prove that $r_0(\mathcal{O}, f)$ is infinite. In view of the above we must steer clear of tori and the precise setting in which the affine linear sieve is developed (see [**29**]) is for linear L's for which the radical (the largest normal solvable subgroup) of the $\mathbb{Q}$ linear algebraic group $G := Zcl(L)$, contains no tori (the unipotent radical causes no difficulties). Applying this theory to the examples of orthogonal groups in (6) and (7) we obtain the following. Let $F(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$ and $L \leq O_F(\mathbb{Z})$. Assume that L is not an elementary group (in particular not finite or abelian, in fact precisely that $Zcl(L)$ is either of the linear algebraic groups O_F or SO_F). If $\mathcal{O} = (3,4,5)L$, then $Zcl(\mathcal{O}) = C$ the affine cone; $F = 0$. For $f \in \mathbb{Z}[x_1, x_2, x_3]$ the results in [**5**] imply that $r_0(\mathcal{O}, f) < \infty$. In particular this applies to $f(x) = A(x) = x_1x_2/2$, the area. This says that given such an orbit of Pythagorean triangles (which may be very sparse!) there is an $r < \infty$ such that the set of triangles in $\mathcal{O}$ whose areas have at most r prime factors is Zariski dense

in C. It is elementary that $\gcd(A(O)) = 6$. From the ancient parametrization of all the Pythagorean triples P (i.e. the $\mathbb{Q}$ morphism of A^2 into C) these are all of the form $(x_1, x_2, x_3) = (a^2 - b^2, 2ab, a^2 + b^2)$ with $a, b \in \mathbb{Z}$, $(a, b) = 1$ and not both odd, one sees that $A/6 = (a-b)(a+b)(ab)/6$. Now the last has at most two prime factors for only finitely many pairs (a, b). The set of (a, b) for which it has at most 3 prime factors lie in a finite union of curves in C (and if HLC is true for $\mathcal{O} = (2, 2, 0) + (3, 3, 1)\mathbb{Z}$, i.e. this rank one orbit in $\mathbb{Z}^3$, then these curves contain infinitely many points with $A/6$ having 3 prime factors). Hence for any $\mathcal{O}$ as above $r_0(\mathcal{O}, A/6) \geq 4$. The general local to global conjectures [**5**] then assert that $r_0(\mathcal{O}, A/6) = 4$ for any such orbit. Interestingly the recent advance in [**15**] mentioned in (5) above just suffices to prove that for the full set of Pythagorean triples P, $r_0(P, A/6) = 4$. Put another way the minimal divisibility of the areas of a Zariski dense set of Pythagorean triangles is 6 (here we include the forced factors 3 and 2). The deduction is immediate, set $a = 2x$ and $b = 3y$ in the ancient parametrization. Then $A/6 = xy(2x + 3y)\ (2x - 3y)$ and apply [**15**] to $\mathcal{O} = L = (1, 0, 2, 2)\mathbb{Z} + (0, 1, 3, -3)\mathbb{Z}$. For some other applications of [**16**] see Granville [**14**].

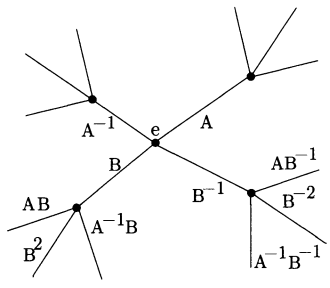
As an example of an application of the affine linear sieve in the context of an L which is a congruence group, consider an integral quadratic form $F(x)$ in 3-variables. That is $F(x) = x^t A x$ where A is 3×3 symmetric and is integral on the diagonal and half integral on the off-diagonal. We assume that F is indefinite over the reals but that it is anisotropic over $\mathbb{Q}$ (so $F(x) = 0$ for $x \in \mathbb{Z}^3$ implies that $x = 0$) and that $\det A$ is square free (so $F(x_1, x_2, x_3) = x_1^2 + x_2^2 - 7x_3^2$ is an example, the anisotropy following from looking at $F(x) \equiv 0 \bmod 8$). Let $0 \neq t \in \mathbb{Z}$ for which $V_t(\mathbb{Z}) = \{x \in \mathbb{Z}^3 : F(x) = t\}$ is nonempty, which according to the work of Hasse and Siegel will happen iff there are no local congruence obstructions to solving $F(x) \equiv t (\bmod q)$ for $q \geq 1$. In this case $V_t(\mathbb{Z})$ is a finite union of $O_F(\mathbb{Z})$ orbits and $Zcl(V_t(\mathbb{Z})) = V_t$, the affine quadric $\{x : F(x) = t\}$. We seek points in $V_t(\mathbb{Z})$ whose coordinates have few prime factors, i.e. to estimate $r_0(V_t(\mathbb{Z}), x_1x_2x_3)$. By the general finiteness theorem, $r_0(V_t(\mathbb{Z}), x_1x_2x_3)$ is finite. However by developing optimal weighted counting results on such quadrics and also exploiting the best bounds known towards the Ramanujan-Selberg Conjecture, it is shown in [**24**] that $r_0(V_t(\mathbb{Z}), x_1x_2x_3) \leq 26$.

We turn to the Apollonian packing. An extension of the $(\mathcal{O}, f)$ finiteness theorem in [**5**] applies to the orbit $\mathcal{O} = aL$ for any L which is Zariski dense in O_F, where F is the quadratic form in 4-variables in paragraph (7). In particular it applies to the Apollonian group A with $f(x) = x_1x_2x_3x_4$. This asserts that in any given integral packing there is an $r < \infty$ such that the set of 4 mutually tangent circles in the packing for which all 4 curvatures have at most r prime factors is Zariski dense in $Zcl(\mathcal{O}) = C = \{x : F(x) = 0\}$. One can determine r_0 for $\mathcal{O} = a.A$ and some special f's using some ad hoc and elementary methods together with (ii) of paragraph (4).

In [**32**] it is shown that $r_0(\mathcal{O}, x_1) = 1$ and $r_0(\mathcal{O}, x_1 x_2) = 2$, from which it follows that in any such packing there are infinitely many circles whose curvatures are prime and better still there are infinitely many pairs of tangent circles both of whose curvatures are prime.

As a final example of an interesting pair $(\mathcal{O}, f)$ for which we can determine r_0, consider the variety V_t in affine n^2-space given by $V_t = \{X = (x_{ij})_{\substack{i=1,\dots n \\ j=1,\dots n}} : \det X = t\}$. For t a nonzero integer $V_t(\mathbb{Z})$ consists of a finite union of $L = \mathrm{SL}_n(\mathbb{Z})$ orbits where the action of g is by $X \longrightarrow X.g$. In [**28**] we show using Vinogradov's methods mentioned in (5), that if $n \geq 3$ then $r_0(V_t(\mathbb{Z}), \prod_{i,j} x_{ij}) = n^2$ if $\prod_{i,j} x_{ij}$ is primitive on $V_t(\mathbb{Z})$. Examining in detail when this happens, we deduce that the set of $n \times n$ integral matrices of determinant t all of whose entries are prime, is Zariski dense in V_t iff $t \equiv 0(\mod 2^{n-1})$. This should of course, also hold for $n = 2$ where it is concerned with the equation $x_{11}x_{22} - x_{12}x_{21} = t$ and the x_{ij}'s are to be primes. The best that appears to be known concerning this is the recent development by [**12**] from which it follows that for this $n = 2$ case, $r_0(V_t(\mathbb{Z}), x_{11}x_{12}x_{21}x_{22}) = 4$, for at least one t in $\{2,4,6\}$.

Comments about Proofs. — I end the lecture with a very brief hint as to what is involved in developing a combinatorial sieve in the affine linear context. This entails getting a little more technical. Let $\mathcal{O} = cL$ be our orbit and $f \in \mathbb{Z}[x_1, \dots, x_n]$. After some algebro-geometric reductions of the problems (using the $\mathbb{Q}$ dominant morphisms from $G = Zcl(L)$ to $V = Zcl(\mathcal{O})$ and $\widetilde{G}$ to G where $\widetilde{G}$ is the simply connected cover of G) we can assume that $\mathcal{O}$ is the group L itself (as a group of matrices in the affine space of $n \times n$ matrices) and $Zcl(\mathcal{O}) = Zcl(L) = G$ is a simply connected $\mathbb{Q}$-group. To do any kind of sieving we need to order the elements of L, so as to carry out some truncated inclusion-exclusion procedure, this being at the heart of Brun's method. Usually one orders by archimedian size perhaps with positive weights, however in this general setting we don't know how to do this, so we order L combinatatorially instead. For the groups that we are considering and for the purpose of proving that $r_0(\mathcal{O}, f)$ is finite, we can (according to a theorem of Tits) assume that L is free on two generators A and B. We use the tree structure of the Cayley graph $T = (L, S)$ of L with respect to the generators $S = \{A, A^{-1}, B, B^{-1}\}$. T is a 4-regular tree;



For $x, y \in T$ let $d(x,y)$ denote the distance from x to y in the tree. The key sums that arise in sieving on L for divisibility of f are:

For $d \geq 1$ square-free and $x_0 \in T$,

$$S(Y,d) := \sum_{\substack{x \in L \\ d(x,x_0) \leq Y \\ f(x) \equiv 0(d)}} 1,$$

or perhaps with 1 replaced by positive weights.

We are interested in $S(Y,d)$ when Y is large and d as large as $e^{\alpha Y}$ for some $\alpha > 0$. The larger the α for which S can be understood the better. To study such sums a couple of key features intervene:

(i) Algebraic stabilization: This is the analogue of the Chinese remainder theorem. We state it for the basic case of $G = \mathrm{SL}_n$, it is valid for G semisimple and simply connected. It is due (originally) to Matthews-Vaserstein and Weisfeiler [**27**] who employ the classification of finite simple groups in the proof. Let $L \leq \mathrm{SL}_n(\mathbb{Z})$ be Zariski dense in SL_n. Then there is a positive integer $\nu = \nu(L)$ such that for d with $(d,\nu) = 1$ the reduction $L \longrightarrow SL_n(\mathbb{Z}/d\mathbb{Z})$ is onto.

This eventually allows us to bring in more standard tools from arithmetic algebraic geometry, in order to identify the main term in the form

$$S(Y,d) = \beta(d)\, S(Y,1) + R(Y,d).$$

Here $\beta(d)$ is a multiplicative arithmetical function associated with counting points mod d on the variety $G \cap \{f = 0\}$ and R is the remainder which is expected to be smaller. The demonstration of the latter for the purpose of sieving far enough to get the finiteness of $r_0(\mathcal{O}, f)$, is essentially equivalent to the second feature.

(ii) The (finite) Cayley graphs $(\mathrm{SL}_n(\mathbb{Z}/d\mathbb{Z}), S)$ are an expander family as $d \longrightarrow \infty$ (see [**30**] for a definition of expanders and [**25**] where this is conjectured). As yet, this expander property has not been established in general and this is the main reason that the finiteness of $r_0(\mathcal{O}, f)$ has not been established in general for the affine linear sieve. It is proven for SL_2 and related groups for d square free, in [**5**]. The proof uses a variety of inputs some of which were to me at least, quite unexpected. We list them for the simpler case that $d = p$ is prime:

(a) The dichotomy that an irreducible complex representation of $G(\mathbb{Z}/p\mathbb{Z})$ is either 1-dimensional or is of very large dimension (here $p \to \infty$) coupled with a "softer" *upper* bound density theorem for multiplicities of exceptional eigenvalues of the Cayley graphs, leads to a proof of the key spectral gap defining an expander [**33**]. For the soft upper bound we use techniques from arithmetic combinatorics.

(b) Sum-Product Theorem [**6**]: This is an elementary and very useful theorem concerning mixing the additive and multiplicative structures of a finite field. Let $\epsilon > 0$ be given, there is a $\delta > 0$, $\delta = \delta(\epsilon)$, such that if $A \subset \mathbb{F}_p$ and $|A| \leq p^{1-\epsilon}$ then $|A + A| + |A \cdot A| \geq |A|^{1+\delta}$ (here p is sufficiently large).

(c) Helfgott's $\mathrm{SL}_2(\mathbb{F}_p)$ Theorem [**21**]: Let $\epsilon > 0$ there is $\delta = \delta(\epsilon) > 0$ such that if $A \subset \mathrm{SL}_2(\mathbb{F}_p)$, A is not contained in a proper subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$ and $|A| \leq |\mathrm{SL}_2(\mathbb{F}_p)|^{1-\epsilon}$, then $|A \cdot A \cdot A| \geq |A|^{1+\delta}$.

(d) Balog-Szemerédi, Gowers Theorem: This is a purely combinatorial theorem from graph theory which is used in [**3**] to give the required upper bounds on counting closed circuits in the graph, and leads to a proof that $(\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), S)$ is an expander family.

A point worth noting is that once the affine sieve is set up and gives lower bounds in our combinatorial group theoretic ordering, for points in $\mathcal{O}$ for which f has at most r prime factors, the expander property is used again and in a different way to demonstrate the Zariski density of these points.

To end let me highlight the fundamental difference between the additive translational counting and the affine linear counting which necessitates the introduction of expanders. In $\mathbb{Z}$ the boundary of a large interval is small compared with the size of the interval and the same is true uniformly for an arithmetic progression of common difference q in the interval, even for q almost as large as the interval length. On the other hand on a k-regular tree ($k \geq 3$) this is not true. Given a big ball B (or any large finite set), the size of the boundary ∂B is of the same order of magnitude as B. It is exactly the expander property that allows one to draw an effective approximation for the number of points in B lying in the orbit with a congruence condition.

Acknowledgements. — In developing this theory of an affine sieve and the geometric view point described in this lecture, I have benefited from discussions with many people. First and foremost with my collaborators on the different aspects of the theory, Bourgain and Gamburd, Liu, Nevo and Salehi. Also with Lubotzky, Katz, Lindenstrauss and Mazur. Finally thanks to Lagarias who pointed me to his joint works on integral Apollonian packings and their subtle diophantine features.

References

[1] A. Balog – "Linear equations in primes", *Mathematika* **39** (1992), p. 367–378.

[2] V. Bergelson & A. Leibman – "Polynomial extensions of van der Waerden's and Szemerédi's theorems", *J. Amer. Math. Soc.* **9** (1996), p. 725–753.

[3] J. Bourgain & A. Gamburd – "Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$", *Ann. of Math.* **167** (2008), p. 625–642.

[4] J. BOURGAIN, A. GAMBURD & P. SARNAK – "Sieving and expanders", *C. R. Math. Acad. Sci. Paris* **343** (2006), p. 155–159.

[5] ———, "Affine sieve, expanders and sum product", `http://www.math.princeton.edu/sarnak/sespM8.pdf`.

[6] J. BOURGAIN, N. KATZ & T. TAO – "A sum-product estimate in finite fields, and applications", *Geom. Funct. Anal.* **14** (2004), p. 27–57.

[7] D. W. BOYD – "The sequence of radii of the Apollonian packing", *Math. Comp.* **39** (1982), p. 249–254.

[8] Y. BUGEAUD, F. LUCA, M. MIGNOTTE & S. SIKSEK – "On Fibonacci numbers with few prime divisors", *Proc. Japan Acad. Ser. A Math. Sci.* **81** (2005), p. 17–20.

[9] L. CLOZEL – "Démonstration de la conjecture τ", *Invent. Math.* **151** (2003), p. 297–328.

[10] J. FRIEDLANDER & H. IWANIEC – "The polynomial X^2+Y^4 captures its primes", *Ann. of Math.* **148** (1998), p. 945–1040.

[11] W. M. GOLDMAN – "The modular group action on real SL(2)-characters of a one-holed torus", *Geom. Topol.* **7** (2003), p. 443–486.

[12] D. GOLDSTON, J. GRAHAM, J. PINTZ & C. YILDRIM – "Small gaps between products of two primes", preprint, arXiv:math/0609615v1, 2006.

[13] R. L. GRAHAM, J. C. LAGARIAS, C. L. MALLOWS, A. R. WILKS & C. H. YAN – "Apollonian circle packings: number theory", *J. Number Theory* **100** (2003), p. 1–45.

[14] A. GRANVILLE – "Prime number patterns", *Amer. Math. Monthly* **115** (2008), p. 279–296.

[15] B. GREEN & T. TAO – "Linear equations in primes", to appear in *Annals of Math.*, arXiv:math/0606088v2, 2006.

[16] ———, "The primes contain arbitrarily long arithmetic progressions", *Ann. of Math.* **167** (2008), p. 481–547.

[17] H. HALBERSTAM & H. RICHERT – *Sieve methods*, Academic Press, 1974.

[18] G. HARDY & J. LITTLEWOOD – "Some problems of 'Partitio Numerorum.' III. On the expression of a number as a sum of primes", *Acta Math.* **44** (1922), p. 1–70.

[19] D. R. HEATH-BROWN & B. Z. MOROZ – "Primes represented by binary cubic forms", *Proc. London Math. Soc.* **84** (2002), p. 257–288.

[20] K. HEEGNER – "Diophantische Analysis und Modulfunktionen", *Math. Z.* **56** (1952), p. 227–253.

[21] H. A. HELFGOTT – "Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$", *Ann. of Math.* **167** (2008), p. 601–623.

[22] B. HOST & B. KRA – "Nonconventional ergodic averages and nilmanifolds", *Ann. of Math.* **161** (2005), p. 397–488.

[23] H. IWANIEC – "Primes represented by quadratic polynomials in two variables", *Acta Arith.* **24** (1973/74), p. 435–459.

[24] J. LIU & P. SARNAK – "Integral points on quadrics in three variables whose coordinates have few prime factors", to appear in *Israel Jnl. of Math.*, `http://www.math.princeton.edu/sarnak/few-final.pdf`.

[25] A. LUBOTZKY – "Cayley graphs: eigenvalues, expanders and random walks", in *Surveys in combinatorics, 1995 (Stirling)*, London Math. Soc. Lecture Note Ser., vol. 218, Cambridge Univ. Press, 1995, p. 155–189.

[26] Y. V. MATIYASEVICH – *Hilbert's tenth problem*, Foundations of Computing Series, MIT Press, 1993.

[27] C. R. MATTHEWS, L. N. VASERSTEIN & B. WEISFEILER – "Congruence properties of Zariski-dense subgroups. I", *Proc. London Math. Soc.* **48** (1984), p. 514–532.

[28] A. NEVO & P. SARNAK – "Prime and almost prime integral points on principal homogeneous spaces", **http://www.math.princeton.edu/sarnak/NS-final-Oct-08.pdf**.

[29] A. SALEHI & P. SARNAK – in preparation.

[30] P. SARNAK – "What is an expander?", *Notices Amer. Math. Soc.* **51** (2004), p. 762–763.

[31] ———, "Notes on the generalized Ramanujan conjectures", in *Harmonic analysis, the trace formula, and Shimura varieties*, Clay Math. Proc., vol. 4, Amer. Math. Soc., 2005, p. 659–685.

[32] ———, *Letter to J. Lagarias*, June 2007, **http://www.math.princeton.edu/sarnak/AppolonianPackings.pdf**.

[33] P. SARNAK & X. X. XUE – "Bounds for multiplicities of automorphic representations", *Duke Math. J.* **64** (1991), p. 207–227.

[34] A. SCHINZEL & W. SIERPIŃSKI – "Sur certaines hypotheses concernant les nombres premiers", *Acta Arith.* **4** (1958), p. 185–208.

[35] S. SHELAH – "Logical dreams", *Bull. Amer. Math. Soc. (N.S.)* **40** (2003), p. 203–228.

[36] T. TAO & T. ZIEGLER – "The primes contain arbitrary long polynomial progressions", to appear in *Acta Math.*, 2006.

[37] R. C. VAUGHAN – *The Hardy-Littlewood method*, Cambridge Tracts in Mathematics, vol. 80, Cambridge University Press, 1981.

[38] A. WILES – "The Birch and Swinnerton-Dyer conjecture", in *The millennium prize problems*, Clay Math. Inst., Cambridge, MA, 2006, p. 31–41.

P. SARNAK, Princeton University & Institute for Advanced Study, Princeton, NJ, USA