The Arithmetic Behind Cryptography

Gerhard Frey

The security of very efficient and widely used public key crypto systems is based on the hardness of mathematical problems. Typically such problems come from arithmetic. Here are three important examples: Find shortest or closest vectors in lattices, factor large numbers, and **compute logarithms in finite groups**.

In this article we shall concentrate on the last example and so cover crypto systems for which the crypto primitive behind them is the discrete logarithm (DL) in cyclic groups of prime order (see the subsection on Diffie-Hellman Problems).

The first proposal for such systems was given by Diffie and Hellman in their groundbreaking article [DH]. As groups they suggested taking roots of unity in the multiplicative group of finite fields.

The generality of the methods provided by algorithmic arithmetic geometry opens immediately a wide range of possibilities. One can replace torsion points in the multiplicative group by torsion points of Jacobian varieties of curves over finite fields. But, on the other side, the strength of the methods allows us to develop very efficient attacks. So most of the suggested candidates for public key systems did not fulfill the expectations, and only DL-systems based on carefully chosen elliptic curves and curves of genus 2 survived without any blame.

This does not mean that the study of curves of arbitrary genus is not important for applications in data security. In many cases we understand partial weaknesses of elliptic curves by making more general objects accessible to computation. The continuous study of consequences of advances in algorithmic arithmetic geometry for the security of used crypto system and failures of attacks give mathematicians a better conscience and users more trust. So even people only involved in designing systems without being interested in the theoretical background can choose (very special) cases, for example, one elliptic curve over a fixed field with explicit addition formulas given in a list of standardized curves, for instance listed in [NIST] or in [BRAIN]. But apart from applications to elliptic curves the higher genus curves have various applications in cryptography which we cannot describe in this short survey.

But it is not only the status quo which is supported. New points of view from the theoretical side lead to advances in the design of hardware as well as in protocols. One of the striking examples is the development of pairing-based cryptography. From its background, namely duality theory in arithmetic geometry, there goes a direct path to very efficient implementations of pairings which allow, for instance, new ways to sign, and here curves of higher genus may play an important role.

Acknowledgment. The author would like to thank the referees for careful reading of the manuscript and for their helpful comments.

Gerhard Frey is professor of mathematics at the University of Duisburg-Essen. His email address is gerhard.frey@gmail.com.

Some Aspects of Arithmetic Geometry

In the section "Construction of DL-Systems" we shall formulate tasks for mathematicians motivated by needs of data security. It turns out that it is surprisingly difficult to find families of groups which are candidates for DL-systems, and that the search for bilinear structures is even more involved.

The only known examples are constructed with the help of advanced methods of arithmetic geometry mostly developed during the last sixty years. We emphasize the remarkable fact that they both enable us to solve old problems like FLT (see the subsection "Digression: FLT") and lead to efficient and secure families of public key crypto systems.

What Is Arithmetic Geometry?

Arithmetic geometry is one of the most powerful ingredients in mathematics. It combines classical algebraic number theory with algebraic geometry. It uses the theory of functions over $\mathbb{C}$, and so analytic geometry, and it transfers this theory to its p-adic counterpart, the p-adic rigid geometry.

The important feature is that objects from number theory, like the ring of integers, and

objects from algebraic geometry, like varieties over finite fields, can be treated in a unified way (as schemes consisting of the set of points with topology and sheaves of functions). For instance, the arithmetic of rings of integers in number fields is very similar to the arithmetic of rings of holomorphic functions on affine curves over finite fields. The analogy is neither only formal nor in all aspects obtained by using a dictionary, and the interplay between the arithmetic world and the geometric world is extremely fruitful for both sides.

The situation becomes very interesting and extremely difficult if both points of view are mixed together, for instance, if we look at the arithmetic of curves C defined over number fields or p-adic fields K. Geometrically these are varieties of dimension 1, but since we can look at them as being defined over the ring of integers O_K of K, they carry a 2-dimensional structure: from C we get an arithmetical surface $\mathcal{C}$. This surface contains for each prime ideal $\mathfrak{p}$ of O_K a closed fiber $C^{\mathfrak{p}}$ (special fiber at $\mathfrak{p}$) which is the reduction of C modulo $\mathfrak{p}$, that is, roughly speaking, the curve obtained by looking at the equations defining C (in a suitable normalization with respect to $\mathfrak{p}$) modulo $\mathfrak{p}$. The arithmetical surface $\mathcal{C}$ contains much more information than its generic fiber C. It is not uniquely determined by C. There is an optimal model, the so-called minimal model, and using this model one can try to get the arithmetical data of C from studying the analogous data of the special fibers. In the case that K is a number field one tries to exploit the local information one gets over the completions at all places of K simultaneously in order to get global information, for example, about rational points on C. (If $K = \mathbb{Q}$ these completions are the reals $\mathbb{R}$ and, for all prime numbers p, the p-adic numbers $\mathbb{Q}_p$.) If this strategy is successful then one has established a local-global principle. One famous example of such a principle is the theorem of Hasse-Minkowski which states that one quadratic polynomial in arbitrarily many variables with coefficients in a number field K has a K-rational solution if and only if it has solutions in all fields obtained as completions with respect to valuations of K.

We cannot expect to get such a principle for general varieties. In fact we already find counterexamples if we look at the set of solutions of two quadratic equations or of polynomials in two variables of degree 3. But there is a Galois theoretical variant which relates local with global information: the density theorem of Čebotarev (Theorem 2.5).

Algorithmic Arithmetic Geometry

Classically algorithmic aspects of number theory mostly deal with lattices and derived objects. A fundamental tool is Minkowski's theorem on points with small norms in lattices and related results, for instance reduction of quadratic forms following Lagrange and Gauss. The enormous growth of computational power made it possible to construct interesting examples in a wide range, and very often one meets the LLL algorithm as a major tool.

The theoretical insights obtained by the approach described in the preceding subsection made rapid and exciting progress possible in the area of algorithmic arithmetic geometry, generalizing considerably both range and techniques of computational number theory. Prominent examples are computation of tables of modular forms including congruences, algorithmic study of modular curves (see, for instance, the Cremona tables [C] listing elliptic curves) and related Galois representations.

Translating arithmetical problems into the geometric language has the immediate consequence that one can apply the methods from arithmetic to the geometric case, too. And so we have now a very advanced theoretical and algorithmic toolkit to deal with the **explicit theory** of varieties over finite fields as a counterpart to the explicit theory of algebraic number fields. We devote the subsection "Arithmetic in Divisor Classes" to an important example.

Complexity Hierarchy. A crucial part of every algorithmic theory is the determination of the complexity of the available algorithms.

Here we can only scratch the surface of this fascinating mathematical subject. We introduce Landau's notation:

For

$$f, g : \mathbb{N} \to \mathbb{R}$$

with g positive define

$$f = \mathcal{O}(g)$$

if there exists $d \in \mathbb{R}_{>0}$ with

$$| f(N) | \leq d g(N)$$

for all N.

Take $\alpha \in [0,1], c \in \mathbb{R}_{>0}$.

Define

$$L_N(\alpha, c) := \exp(c \cdot \log(N)^{\alpha} \cdot \log(N)^{1-\alpha}).$$

For (almost all) $N \in \mathbb{N}$, let

$$f_N : A_N \to B_N$$

be maps from sets A_N to sets B_N. Assume that there is an algorithm which evaluates f_N with (probabilistic) complexity (e.g., number of bit operations needed) $F(N)$.

Then the (probabilistic) asymptotic complexity of the family f_N is called

- *polynomial* if $F = \mathcal{O}(L_N(0, c))$ ("*fast algorithm*")

- *exponential* if $F = \mathcal{O}(L_N(1,c))$ ("*hard algorithm*") and
- *subexponential* if there is $0 < \alpha < 1$ with $F = \mathcal{O}(L_N(\alpha,c))$

in $\log(N)$.

Subexponential complexity is a very interesting case between the two extremes.

Caution: This notion of complexity is an asymptotic estimate of a specific algorithm for the evaluation of f_N. In particular, it is only an upper bound for the hardness of the evaluation of f_N.

Nevertheless it gives a good impression of what one can expect for given instances with concrete N large enough.

Examples for algorithms with polynomial complexity are

- the (extended) Euclidean algorithm and
- exponentiation in groups (expressed in costs for group operation).

The first example implies that the computation of the greatest common divisor of numbers and polynomials as well as the computation of the inverse in finite groups with known group order is of polynomial complexity.

From the second example it follows that exponentiation in finite fields $\mathbb{F}_q$ is, as a function in $\log(q)$, polynomial. The same is true for scalar multiplication in elliptic curves (see the subsection "Curves of Genus 1: Elliptic Curves"). But a highly nontrivial and much more general result is explained in the next subsection.

Arithmetic in Divisor Classes. We take a (projective absolutely irreducible nonsingular) curve C of genus g defined over a field K which has no inseparable algebraic extensions. This means that irreducible polynomials over K have no multiple zeros. Examples are all fields of characteristic 0 and all finite fields.

We assume that we have a K-rational point P_∞ on C and denote by O_C the ring of functions on C which have only poles in P_∞.

As an example take C as projective line. Then O_C is isomorphic to the ring of polynomials in one variable.

The ring O_C is a Dedekind domain and so every ideal $\neq \{0\}$ is, in a unique way, the product of powers of prime ideals.

The quotient field $F_C = Quot(O_C)$ is the function field of C and is independent of the choice of P_∞.

We generalize the notion of ideals of O_C to ideals of F_C by defining: $A \subset F_C$ is an ideal if there is an element $f \in F_C^*$ such that $f \cdot A \subset O_C$ is an ideal of O_C in the usual sense.

The set of ideals of F_C is a commutative group $I(O_C)$ which is freely generated by the set of prime ideals of O_C.

Inside of $I(O_C)$ we have the subgroup $\mathrm{Princ}(O_C)$ of principal ideals $f \cdot O_C$, $f \in F_C^*$. The quotient group

$$\mathrm{Pic}(O_C) := I(O_C)/\mathrm{Princ}(O_C)$$

is the ideal class group of O_C. It is in a natural way isomorphic to $\mathrm{Pic}^0(C)$, the divisor class group of degree 0 of C, and a fundamental theorem of the theory of curves states that $\mathrm{Pic}^0(C)$ is in a functorial way isomorphic to the group of K-rational points of the Jacobian variety J_C of C. This is an abelian variety (i.e., a geometrically connected projective commutative group scheme) of dimension g over K.

The crucial point is that we can add in an explicit way in $J_C(K) = \mathrm{Pic}^0(C) = \mathrm{Pic}(O_C)$! To see this one recalls that the algebraic structure of O_C is similar to the structure of the rings of integers of number fields and that $Pic(O_C)$ is analogous to the ideal class groups of number fields. Computing in these groups is one of the major tasks of computational number theory, and it is done effectively because of Minkowski's theorem on points with small norm in lattices. An immediate consequence is that class groups of number fields are finite.

In the geometric frame Minkowski's theorem is replaced by the even more fundamental theorem of Riemann-Roch. The size of the discriminant of K is replaced by the genus g of C. Again we get as an immediate consequence that for finite fields $K = \mathbb{F}_q$ (the field with q elements) the group $\mathrm{Pic}^0(C)$ is a finite group. In fact, we get a rather sharp estimate from below and above for this size depending on q and g in the subsection "The Local Information".

We state an important and amazing recent result concerning the computation of the sum of two elements in $\mathrm{Pic}^0(C)$.

Theorem 2.1 (Hess, Diem)**.** *Let C be a curve of genus g over the field $\mathbb{F}_q$. The addition and inversion in the divisor class group of degree 0 of C can be performed by an explicitly given algorithm in an expected number of bit operations which is polynomially bounded in g and $\log(q)$, i.e., group operations and scalar multiplication in divisor class groups are of polynomial complexity both in g (with fixed q) and $\log(q)$ (with fixed g).*

Curves of Genus 1: Elliptic Curves. We apply the Riemann-Roch theorem to the special case that C has genus 1 and a rational point P_∞ and come to a well-known object: elliptic curves. We get:

- $E(K)$ is in a natural way an abelian group with neutral element P_∞.
- E is as a variety isomorphic to its Jacobian variety, and so elliptic curves with chosen point P_∞ are abelian varieties of dimension 1.

- The addition in $E(K)$ is given by the following rule: $R = P \oplus Q$ is the unique point on $E(K)$ for which there exists a function in F_E with zeroes of order 1 in P and Q (respectively a zero of order 2 if $P = Q$) and poles of order 1 in P_∞ and R (and of order 2 in P_∞ if $R = P_\infty$).
- We find projective coordinates such that E is given by a cubic homogenous equation

  $$E : Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

  (called the Weierstrass equation) without singular points. The only point with Z-coordinate equal 0 is $P_\infty = (0, 1, 0)$. All other points can be given by affine coordinates as $P = (x, y)$ with $y^2 + a_1xy + a_3y = x^3 + a_2x + a_4x + a_6$.
- If char(K) is prime to 6 we find a short Weierstrass equation

  $$E : Y^2Z = X^3 + AXZ^2 + BZ^3$$

  with $A, B \in K$ and discriminant $\Delta_E := 4A^3 + 27B^2 \neq 0$.

Using the short Weierstrass equation we can describe addition geometrically by the following rule: Take the line through P and Q (take the tangent line if $P = Q$) and compute the third intersection point $-R$ with E. Then R is the point obtained from $-R$ by reflection at the x-axis.

From this description one easily deduces formulas for the addition on $E(K)$: For

$$P_1 = (x_1, y_1, 1), \qquad P_2 = (x_2, y_2, 1)$$

we get "in general"

$$P_3 = (x_3, y_3, 1) := P_1 \oplus P_2$$

with

$$x_3 = -(x_1 + x_2) + ((y_1 - y_2)/(x_1 - x_2))^2.$$

For doubling points there is another explicit formula.

This is a short and simple formula. But, because of its importance, a lot of work has been done to make addition even faster, by using appropriate coordinates, appropriate coefficients, and appropriate normal forms for equations. This is described in full detail in [ACF], Chapter 13.

Remark 2.2. It is obvious that we do not need deep theory to get Theorem 2.1 for elliptic curves. But in many cases one can transfer the group structure of elliptic curves to the addition in divisor class groups of more general curves. By the theorem we know that the complexity hierarchy of the group operations is not changed. This is important for the analysis of attacks to crypto systems (the subsection "Attacks").

Galois Representations Attached to Abelian Varieties

In arithmetic geometry we want to find properties of objects over arithmetically interesting fields K such as number fields, p-adic fields, finite fields. The methods used from algebraic geometry and analysis work better over separably closed fields. The bridge between the two concepts is built by the action of the absolute Galois group $G_K = Aut(K_s/K)$ (K_s the separable closure of K), always assumed to be continuous with respect to the profinite topology on G_K.

Very often this action is studied on free modules over appropriate rings R like $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}_\ell$, the ℓ-adic numbers, and leads to *Galois representations.*

Elements of finite order in abelian varieties A of dimension d are a main source for such representations. A basic result is that, for natural numbers n prime to char(K), the group of torsion points of order n

$$A[n] := \{P \in A(K_s);\ n \cdot P = 0\}$$

is an abelian group isomorphic to $(\mathbb{Z}/n\mathbb{Z})^{2d}$. Hence, after choosing a $\mathbb{Z}/n\mathbb{Z}$-base of $A[n]$, the action of $\sigma \in G_K$ is given by a matrix, and we get a continuous homomorphism

$$\rho_{A,n} : G_K \to M_{2d}(\mathbb{Z}/n) \text{ with } M_{2d} = \text{set of } 2d \times 2d\text{-matrices.}$$

In particular, one can attach to curves C of genus g Galois representations of dimension $2g$ induced by the action of G_K on $J_C[n]$ or, nearer to computation, on the elements of order dividing n in the divisor class group of degree 0 of C regarded as a curve over K_s.

Example 2.3. Let E/K be an elliptic curve and $n \in \mathbb{N}$ prime to the characteristic of K. The action of G_K on $E[n]$ induces a 2-dimensional Galois representation $\rho_{E,n}$ over $\mathbb{Z}/n\mathbb{Z}$.

For $\sigma \in G_K$ the characteristic polynomial is defined by

$$\chi_{\rho_{E,n}(\sigma)}(T) = T^2 - Tr(\rho_{E,n}(\sigma))T + \det(\rho_{E,n}(\sigma)).$$

In many important cases $\rho_{E,n}$ is semisimple and hence it is determined by its characteristic polynomials.

ℓ-adic version. Take a prime ℓ different from char(K) and define the ℓ =adic Tate module of an abelian variety A (e.g., $A = J_C$) by

$$T_\ell(A) := \overleftarrow{lim}\, A[\ell^k].$$

G_K acts continuously (with respect to the ℓ-adic topology) on $T_\ell(A)$ and induces a representation $\tilde{\rho}_{A,\ell}$ over $\mathbb{Z}_\ell$ or, by tensoring with $\mathbb{Q}$, over the field of ℓ-adic numbers.

In particular, we can attach ℓ-adic representations to the divisor class groups of curves.

Galois Representations in Arithmetical Geometry

In this subsection K is a number field, that is, a finite algebraic extension of $\mathbb{Q}$. In number theory we have a hierarchy of fields: The number field K carries various topologies induced by valuations ν which extend the p-adic valuations and the absolute value on $\mathbb{Q}$.

The completion of K with respect to the topology induced by one ν is the local field K_ν which is an algebraic extension field of p-adic numbers $\mathbb{Q}_p$ or an extension field of $\mathbb{R}$. It contains K as a dense subfield.

If ν is an extension of a p-adic valuation then the ring of integers O_K of K is contained in the valuation ring of ν and dense in the ring O_ν of ν-adic integers of K_ν. The residue field of ν is the finite field $\mathbb{F}_\nu$ obtained as quotient of O_K or, equivalently, of O_ν modulo the maximal ideal m_ν of ν. It is a finite algebraic extension of $\mathbb{F}_p$ and so isomorphic to $\mathbb{F}_q$ with $q = p^d$.

Remark 2.4. In *number theory* one tries to solve problems over global fields by looking at them over (all) local fields and then reducing them modulo ν to problems over finite fields. One hopes that one does not lose too much information (local-global-principle), see "What Is Arithmetic Geometry?"

In *cryptography* one is mainly interested in objects over *finite fields* but by studying them one is often led to problems over local fields ("lifting") and even over global fields.

The hierarchy of fields is reflected by the hierarchy of Galois groups.

The global Galois group G_K is big and complicated. It is studied by restriction to subgroups G_ν which consist of elements of G_K acting continuously with respect to the ν-topology. G_ν can be identified with the Galois group G_{K_ν} whose structure is much simpler. In particular, it has a canonical quotient group which is the Galois group of the maximal unramified extension of K_ν in its separable closure. This quotient is canonically isomorphic to $G_{\mathbb{F}_q}$ and so topologically generated by the Frobenius automorphism ϕ_q mapping elements of $\mathbb{F}_{q,s}$ to their qth power.

Via these identifications one can define (conjugacy classes of) Frobenius elements $\sigma_\nu \in G_K$ attached to each ν.

We come back to the Galois representations attached to torsion points, respectively, Tate modules of abelian varieties A.

A consequence of the arithmetic of abelian varieties is that the fixed field $K_s^{ker(\tilde{\rho}_{A,\ell})}$ of the kernel of $\tilde{\rho}_{A,\ell}$ is ramified only in places ν of K dividing either ℓ or at which A has bad reduction. Hence almost all places of K are unramified. Representations satisfying this condition are called geometric. It has turned out that the study of these representations is the key to the great results in number theory achieved during the last thirty years, and, at the same time, it is crucial for finding DL-systems.

Having a geometric representation ρ, we define S_ρ as the finite set of places of K which consist of all extensions of the absolute value and of all places which ramify in $K_s^{ker(\rho)}/K$. For every place $\nu \notin S_\rho$ we can choose a Frobenius element σ_ν. The image of ρ at σ_ν determines the restriction of ρ to G_{K_ν}. It gives the local information about ρ at ν.

Looking at all $\nu \notin S_\rho$ we can bundle this local information. The next big result tells us that we have a local-global principle for semisimple geometric representations.

Theorem 2.5 (Čebotarev's Density Theorem)**.** *Let ρ be a geometric Galois representation of G_K.*

If ρ is semisimple, then ρ is determined by

$$\{\chi_{\rho(\sigma_\nu)}(T); \nu$$

*runs over places of K not contained in S_ρ}.

It is even allowed to omit finitely many additional places of K.

Remark 2.6. To demonstrate the power of this statement we remark that the proof of Mordell's conjecture by Faltings follows from his result that for prime numbers ℓ and all abelian varieties A defined over K the representation $\tilde{\rho}_{A,\ell}$ is semisimple.

The Local Information

We fix a finite field $\mathbb{F}_q$ and take a curve C of genus g defined over $\mathbb{F}_q$. The following result is a landmark (established $\sim$ 1930–40) in arithmetical geometry.

Theorem 2.7 (Weil)**.** *There is a monic polynomial $\chi_C(T) \in \mathbb{Z}[T]$ such that:*

- *All zeroes of $\chi_C(T)$ have (complex) value $\sqrt{q}$.* [1]
- *For all n, the characteristic polynomial of the Frobenius automorphism ϕ_q under the representation $\rho_{J_C,n}$ is congruent to $\chi_C(T)$ modulo n.*
- *For all $\ell \neq p$, the characteristic polynomial of the Frobenius automorphism ϕ_q under the representation $\tilde{\rho}_{J_C,\ell}$ is equal to $\chi_C(T)$.*

By linear algebra we see that $|J_C(\mathbb{F}_q)| = |Pic^0(C)| = \chi_C(1)$.

As consequence we get that $|Pic^0(C) - q^g| = \mathcal{O}(q^{g-1/2})$.

Corollary 2.8. *Let E be an elliptic curve over $\mathbb{F}_q$.*

Then $a_q := |\ E(\mathbb{F}_q)\ | = q + 1 - Tr(\tilde{\rho}_{E,\ell})$ and $|\ |\ E(\mathbb{F}_q)\ | - q - 1| \leq 2\sqrt{q}$ (Hasse bound).

[1] *This is an analogue of the Riemann hypothesis in number theory and so it is sometimes called the "Riemann hypothesis" for curves though it is a proven result.*

The Global Information

We are very short on space here and consider only a special case: $K = \mathbb{Q}$ and $C = E$ an elliptic curve. In the preceding subsection we have computed the characteristic polynomial of the Frobenius endomorphism ϕ_p at least when p is not dividing Δ_E.

We bundle the local information and form the **global L-series** of E:

$$L_E(s) := f^*(s) \cdot \prod_{p \text{ prime to } \Delta_E} (1-(p+1-a_p)p^{-s}+p^{1-s})^{-1}$$

where $f^*(s)$ is a rational function which takes care of bad primes and a_p is, as above, the order of $E(\mathbb{F}_p)$.

This is a Dirichlet series, and the conjecture of Hasse-Taniyama was that it could be extended to an analytic function in the complex plane.

This was known since about the 1950s from results of M. Deuring in the case that E has complex multiplication (CM) and yields an important tool for cryptography.

Digression: FLT. Though it has nothing to do with cryptography, we cannot resist hinting at how to prove FLT.

The conjecture of Hasse-Taniyama has been proved by A. Wiles ([W]) in a relevant special case; in fact he proved the conjecture of Shimura-Taniyama-Weil for semistable elliptic curves:[2] $L_E(s)$ is the Euler product attached to a **modular form of weight** 2**!**

Proof of FLT in Eight Lines

Assume that $A^p - B^p = C^p$ for $p \geq 5$.

I suggested to look at the elliptic curve $E : Y^2 = X(X - A^p)(X - B^p)$.

By global information (Wiles) $\rho_{E,p}$ is attached to a modular form. By local information about E and the theorem of Ribet (Serre's conjecture) $\rho_{E,p}$ is attached to a modular form $\neq 0$ of level 2 and weight 2. Such a form does not exist!

Construction of DL-Systems

We want to explain how the methods sketched in the preceding section can be used for cryptography.

We introduce very shortly the requirements coming from the needs of public key cryptography based on discrete logarithms.

[2] *The general proof was given by Breuil, Conrad, and Diamond. The conjecture is a special case of* Serre's conjecture, *which is now proved, too.*

Diffie-Hellman Problems

Let G be a group of known order n.

We can formulate a *computational problem*:

DHCP: For randomly chosen $a, b \in \{1, \ldots, n\}$, $g \in G$ and given $g_1 = g^a, g_2 = g^b$, compute $g^{a \cdot b}$.

DHCP is called the Diffie-Hellman computational problem.

It is obvious that we can solve DHCP if we can solve the following task:

For randomly chosen $g_1, g_2 \in G$ decide whether g_2 lies in the cyclic group generated by g_1, and if so, compute $k \in \mathbb{N}$ with

$$g_2 = g_1^k.$$

The residue class of such a k modulo n is the discrete logarithm (DL) $log_{g_1}(g_2)$. Highly nontrivial is a kind of converse: there is a subexponential algorithm due to Maurer-Wolf that can compute the (DL) in G if one knows how to solve DHCP. Thus, the complexity of (DL) is an upper bound for the complexity of DHCP and, up to subexponential algorithms, the crypto primitive determining security of the Diffie-Hellman key exchange and encryption, as well as of the El Gamal signature ([ACF]), is the discrete logarithm.

By elementary number theory (Chinese remainder theorem and p-adic expansion of numbers) one sees immediately that without loss of generality we can and shall assume that n is a prime number ℓ. So G is a cyclic group generated by an element g_0, and (DL) is equivalent to the computation of $log_{g_0}(g)$ for random $g \in G$.

There are "derived" cryptographic schemes for which the hardness of the *Diffie-Hellman decision problem* **DHDP** determines security. The DHDP asks—for randomly given elements g_0, g_1, g_2, g_3—for a decision whether

$$\log_{g_0}(g_1) \cdot \log_{g_0}(g_1) = \log_{g_0}(g_3).$$

DL-Systems

To use (family of) groups G for public key systems we have to solve three crucial tasks:

(1) Store the elements in G in a computer in a compact way (ideally $\mathcal{O}(\log(\mid G \mid))$ bits should be enough).
(2) The group composition is given by an algorithm which is easily implemented and very fast (at most polynomially bounded time and space is allowed). So exponentiation is of polynomial complexity, too.
(3) The computation of the DL in G (for random elements) is (to the best of our knowledge) very hard and so unfeasible in practice (ideally exponential in $\mid G \mid$).

Groups G with generator g_0 satisfying these conditions are called DL-systems.

Remark 3.1. We use the structure "group" to define the crypto primitive. This already implies that the complexity of the DHCP is at most $\sim \sqrt{\ell}$ since there are (deterministic and probabilistic) algorithms for the computation of discrete logarithms applicable to *all* groups (e.g., Shank's Baby-Step-Giant-Step or Pollard's ρ-algorithm) of this complexity. A deep fact is that in generic groups no faster algorithm is available. Hence $\sqrt{\ell}$ is the benchmark for the hardness of DL and DHCP.

Bilinear Structures

Discrete logarithms concern the $\mathbb{Z}/\ell$-linear structure of cyclic groups. In every elementary course on linear algebra one learns that there are multilinear aspects coming in a natural way from the theory of linear maps. The principle behind this is duality.

During the last ten years this aspect has become more and more important in public key cryptography, and there is much ongoing research in this area.

Definition 3.2. Let G be a cyclic group of prime order ℓ. Assume that there are $\mathbb{Z}/\ell$-modules B and C and a bilinear map $Q : G \times B \to C$ with

- **i):** the group composition laws in G, B, and C, as well as the map Q, are fast (e.g., polynomial time).
- **ii):** For random $b \in B$ we have $Q(g_1, b) = Q(g_2, b)$ iff $g_1 = g_2$.

We call (G, Q) a *DL-system with bilinear structure.*

Since we can transfer the computation of discrete logarithms from G to C via Q, the existence of bilinear structures may weaken DL-systems. Moreover, if $G = B$, then DHDP becomes trivial.

But there are very interesting constructive features, too. In the center of interest are short signatures and identity-based protocols.

Candidates for DL-Systems

We want to find groups G satisfying the conditions of subsection "DL-Systems" and analyze bilinear structures.

As mentioned earlier, Diffie and Hellman suggested taking a prime ℓ dividing $q - 1$ and G as the group of ℓth roots of unity in $\mathbb{F}_q^*$. It is not difficult to find instances where ℓ is of the same magnitude as q, and conditions i and ii of Definition 3.2 are satisfied. But the hardness of the DL (which is the classical one already studied by Jacobi) is disappointing. It is only of subexponential complexity; the reason is that points on the projective line over $\mathbb{F}_q$ are easily lifted to points on the line over $\mathbb{Z}$, and then a powerful method, index calculus (cf. [ACF]), can be applied.

Structural theorems about abelian varieties over number fields imply that such a lifting is very difficult if we take Jacobians of curves of genus larger than 0.

This was the motivation for V. Miller [M1] to suggest in 1985 using elliptic curves over finite fields for DL-systems. Independently N. Koblitz [K1] suggested this at the same time, and in 1989 [K2] he went further to propose class groups of hyperelliptic curves, too.

With the results obtained in Theorem 2.1, we can go even further and try to use divisor class groups of curves of any genus $g > 0$ over $\mathbb{F}_q$. Fixing a size of magnitude M for ℓ (e.g., 10^{80}) we can use Weil's theorem and take q and g such that $M \approx q^g$ (e.g., $g \cdot \log_{10}(q) \approx 80$). Then the diophantine task is to find a field $\mathbb{F}_q$ and a curve C of genus g defined over $\mathbb{F}_q$ such that $|\mathrm{Pic}^0(C)|$ is (maybe up to a small cofactor) a prime number ℓ.

There are theorems from analytic number theory which predict that the chances to find such pairs are rather good; and so the strategy is to choose random curves and test whether they satisfy the condition.

To do this we need a fast algorithm to determine $|\mathrm{Pic}^0(C)|$.

Point Counting

This cannot be done naively. One has to use the action of Frobenius automorphisms ϕ_q and compute its characteristic polynomial. For this one uses all the techniques sketched in the section "Some Aspects of Arithmetic Geometry" in very advanced forms:

(1) Class field theory of CM-fields leads to the so-called CM method worked out for genus 1,2,3 by Atkin, Enge, Morain, Spallek, Weng, and many others. Hence we use global Galois theory. We remark that for $g = 1$ this leads to the explicit class field theory of imaginary quadratic fields.
(2) Etale cohomology groups, i.e., ℓ-adic representations attached to Tate modules of J_C (Schoof, Atkin, Elkies, Pila, ...).
(3) p-adic cohomology (Satoh, Kedlaya, Mestre,...) for moderate sizes of p (lifting to p-adic fields).

For details we refer to [ACF]. As state of the art, we get

Result 1. In cryptographically relevant areas

- we can count points on random elliptic curves,
- we can count points on Jacobians of random curves over fields of small (and even medium) characteristic,
- we still have problems with random curves of genus 2 over prime fields (but see the work of Gaudry and Schost [GS]); we can

use class field theory of CM-fields to find an abundance of curves of genus 2 suitable for DL-systems,
- and, of course, we have many *special* families of curves whose members are accessible for point counting.

So for every $g \in \mathbb{N}$ we find divisor class groups of curves of genus g which satisfy conditions 1 and 2 of the subsection "DL-Systems".

Attacks

Curves of Genus > 2. The main motivation to take divisor class groups was to avoid index-calculus attacks enabled by lifting points on curves from finite fields to number fields.

But it was soon discovered that the internal structure of divisor class groups of curves of large genus yields another type of index-calculus attack.

Even worse: variants of this attack are applicable to curves of moderate genus. The sharpest result nowadays is

Theorem 3.3 (Diem, Gaudry, Thomé, Thériault)**.** *There exists an algorithm which computes, up to* $\log(q)$*-factors, the DL in the divisor class group of curves of genus* g *in expected time of* $\mathcal{O}(q^{(2-2/g)})$.

If C *is given by a plane curve of degree* d *(singularities allowed), then the DL in the group of divisor classes of degree* 0 *is of complexity* $\mathcal{O}(q^{2-\frac{2}{d-2}})$.

Recall that the generic algorithms have complexity $\mathcal{O}(q^{g/2})$. Hence curves of genus larger than 4 are not advisable.

The results of the theorem do not imply that systems using hyperelliptic curves of genus 4 are insecure, but the parameters of the systems have to be larger than for generic groups.

Particularly interesting is the situation for curves of genus 3. Surprisingly, nonhyperelliptic curves are less secure since they have a plane model of degree 4. So one has to exclude hyperelliptic curves of genus 3 which have computable isogenies to nonhyperelliptic curves. Unfortunately there are many such curves, and at the moment it is not clear how to find criteria for the nonexistence of such isogenies.

So it may be wise to use only curves of genus 1 and 2 if there are no very good reasons for deciding otherwise.

Elliptic Curves. There is no direct index-calculus attack known which is effective on elliptic curves. But if the ground field is not a prime field, we can apply Weil descent (a well-known method in algebraic geometry [Fr1]) and transfer the DL in $E(\mathbb{F}_q)$ to the DL in an abelian variety of higher dimension and defined over a smaller ground field. Again, one can try to apply index-calculus in these abelian varieties, and there are many cases where one succeeds (e.g., the field $\mathbb{F}_{2^{155}}$ is not good for cryptographical use). To avoid this it is suggested that one use as ground field either $\mathbb{F}_p$ or $\mathbb{F}_{2^n}$, where n is a prime which is not a Mersenne prime.

Bilinear Structures

In the last section we saw that divisor classes of carefully enough chosen curves of genus 1 (elliptic curves) and genus 2 cannot be attacked nowadays by index-calculus methods. Now we want to discuss transfers by duality theorems coming from class field theory of local and global fields [Fr3].

The Lichtenbaum-Tate Pairing on Elliptic Curves

Divisor class groups carry a natural duality induced by class field theory of local fields. For Jacobian varieties this is made explicit by the Lichtenbaum-Tate pairing.

We state a version of this pairing for elliptic curves E over $\mathbb{F}_q$. (For the general background see [Fr3], for the general version see [FR].)

Theorem 4.1. *Let* ℓ *be a prime dividing* $|E(\mathbb{F}_q)|$. *Let* k *be the smallest natural number such that* $\ell|(q^k-1)$. *Let* ζ_ℓ *be an* ℓ*th root of unity in* $\mathbb{F}_{q^k}$.

Define $G := E[\ell] \cap E(\mathbb{F}_q)$ *and* $G^\perp := \{Q \in E(\mathbb{F}_q^k) \cap E[\ell];\ \phi_q(Q) = q \cdot Q\}$.

There is a nondegenerate pairing

$$Q : G \times G^\perp \to <\zeta_\ell>$$

which can be computed with complexity polynomial in $k \cdot \log q$.

This pairing is given very explicitly. For an algorithm see [M2] and [ACF], Chapter 16. Because of its importance there are many versions trying to accelerate the computation of pairings related to Q.

Corollary 4.2. *The DL in elliptic curves over* $\mathbb{F}_q$ *is transferred to the discrete logarithm in* $\mathbb{F}_{q^k}^*$ *and hence has complexity which is subexponential in* $k \cdot \log q$.

Pairing-Friendly Curves

Theorem 4.1 has computational relevance only if k is not large.

It is a dangerous result if k is small (say ≤ 6), and it can be used constructively if $k \approx 12$.

If we take random elliptic curves we shall expect that k is of the same size as ℓ. An elliptic curve E is called pairing-friendly if k is small.

There are curves which are too friendly: if E is supersingular (i.e., there are no algebraic points of order p on E), then $k \leq 6$ and even $k \leq 2$ if $p > 3$. Hence the DL on supersingular curves is not harder than the classical DL. For this reason supersingular curves play only a minor role in DL-systems based on elliptic curves. The nice thing is that they deliver easy examples for groups with gaps between DHCP and DHDP.

To use the constructive aspects of pairings described in the subsection "Bilinear Structures", for example, short signatures, one has to solve interesting diophantine problems in order to find elliptic curves with small, but not too small, k. In [BN] one finds (conjecturally infinitely) many elliptic curves with $k = 12$.

Remark 4.3. It would be interesting to be able to construct nonsupersingular pairing-friendly curves of genus 2.

Conclusion

We end by giving the equation of an elliptic curve which passed all security checks and travels on passports. Its equation is

$$Y^2Z = X^3 + AXZ^2 + BZ^3$$

with

$A = 7D5A0975FC2C3057EEF67530417AFFE7FB8055C126DC5C6CE94A4B44F330B5D9,$

$B = 26DC5C6CE94A4B44F330B5D9BBD77CBF958416295CF7E1CE6BCCDC18FF8C07B6$

defined over $\mathbb{F}_{A9FB57DBA1EEA9BC3E660A909D838D726E3BF623D52620282013481D1F6E5377}$.

$|E(\mathbb{F}_p)|$ is a prime ℓ

with $\ell = A9FB57DBA1EEA9BC3E660A909D838D718C397AA3B561A6F7901E0E82974856A7$.

The name of the curve is **brainpoolP256r1**.

References

[ACF] H. COHEN and G. FREY (eds.), *Handbook of Elliptic and Hyperelliptic Curve Cryptography*, CRC, 2005.

[BN] P. S. L. M. BARRETO and M. NÄHRIG, *Pairing-Friendly Elliptic Curves of Prime Order*, SAC'2005, LNCS 3897, Springer, 319–331, 2006.

[BRAIN] http://www.ecc-brainpool.org/ecc-standard.htm.

[C] J. CREMONA, *Algorithms for Modular Elliptic Curves*, 2nd edition, Cambridge University Press, 1997.

[DH] W. DIFFIE and M. E. HELLMAN, New Directions in Cryptography, *IEEE Transactions on Information Theory*, **22**(6) (1976), 644–654.

[Di] C. DIEM, *On arithmetic and the discrete logarithm problem in class groups of curves*, Habil. Thesis, Leipzig 2009.

[Fr1] G. FREY, *How to disguise an elliptic curve*, slides, http://cacr.math.uwaterloo.ca/conferences/1998/ecc98/slides.html.

[Fr3] ———, Discrete Logarithms, Duality, and Arithmetic in Brauer groups, in: *Algebraic Geometry and its Applications*, J. Chaumine, J. Hirschfeld, R. Rolland eds., 241–272, World Scientific, 2008.

[FR] G. FREY and H. G. RÜCK, A remark concerning m-divisibility and the discrete logarithm problem in the divisor class group of curves, preprint IEM, Essen, 7, 1991, appeared in: *Math. Comp.* **62** (1994), 865–874.

[GS] P. GAUDRY and E. SCHOST, Construction of secure random curves of genus 2 over prime fields, in: *Advances in Cryptology*, Eurocrypt 2004, LNCS 3027, 239–256, 2004.

[He] F. HESS, Computing Riemann-Roch spaces in algebraic function fields and related topics, *J. Symbolic Comp.* **33**(4) (2002), 425–445.

[K1] N. KOBLITZ, Elliptic curve cryptosystems, *Math. Comp.* **48** (1987), 203–209.

[K2] ———, Hyperelliptic cryptosystems, *J. Cryptology* **1** (1989), 139–150.

[M1] V. MILLER, *Use of elliptic curves in cryptography*, Advances in cryptology—CRYPTO 85, Springer Lecture Notes in Computer Science, vol. 218, 1985.

[M2] V. S. MILLER, Short programs for functions on curves, IBM, Thomas J. Watson Research Center, 1986; http://crypto.sanford.edu/miller/.

[NIST] National Institute of Standard and Technology, Recommended elliptic curves for federal use, 1999, http://www.csrc.nist.gov/encryption/.

[W] A. WILES, Modular elliptic curves and Fermat's last theorem, *Ann. Math.* **141**(3) (1995), 443–551.