Journal de Théorie des Nombres
de Bordeaux **19** (2007), 221–229

Sieve methods for varieties over finite fields and arithmetic schemes

par Bjorn POONEN

Résumé. Des méthodes du crible classiques en théorie analytique des nombres ont été récemment adaptées à un cadre géométrique. Dans ce nouveau cadre, les nombres premiers sont remplacés par les points fermés d'une variété algébrique sur un corps fini ou plus généralement un schéma de type fini sur $\mathbb{Z}$. Nous présentons la méthode et certains des résultats surprenants qui en découlent. Par exemple, la probabilité qu'une courbe plane sur $\mathbb{F}_2$ soit lisse est asymptotiquement $21/64$ quand son degré tend vers l'infini. La plus grande partie de cet article est une exposition des résultats de [Poo04] et [Ngu05].

Abstract. Classical sieve methods of analytic number theory have recently been adapted to a geometric setting. In the new setting, the primes are replaced by the closed points of a variety over a finite field or more generally of a scheme of finite type over $\mathbb{Z}$. We will present the method and some of the surprising results that have been proved using it. For instance, the probability that a plane curve over $\mathbb{F}_2$ is smooth is asymptotically $21/64$ as its degree tends to infinity. Much of this paper is an exposition of results in [Poo04] and [Ngu05].

1. Squarefree integers

Before discussing our geometric sieve, let us recall a simple application of classical sieve techniques, to the counting of squarefree integers.

Consider the problem of determining the "probability" that a "random" positive integer is squarefree. To make sense of this problem, we should clarify what is meant by probability, since the set of positive integers is countably infinite. For any subset $S \subseteq \mathbb{N}$, define the *density* of S as the

Manuscrit reçu le 9 janvier 2006.
Mots clefs. Bertini theorem, finite field, Lefschetz pencil, squarefree integer, sieve.
This is an extension of a plenary lecture given at the 24th Journées Arithmétiques in Marseilles, July 4–8, 2005. The research was supported by NSF grant DMS-0301280, a Packard Fellowship, and the Miller Institute for Basic Research in Science.

limit

$$\mu(S) := \lim_{B\to\infty} \frac{\#(S \cap [1, B])}{B}.$$

In other words, we compute the fraction of the integers from 1 to B that belong to S, and then let B tend to ∞.

From now on, we interpret "the probability that a positive integer is squarefree" as the density of the set S of squarefree positive integers. We can guess the answer by using the following reasoning. An integer n is squarefree if and only if for all primes p, the integer p^2 does not divide n. For each prime p, the probability that an integer is divisible by p^2 is $1/p^2$, so the probability that the integer is *not* divisible by p^2 is $1 - 1/p^2$. These conditions for different p should be independent, by the Chinese remainder theorem. Therefore one predicts that the density of squarefree integers equals

$$\prod_{\text{prime } p} (1 - p^{-2}) = \zeta(2)^{-1} = 6/\pi^2,$$

where $\zeta(s)$ is the Riemann zeta function, defined by

$$\zeta(s) := \sum_{n\geq 1} n^{-s} = \prod_{\text{prime } p} (1 - p^{-s})^{-1}$$

for $\mathrm{Re}(s) > 1$ (which is all we need).

It is not immediate that this heuristic prediction can be made rigorous. The Chinese remainder theorem does imply that for any finite set S of primes, the density of positive integers not divisible by p^2 for any $p \in S$ equals $\prod_{p\in S}(1 - p^{-2})$. But the argument breaks down if we try to apply the Chinese remainder theorem to infinitely many primes. In other words, the difficulty is that to prove a density result for squarefree integers, we must let S grow to include all primes *before* letting B tend to infinity, but the argument so far shows only that when B is *sufficiently large relative to the primes in* S, then the number of such integers in $[1, B]$ is approximately $B \prod_{p\in S}(1 - p^{-2})$.

One approach that works is to approximate the condition of being squarefree by the condition of being "squarefree as far as the primes $\leq r$ are concerned", and then to estimate the error in this approximation. As $B \to \infty$ for fixed r, the fraction of integers not divisible by p^2 for any prime $p \leq r$ indeed equals $\prod_{\text{prime } p \leq r}(1 - p^{-2})$, by the Chinese remainder theorem. Bounding the error amounts to bounding the fraction of integers in $[1, B]$ divisible by p^2 for a large prime p (that is, a prime $p > r$). More precisely, it remains to show that

$$\lim_{r\to\infty} \lim_{B\to\infty} \left(\frac{\#\{n \leq B : n \text{ is divisible by } p^2 \text{ for some } p > r\}}{B} \right) = 0.$$

This is easy to prove:

$$\begin{aligned}
&\#\{n \le B : n \text{ is divisible by } p^2 \text{ for some } p > r\} \\
&\le \sum_{\text{prime } p > r} \#\{n \le B : n \text{ is divisible by } p^2\} \\
&= \sum_{\text{prime } p > r} \lfloor B/p^2 \rfloor \\
&\le \sum_{\text{integers } n > r} B/n^2 \\
&\le B \int_r^\infty \frac{1}{x^2}\, dx \\
&= B/r,
\end{aligned}$$

so if we divide by B, take the limit as $B \to \infty$, and then take the limit as $r \to \infty$, we get 0.

Thus the density of squarefree integers equals

$$\lim_{r\to\infty} \prod_{\text{prime } p \le r} (1 - p^{-2}) = \prod_{\text{prime } p} (1 - p^{-2}) = \zeta(2)^{-1}.$$

2. Squarefree values of polynomials

For more general problems, the hard part is in bounding the error arising from ignoring the large primes. Consider for instance the following problem: Given a polynomial $f(x) \in \mathbb{Z}[x]$, compute the density of integers n such that $f(n)$ is squarefree. The naïve heuristic above suggests that the answer should be $\prod_{\text{prime } p}(1 - c_p/p^2)$ where c_p equals the number of integers $n \in [0, p^2 - 1]$ for which $p^2 \mid f(n)$.

For fixed r, the density of integers n satisfying the conditions for primes $\le r$ can be computed as before, by using the Chinese remainder theorem. Assuming r exceeds the discriminant of f, Hensel's lemma shows that there are at most $\deg f$ solutions $x \bmod p^2$ to $f(x) \equiv 0 \pmod{p^2}$, so for any primes $p > r$, we can bound the number of integers $n \in [1, B]$ for which $p^2 | f(n)$ by $(\deg f)\lceil B/p^2 \rceil$. But $f(n)$ for $n \le B$ could be as large as (a constant times) $B^{\deg f}$, so we must consider all p up to about $B^{(\deg f)/2}$, and unfortunately the sum of $(\deg f)\lceil B/p^2 \rceil$ over these primes will be small compared to B only if $\deg f \le 2$.

Thus controlling the error is easy only if $\deg f \le 2$. A more complicated argument [Hoo67] shows that the error can be controlled and the predicted density proved correct also in the case $\deg f = 3$, but for irreducible f of degree ≥ 4, there is no known unconditional proof that the conjectural density is correct (except in cases where there is an obstruction coming from a single prime, in which case the density is 0). There is only a theorem of

A. Granville [Gra98] saying that the expected result follows from the *abc* conjecture.

3. A function field analogue

There is an obvious function field analogue of the result about the density of squarefree integers. Namely, for a fixed finite field $\mathbb{F}_q$, one can ask what fraction of elements of the polynomial ring $\mathbb{F}_q[t]$ are squarefree. In this context one defines the *density* of a subset $S \subseteq \mathbb{F}_q[t]$ as the limit

$$\mu(S) := \lim_{d\to\infty} \frac{\#(S \cap \mathbb{F}_q[t]_{\leq d})}{\#\mathbb{F}_q[t]_{\leq d}},$$

if the limit exists, where $\mathbb{F}_q[t]_{\leq d}$ is the set of polynomials in $\mathbb{F}_q[t]$ of degree $\leq d$.

The sieve argument works as before. One need only replace integer primes p by monic irreducible polynomials P of $\mathbb{F}_q[t]$. We find that the density of squarefree elements of $\mathbb{F}_q[t]$ equals

$$\prod_P \left(1 - q^{-2\deg P}\right),$$

which turns out to equal $1-1/q$, as we will explain later using zeta functions.

4. Closed points and zeta functions

To generalize, we will need to reinterpret the set of monic irreducible polynomials in $\mathbb{F}_q[t]$ in geometric terms. Namely, the following sets are in bijection:

(1) the set of monic irreducible polynomials of $\mathbb{F}_q[t]$,
(2) the set of maximal ideals of $\mathbb{F}_q[t]$, and
(3) the set of $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$-orbits in $\mathbb{A}^1(\overline{\mathbb{F}}_q)$.

Namely, given a monic irreducible polynomial, one can take the ideal it generates in $\mathbb{F}_q[t]$, or one can take its set of zeros in $\mathbb{A}^1(\overline{\mathbb{F}}_q)$.

A closed point on a variety (or scheme of finite type) X over $\mathbb{F}_q$ corresponds to a maximal ideal of the affine coordinate ring of an affine open subscheme of X. The closed points of X are in bijection with the $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$-orbits in $X(\overline{\mathbb{F}}_q)$. The *degree* of a closed point is the degree over $\mathbb{F}_q$ of the residue field of the corresponding maximal ideal: it equals the size of the corresponding $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$-orbit. The *zeta function* of X can be defined either as an *Euler product* over closed points, or as a *generating function* for the sequence of integers $\#X(\mathbb{F}_q)$, $\#X(\mathbb{F}_{q^2})$, $\#X(\mathbb{F}_{q^3})$, ...:

$$\zeta_X(s) = Z_X(q^{-s}) := \prod_{\text{closed } P\in X} \left(1 - q^{-s\deg P}\right)^{-1} = \exp\left(\sum_{r=1}^{\infty} \frac{\#X(\mathbb{F}_{q^r})}{r} q^{-rs}\right)$$

for $\mathrm{Re}(s) > \dim X$. The power series $Z_X(T) \in \mathbb{Z}[[T]]$ is (the Taylor series of) a rational function [Dwo60], as was conjectured by A. Weil [Wei49]. The Euler product definition extends also to schemes X of finite type over $\mathbb{Z}$.

The density of squarefree elements of $\mathbb{F}_q[t]$ is $\zeta_{\mathbb{A}^1}(2)^{-1}$ as given by the product definition. The other definition shows that $Z_{\mathbb{A}^1}(T) = 1/(1-qT)$, so this density equals

$$Z_{\mathbb{A}^1}(q^{-2})^{-1} = 1 - qq^{-2} = 1 - 1/q.$$

5. Smooth plane curves

We now consider a more geometric problem. What is the density of homogeneous polynomials $f \in \mathbb{F}_q[x,y,z]$ such that the plane curve $f = 0$ in $\mathbb{P}^2$ is smooth (of dimension 1)? (Density is defined as the limit as $d \to \infty$ of the fraction of the degree-d homogeneous polynomials which satisfy the desired condition.)

Smoothness can be tested locally. Therefore we will start with all homogeneous polynomials f of degree d and sieve out, for each closed point $P \in \mathbb{P}^2$, those f for which the curve $f = 0$ has a singularity at P. The condition that $f = 0$ has a singularity at P amounts to 3 linear conditions on the Taylor coefficients of a dehomogenization $\bar{f}$ of f at P (namely, the vanishing of $\bar{f}$ and its two partial derivatives at P), and these linear conditions are over the residue field of P. It follows that the density of f such that $f = 0$ has a singularity at P equals $q^{-3\deg P}$. This suggests the guess that the density of f defining a smooth plane curve equals

$$\prod_{\text{closed } P \in \mathbb{P}^2} \left(1 - q^{-3\deg P}\right) = \zeta_{\mathbb{P}^2}(3)^{-1}$$
$$= (1-q^{-1})(1-q^{-2})(1-q^{-3}),$$

where the last equality comes from substituting $T = q^{-3}$ in

$$Z_{\mathbb{P}^2}(T)^{-1} = (1-T)(1-qT)(1-q^2T).$$

Taking $q = 2$ gives $21/64$.

The guess turns out to be correct, although the proof is much more difficult than the proof for squarefree integers or polynomials. To control the error arising from ignoring the conditions from closed points of high degree, we exploit the fact that p-th powers in characteristic p have derivative 0 in order to decouple the partial derivatives; the argument also uses Bézout's theorem. See [Poo04] for details.

6. Bertini theorems

A generalization of the argument of the previous section yields a "Bertini smoothness theorem" over finite fields.

Suppose first that k is an *infinite* field. The Bertini smoothness theorem says that if a subvariety $X \subseteq \mathbb{P}^n$ over k is smooth, then for a sufficiently general hyperplane $H \subset \mathbb{P}^n$, the variety $H \cap X$ is smooth too. "Sufficiently general" here means inside a Zariski dense open subset U of the dual projective space that parametrizes hyperplanes in $\mathbb{P}^n$. Since k is infinite, the set $U(k)$ is nonempty, so there exists a hyperplane H over k with $H \cap X$ smooth. But if k is finite, this last result can fail: for example, if X is the hypersurface

$$\sum_{i=1}^{n+1}(X_i^q Y_i - X_i Y_i^q) = 0$$

in $\mathbb{P}^{2n+1}$ over $k = \mathbb{F}_q$, then every hyperplane over k is tangent to X somewhere [Kat99, pp. 621–622].

N. Katz [Kat99, Question 13] asked whether the Bertini theorem over finite fields could be salvaged by allowing hypersurfaces of unbounded degree in place of hyperplanes. The closed point sieve yields such a result, and even gives an asymptotically positive fraction of good hypersurfaces of degree d, as $d \to \infty$. (The existence of a good hypersurface, for d sufficiently large and divisible by the characteristic p, was shown independently by O. Gabber [Gab01, Corollary 1.6].)

The result is that if X is a smooth quasiprojective subvariety of $\mathbb{P}^n$ of dimension m over $\mathbb{F}_q$, then the density of f such that $\{f = 0\} \cap X$ is smooth of dimension $m-1$ equals $\zeta_X(m+1)^{-1}$ [Poo04].

Perhaps surprisingly, the density is an intrinsic property of X, independent of how X is embedded in projective space. Taking $X = \mathbb{A}^1 \subseteq \mathbb{P}^1$, we recover the result that the density of squarefree polynomials in $\mathbb{F}_q[t]$ equals $\zeta_{\mathbb{A}^1}(2)^{-1}$.

Here are a few applications of the Bertini theorem and its variants:

- *Space-filling curves* (answering questions of N. Katz [Kat99]): Given a smooth projective geometrically irreducible variety X of positive dimension over $\mathbb{F}_q$, there exists a smooth projective geometrically irreducible curve $Y \subseteq X$ passing through all the $\mathbb{F}_q$-points of X.
- *Space-avoiding varieties*: Given X as above, and an integer y satisfying $1 \leq y < \dim X$, there exists a smooth projective geometrically irreducible variety $Y \subseteq X$ of dimension y such that $Y(\mathbb{F}_q) = \emptyset$.
- *Abelian varieties as quotients of Jacobians*: For every nontrivial abelian variety A over $\mathbb{F}_q$, there is a smooth projective geometrically irreducible curve Y in A such that the induced map from the Jacobian of Y to A is surjective.
- *Brauer groups of surfaces*: Q. Liu, D. Lorenzini, and M. Raynaud [LLR05] used the Bertini theorem (and several other ingredients) to

show that if X is a smooth projective geometrically irreducible surface over $\mathbb{F}_q$, then the order of $\operatorname{Br} X$ is a perfect square.

The Bertini theorem also has a conjectural arithmetic analogue: If X is a quasiprojective subscheme of $\mathbb{P}^n_{\mathbb{Z}}$ that is regular of dimension m, then the density (suitably defined) of $f \in \mathbb{Z}[x_0, \ldots, x_n]$ such that $\{f = 0\} \cap X$ is regular of dimension $m-1$ equals $\zeta_X(m+1)^{-1}$. This is proved in [Poo04] assuming the *abc* conjecture and one other conjecture, by making use of a multivariable extension [Poo03] of Granville's conditional result [Gra98] on squarefree values of polynomials. This statement implies both the finite field Bertini theorem and the fact that the squarefree integers have density $\zeta(2)^{-1}$ (take $X = \operatorname{Spec} \mathbb{Z}$ in $\mathbb{P}^0_{\mathbb{Z}} = \operatorname{Spec} \mathbb{Z}$).

7. Whitney embedding theorems

If X is a smooth projective curve over an infinite field k, then there is a closed immersion $X \hookrightarrow \mathbb{P}^3$. To prove this, one starts with X in some large projective space $\mathbb{P}^N$, and iteratively performs projections. One shows that if $N > 3$, then composing the embedding $X \hookrightarrow \mathbb{P}^N$ with a sufficiently generic projection $\mathbb{P}^N \dashrightarrow \mathbb{P}^{N-1}$ yields an embedding $X \hookrightarrow \mathbb{P}^{N-1}$.

The analogous statement for a finite field $\mathbb{F}_q$ is false. There are some obvious obstructions to embedding a smooth curve X in $\mathbb{P}^3$. Namely, it can happen that X has more points over $\mathbb{F}_q$ than $\mathbb{P}^3$ does! Even if $\#X(\mathbb{F}_q) \leq \#\mathbb{P}^3(\mathbb{F}_q)$, it could happen that X has more closed points of degree 2 than $\mathbb{P}^3$ does.

N. Nguyen [Ngu05] used the closed point sieve to prove that the obvious obstructions are the only ones. Namely, he proved that given a smooth curve X over $\mathbb{F}_q$ and an integer $n \geq 3$, there exists a closed immersion $X \hookrightarrow \mathbb{P}^n$ if and only if for every $e \geq 1$ the number of closed points of degree e on X is less than or equal to the number of closed points of degree e on $\mathbb{P}^n$. In fact, he also proved the higher-dimensional analogue: given a smooth variety X of dimension m and an integer $n \geq 2m+1$, there is a closed immersion $X \hookrightarrow \mathbb{P}^n$ if and only if the conditions on the number of closed points are satisfied. This proof was even more involved than the proof of the Bertini theorem, because the conditions on homogeneous polynomials $f_0, \ldots, f_n$ for the rational map $(f_0 : \cdots : f_n) \colon \mathbb{P}^N \dashrightarrow \mathbb{P}^n$ to restrict to a closed immersion $X \hookrightarrow \mathbb{P}^n$ are not local, as were the conditions defining smoothness. Nguyen had to sieve over *pairs* of closed points to get his result.

These embedding results are algebraic analogues of the *Whitney embedding theorem*, which states that every m-dimensional real manifold X can be embedded in $\mathbb{R}^{2m+1}$ (in fact, Whitney proved that $\mathbb{R}^{2m}$ suffices, but his methods for this stronger result are not algebraic, and indeed this result fails in the algebraic setting, even over infinite fields).

8. Lefschetz pencils

One fruitful way to study a variety $X \subseteq \mathbb{P}^n$ is to choose a dominant rational map $(f : g) \colon X \dashrightarrow \mathbb{P}^1$, say defined by a pair of homogeneous polynomials $f, g \in k[x_0, \ldots, x_n]$ of the same degree. The fibers (after blowing up the indeterminacy locus) form a family of hypersurface sections in X, namely $\{\lambda_1 f - \lambda_0 g = 0\} \cap X$ for $(\lambda_0 : \lambda_1) \in \mathbb{P}^1$. Ideally, questions about X can then be reduced to questions about these hypersurface sections, which are of lower dimension.

Unfortunately, even if X is smooth and the rational map is chosen generically, some of the hypersurfaces sections may fail to be smooth. The best one can reasonably expect is that there will be at most finitely many singular fibers, and each such fiber has the simplest kind of singularity. More precisely, $(f : g) \colon X \dashrightarrow \mathbb{P}^1$ defines a *Lefschetz pencil* if all of the following hold (after base extension to an algebraically closed field):

(1) The *axis* $f = g = 0$ intersects X transversely.
(2) All but finitely many hypersurface sections in the family are smooth.
(3) Each non-smooth hypersurface section has only one singularity, and that singularity is an ordinary double point.

Over an infinite field k, a dimension-counting argument proves the existence of Lefschetz pencils for any smooth $X \subseteq \mathbb{P}^n$: see [Kat73]. This was famously used by P. Deligne to prove the Riemann hypothesis for varieties over finite fields [Del74, Del80]: for his application, he had the freedom to enlarge the ground field if necessary, so he needed only the existence of Lefschetz pencils over an algebraic closure of a finite field.

In any case, the question remained as to whether Lefschetz pencils over k existed for varieties over k in the case where k is finite. N. Nguyen [Ngu05] proved such an existence result using the closed point sieve. Again, because the conditions in the definition of Lefschetz pencil are not all local, he had to sieve over pairs of closed points.

9. Questions

(1) There seems to be a general principle that if an existence result about polynomials or n-tuples of polynomials over an infinite field can be proved by dimension counting, then a corresponding result over finite fields can be proved by the closed point sieve. Can this principle be formalized and proved?
(2) The closed point sieve we have discussed is the geometric analogue of the simplest kind of sieve appearing in analytic number theory. Are there also geometric analogues of more sophisticated sieves like the *large sieve*, and do these have applications?

(3) What other theorems currently require the hypothesis "Assume that k is an infinite field"? Hopefully the closed point sieve could be used to eliminate the hypothesis in many of these.

References

[Del74] Pierre Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307 (French). MR0340258 (49 #5013)

[Del80] Pierre Deligne, *La conjecture de Weil. II*, Inst. Hautes Études Sci. Publ. Math. (1980), no. 52, 137–252 (French). MR601520 (83c:14017)

[Dwo60] Bernard Dwork, *On the rationality of the zeta function of an algebraic variety*, Amer. J. Math. **82** (1960), 631–648. MR0140494 (25 #3914)

[Gab01] O. Gabber, *On space filling curves and Albanese varieties*, Geom. Funct. Anal. **11** (2001), no. 6, 1192–1200. MR1878318 (2003g:14034)

[Gra98] Andrew Granville, *ABC allows us to count squarefrees*, Internat. Math. Res. Notices (1998), no. 19, 991–1009. MR1654759 (99j:11104)

[Hoo67] C. Hooley, *On the power free values of polynomials*, Mathematika **14** (1967), 21–26. MR0214556 (35 #5405)

[Kat73] Nicholas M. Katz, *Pinceaux de Lefschetz: théorème d'existence*, Groupes de monodromie en géométrie algébrique. II, Springer-Verlag, Berlin. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II); Dirigé par P. Deligne et N. Katz, Lecture Notes in Mathematics, Vol. 340, Exposé XVII, 1973, pp. 212–253.

[Kat99] ———, *Space filling curves over finite fields*, Math. Res. Lett. **6** (1999), no. 5-6, 613–624. MR1739219 (2001e:11067)

[LLR05] Qing Liu, Dino Lorenzini, and Michel Raynaud, *On the Brauer group of a surface*, Invent. Math. **159** (2005), no. 3, 673–676. MR2125738

[Ngu05] Nghi Huu Nguyen, *Whitney theorems and Lefschetz pencils over finite fields*, May 2005. Ph.D. thesis, University of California at Berkeley.

[Poo03] Bjorn Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J. **118** (2003), no. 2, 353–373. MR1980998 (2004d:11094)

[Poo04] ———, *Bertini theorems over finite fields*, Annals of Math. **160** (2004), no. 3, 1099–1127.

[Wei49] André Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508. MR0029393 (10,592e)

Bjorn Poonen
Department of Mathematics
University of California
Berkeley, CA 94720-3840, USA
E-mail : poonen@math.berkeley.edu
URL: http://math.berkeley.edu/~poonen