Snapshots of modern mathematics
from Oberwolfach

№ 10/2016

Prime Tuples in Function Fields

Lior Bary-Soroker

How many prime numbers are there? How are they distributed among other numbers? These are questions that have intrigued mathematicians since ancient times. However, many questions in this area have remained unsolved, and seemingly unsolvable in the forseeable future.

In this snapshot, we will discuss one such problem, the Twin Prime Conjecture, and a quantitative version of it known as the Hardy–Littlewood Conjecture. We will also see that these and other questions about prime numbers can be extended to questions about *function fields*, and discuss recent progress which has been made to answer them in this context.

1 The Prime Number Theorem

For any number x, we use $\pi(x)$ to denote $\#\{p \leq x\}$, the number of prime numbers p up to x. Euclid's Theorem (c. 300 BC) asserts that there are infinitely many prime numbers, that is

$$\lim_{x\to\infty} \pi(x) = \infty.$$

It took more than 2000 years and the development of complex analysis to obtain a *quantitative version* of Euclid's theorem, that is, a statement of just how fast $\pi(x)$ grows.

The Prime Number Theorem, which was conjectured by Gauß and proved independently by Hadamard and by de la Vallée-Poussin in 1896,[1] asserts that

$$\pi(x) \sim \frac{x}{\log x}, \qquad x \to \infty.$$

Here $f(x) \sim g(x)$ means that $\lim_{x\to\infty} f(x)/g(x) = 1$.

It will be convenient for our discussion to adopt a more statistical or probabilistic point of view. The *expectation* of a function f on a finite set A, which can be thought of either as its average value, or as the typical value $f(a)$ one would expect if one repeatedly picked a at random from A, is computed by:

$$\langle f(a)\rangle_{a\in A} = \frac{1}{\#A}\sum_{a\in A} f(a).$$

We define the *prime characteristic function*, χ, so that $\chi(n) = 1$ if n is prime and $\chi(n) = 0$ if n is not prime. Using this function, and observing that there are approximately x natural numbers n in the interval $[1, x]$, we can write the Prime Number Theorem in the (statistical) form:

$$\langle \chi(n)\rangle_{n\in[1,x]} \sim \frac{1}{\log x}, \qquad x \to \infty. \tag{1}$$

2 The Hardy–Littlewood Prime Tuple Conjecture

Twin primes are pairs $(n, n+2)$ such that both n and $n+2$ are prime; e.g., $(3, 5)$, $(101, 103)$, and $(30089, 30091)$. The Twin Prime Conjecture — which remains an open question — asserts that there are infinitely many twin primes. There is a quantitative version of this conjecture, called the Hardy–Littlewood Conjecture, that predicts the auto-correlation of the prime characteristic function χ:

$$\langle \chi(n)\chi(n+2)\rangle_{n\in[1,x]} \sim \mathfrak{S}_2\langle \chi(n)\rangle^2_{n\in[1,x]} \sim \mathfrak{S}_2\frac{1}{(\log x)^2}, \tag{2}$$

as $x \to \infty$, where $\mathfrak{S}_2 \approx 1.32$ is an absolute constant.

This Conjecture has two interesting corollaries:

1. Since $x \cdot \langle \chi(n)\chi(n+2)\rangle_{n\in[1,x]}$ asymptotically equals the number of $n \le x$ such that n and $n+2$ are both primes, (2) implies that the number of twin primes is asymptotically proportional to $\frac{x}{(\log x)^2}$ and in particular tends to infinity. So (2) implies the Twin Prime Conjecture.

[1] Carl Friedrich Gauß (1777–1855); Jacques Salomon Hadamard (1865–1963); Charles-Jean Étienne Gustave le Vieux de la Vallée-Poussin (1866–1962)

2. The statistical interpretation of (2) is that the events of n being a prime and of $n+2$ being a prime are *asymptotically independent*, up to the correlation factor $\mathfrak{S}_2$. That is, if $\mathfrak{S}_2$ were exactly 1, we would know that these events are getting closer to being independent the larger n gets. Since $\mathfrak{S}_2 \approx 1.32 > 1$, the events depend on each other; so, if n is prime, the probability that $n+2$ is a prime is bigger than the probability that a random number (in $[1, x]$) is a prime. This make sense, since, if $n > 2$ is a prime, it is odd, hence $n+2$ is odd, thus more likely to be a prime than a random number.

Though we have not encountered much trouble to state them, the Hardy–Littlewood and the Twin Prime Conjectures turn out to be some of the most difficult problems in mathematics.

3 The Function Field Setting

Up till now, we have considered integers. Now we introduce an ancient analogy between the integers and polynomials over a finite field. A finite field is a finite set with the four operations of addition, multiplication, subtraction, and division. Note that addition and multiplication determine the other two operations. The smallest finite field has 2 elements, and is denoted $\mathbb{F}_2 = \{0, 1\}$. The operations are given by

$$\begin{aligned} 0+0=0, \qquad & 0+1=1, \qquad & 1+1=0, \\ 0\cdot 0=0, \qquad & 0\cdot 1=0, \qquad & 1\cdot 1=1, \end{aligned}$$

Slightly more generally, if p is a prime, then the set of residues modulo p, denoted by $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$, with operations modulo p is also a finite field (here the division function is given by Euclid's algorithm, which says that for each $1 \le x \le p-1$ there exist a, b so that $ax + bp = 1$, and so a is the inverse of x modulo p). The general theory of finite fields says that for every power of a prime number $q = p^k$, there exists a unique (up to isomorphism) finite field $\mathbb{F}_q$ with q elements, and that there are no others. These finite fields play key role in many theories; e.g., coding theory, cryptography, and most notably in our setting, in number theory.

For a finite field $\mathbb{F}_q$, we consider the set of polynomials over it, denoted $\mathbb{F}_q[T]$. A polynomial f in $\mathbb{F}_q[T]$ is a formal expression

$$f(T) = a_n T^n + \cdots a_1 T + a_0$$

where T is a formal variable. If $a_n \neq 0$, we say that f has degree n (symbolically, $\deg(f) = n$). We can add and mulitply polynomials in the standard way. Since at least the 19th century, it has been observed that $\mathbb{F}_q[T]$ is intimately connected

with the integers $\mathbb{Z}$; in particular the *monic* polynomials — those with leading coefficient $a_n = 1$ — play the role the positive integers take in $\mathbb{Z}$. We denote by $M_{n,q} \subseteq \mathbb{F}_q[T]$ the set of monic polynomials of degree n.

We call a polynomial P *irreducible* if it cannot be written as a product of two polynomials of positive degrees. The irreducible monic polynomials play the role of prime numbers. So we call a polynomial $P \in \mathbb{F}_q[T]$ *prime polynomial* if it is irreducible and monic.

In $\mathbb{Z}$, the absolute value of an integer $n \neq 0$ may be viewed as the number of residues modulo n. This leads one to define the *norm* of a nonzero polynomial f in $\mathbb{F}_q[T]$ as the number of residues modulo f:

$$|f| = q^{\deg f}.$$

Let χ_q be the characteristic function of the prime polynomials; that is, set $\chi_q(f) = 1$ if f is a prime polynomial and $\chi_q(f) = 0$ if f is not a prime polynomial. The Prime Polynomial Theorem, which is the analogue of the Prime Number Theorem, asserts that

$$\langle \chi_q(f) \rangle_{f \in M_{n,q}} \sim \frac{1}{n}, \qquad q^n \to \infty. \tag{3}$$

It turns out that the proofs involved in establishing (3) are easy, in contrast to those of (1). There are two things to notice about (3).

1. If we set $x = q^n$, then $\log_q(x) = n$, so we see that the form of (3) is in perfect agreement with (1).
2. The asymptotic formula (3) holds whenever the parameter q^n is large. In fact there are two ways for q^n to be large: q could be large, even if n is not, or n could be large even if q is not. Usually asymptotic formulas behave differently in these limits, but not (3).

Now we can try to ask for the analogy of twin primes in $\mathbb{F}_q[T]$. As there is nothing special about the number 2 in the function field setting, we will consider pairs $(f, f+A)$, where $A \in \mathbb{F}_q[T]$; or the more general tuples $(f+A_1, \ldots, f+A_k)$, with distinct $A_1, \ldots, A_k \in \mathbb{F}_q[T]$.

In light of the above remarks, we will consider the limits $q \to \infty$ and $n \to \infty$ separately.

3.1 The limit $q \to \infty$

Recently, the function field version of the Hardy–Littlewood Conjecture was completely resolved in this limit:

Theorem 1 *Let $n > 2$ and $k > 0$ be fixed integers. Then*

$$\left\langle \prod_{i=1}^{k} \chi_q(f + A_{i,q}) \right\rangle_{f \in M_{n,q}} \sim \frac{1}{n^k}, \qquad q \to \infty$$

for any choice of $A_{i,q} \in \mathbb{F}_q[T]$ such that for each q the polynomials $A_{1,q}, \ldots, A_{k,q}$ are distinct and of degree $< n$.

Bender and Pollack [3] proved the theorem in the case $k = 2$ and q odd. The author [2] used a more general approach to prove the result for any k when q is odd. Carmon, in his work [4] on the Chowla conjecture in characteristic $q = 2$ (a different but related problem), developed the necessary tools to extend the author's method to even characteristic. Finally, Bank and the author [1] study a more general problem on simultaneous prime values of linear functions which in particular allows the $A_{i,q}$ to be of arbitrary bounded degrees.

The key points in this result are to calculate an algebraic invariant called the Galois group of a "generic" polynomial $\mathcal{F}$ and then to apply a high dimensional version of Chebotarev's theorem over finite fields. We will not discuss this any further here, but we will say that one chooses $\mathcal{F} = \prod_{i=1}^{n}(F + A_i)$, where F is the monic polynomial of degree n whose coefficients are independent variables, and its Galois group tend to be a direct product of n copies of the symmetric group on n letters.[2]

3.2 The limit $n \to \infty$

In this limit the Hardy–Littlewood Conjecture is completely open.

In fact, in 2006, Hall [6] showed that for a nonzero $A \in \mathbb{F}_q[T]$ with $\deg A = 0$ (i.e. $A \in \mathbb{F}_q$) there exist infinitely many polynomials $f \in \mathbb{F}_q[T]$ with both f and $f + A$ irreducible ($q > 3$). However, Hall's proof gives a sequence of f's with exponentially increasing degrees. In particular, it does not imply that the auto-correlation is positive. We refer the reader to [8] for a more detailed exposition of Hall's result and further developments in this direction.

[2] For more on the symmetric group on n letters, see Snapshot $\mathcal{N}^o$ 5/2016 *Symmetry and Characters of Finite Groups* by Eugenio Gianelli and Jay Taylor.

4 Further works

This snapshot is focused on just one problem in the number theory of function fields. In fact there is an abundance of recent works in the number theory of function fields. We refer the interested reader to the survey paper by Rudnick [9] from International Congress of Mathematicians 2014.[3] Even since this survey was written, more interesting results have been published. Two examples are:

1. Entin's work [5] on prime *values* of polynomials
2. Keating and Roddity-Gershon's work [7] on the average of the error term

$$E = \langle \chi_q(f)\chi_q(f+A)\rangle_{M_{n,q}} - 1/n^2$$

in Theorem 1 (where A ranges over suitable sets) which implies that E cannot be too small.

Acknowledgements

The author thanks Arno Fehm for a careful reading of an earlier version of the manuscript and the *Snapshots* editorial team for many suggestions and changes that improved the presentation of this paper. Some of the research discussed in this paper was partially supported by the Israel Science Foundation (grant No. 925/14).

[3] The International Congress of Mathematicians is the major meeting of the International Mathematical Union, held every four years. In 2014 the ICM was held in Seoul, Korea.

References

[1] E. Bank and L. Bary-Soroker. *Prime polynomial values of linear functions in short intervals.* Journal of Number Theory, 151:263–275, 2015.

[2] L. Bary-Soroker. *Hardy–Littlewood tuple conjecture over large finite fields.* International Mathematics Research Notices IMRN, (2):568–575, 2014.

[3] A. O. Bender and P. Pollack. *On quantitative analogues of the Goldbach and twin prime conjectures over* $F_q[t]$. arXiv:0912.1702v1, 2009.

[4] D. Carmon. *The autocorrelation of the Möbius function and Chowla's conjecture for the rational function field in characteristic* 2. Philosophical Transactions of the Royal Society A, 373(2040), 2015.

[5] A. Entin. *On the Bateman–Horn conjecture for polynomials over large finite fields.* arXiv:1409.0846v2, 2014.

[6] C. Hall. *L-functions of twisted Legendre curves.* Journal of Number Theory 119(1):128–147, 2006.

[7] E. Roditty-Gershon and J. Keating. *Arithmetic Correlations over Large Finite Fields.* preprint, 2015.

[8] P. Pollack. *Arithmetic properties of polynomial specializations over finite fields.* Acta Arithmetica, 136(1):57–79, 2009.

[9] Z. Rudnick. *Some problems in analytic number theory for polynomials over a finite field.* Proceedings of the International Congress of Mathematicians 2014, vol. 1, 2014.

Lior Bary-Soroker *is an associate professor of pure mathematics at the Tel Aviv University.*

Mathematical subjects
Algebra and Number Theory

License
Creative Commons BY-NC-SA 4.0

DOI
10.14760/SNAP-2016-010-EN

Snapshots of modern mathematics from Oberwolfach are written by participants in the scientific program of the Mathematisches Forschungsinstitut Oberwolfach (MFO). The snapshot project is designed to promote the understanding and appreciation of modern mathematics and mathematical research in the general public worldwide. It started as part of the project "Oberwolfach meets IMAGINARY" in 2013 with a grant by the Klaus Tschira Foundation. The project has also been supported by the Oberwolfach Foundation and the MFO. All snapshots can be found on www.imaginary.org/snapshots and on www.mfo.de/snapshots.

Junior Editor
Andrew Cooper
junior-editors@mfo.de

Senior Editor
Carla Cederbaum
senior-editor@mfo.de

Mathematisches Forschungsinstitut
Oberwolfach gGmbH
Schwarzwaldstr. 9–11
77709 Oberwolfach
Germany

Director
Gerhard Huisken