Project-Team TANC

Théorie Algorithmique des Nombres pour la Cryptologie

Saclay - Île-de-France

Theme : Algorithms, Certification, and Cryptography
# Table of contents

1. **Team** .............................................................. 1  
2. **Overall Objectives** ............................................... 1  
   2.1. Main topics  
   2.2. Exploratory topics  
3. **Scientific Foundations** ........................................... 2  
   3.1. General overview  
   3.2. Algebraic curves over finite fields  
      3.2.1. Effective group laws  
      3.2.2. Cardinality  
      3.2.3. Computing isogenies  
      3.2.4. The discrete logarithm problem  
      3.2.5. Pairings on algebraic curves  
   3.3. Complex multiplication  
      3.3.1. Genus 1  
      3.3.2. Genus 2  
   3.4. Algebraic Geometry codes  
4. **Application Domains** ............................................ 7  
5. **Software** .......................................................... 7  
   5.1. ECPP  
   5.2. mpc  
   5.3. mpfrcx  
   5.4. TIFA  
   5.5. FAAST  
6. **New Results** ..................................................... 9  
   6.1. Algebraic curves over finite fields  
      6.1.1. Cardinality  
      6.1.2. Isogenies  
      6.1.3. Discrete logarithms on curves  
      6.1.4. Pairings  
   6.2. Complex multiplication  
   6.3. Algebraic codes  
      6.3.1. List decoding of Reed–Solomon codes  
      6.3.2. List decoding of Algebraic Geometry codes  
      6.3.3. List decoding of binary codes  
      6.3.4. Homomorphic encryption  
   6.4. Number fields  
7. **Contracts and Grants with Industry** ............................. 12  
   7.1. Gemplus  
   7.2. Industrial ANR  
8. **Other Grants and Activities** .................................... 13  
   8.1. Network of excellence  
   8.2. ANR  
9. **Dissemination** .................................................... 13  
   9.1. Programme committees  
   9.2. Teaching  
   9.3. Seminars and talks  
   9.4. Vulgarization and Summer schools  
   9.5. Editorship  
   9.6. Awards
9.7. Thesis committees ................................. 14
9.8. Research administration ................. 14
10. Bibliography ......................................................... 14
1. Team

Research Scientist
Andreas Enge [ CR1, HdR ]
Daniel Augot [ DR2, HdR ]
Benjamin Smith [ CR2 ]

Faculty Member
François Morain [ Team leader, Professor at École polytechnique, HdR ]

Technical Staff
Jérôme Milan [ Ingénieur de Développement Digiteo ]

PhD Student
Luca De Feo [ École polytechnique since 2007-09-01 ]
Jean-François Biasse [ DGA since 2007-09-01 ]
Morgan Barbier [ École polytechnique since 2008-10-01 ]
Guillaume Quintin [ DGA since 2009-10-01 ]

Post-Doctoral Fellow
Alain Couvreur [ Since 2009-10-01 ]

Administrative Assistant
Évelyne Rayssac [ École polytechnique ]

2. Overall Objectives

2.1. Main topics

TANC is located in the Laboratoire d'Informatique de l'École polytechnique (LIX). The project was created on 2003-03-10.

The aim of the TANC project is to promote the study, implementation and use of robust and verifiable asymmetric cryptosystems based on algorithmic number theory.

It is clear from this statement that we combine high-level mathematics and efficient programming. Our main area of competence and interest is that of algebraic curves over finite fields, and most notably the computational aspects of these objects, which appear as a substitute for modular arithmetic in new analogues of old-fashioned cryptography. One reason for this change is that we can achieve an equivalent security level with a much smaller key size. Our research contributes to the effort to find a diverse range of secure substitutes for the famous RSA (Rivest–Shamir–Adleman) cryptosystem, in case some attack appears and destroys the products that use it.

Whenever possible, we produce certificates (proofs) of validity for the objects and systems we build. For instance, an elliptic curve has many invariants, and their values need to be proved, since they may be difficult to (re-)compute.

Our research area includes:

- Fundamental number theoretic algorithms: We are interested in primality proving algorithms based on elliptic curves, integer factorization, and the computation of discrete logarithms over finite fields. These problems lie at the heart of the security of arithmetic based cryptosystems.
- Algebraic curves over finite fields: We tackle algorithmic problems involving efficiently computing group laws on Jacobians of curves, evaluating the cardinality of these objects, and studying the security of the discrete logarithm problem in such groups. These topics are crucial to the applicability of these objects in real crypto products.
• **Complex multiplication:** The theory of Complex Multiplication is a meeting point of algebra, complex analysis and algebraic geometry. Its applications range from primality proving to the efficient construction of elliptic and hyperelliptic curve-based cryptosystems.

• **Pairings:** The new number theoretic primitive of pairings (i.e. bilinear functions) on algebraic curves enables many novel applications, and poses algorithmic challenges concerning efficient implementation and the creation of secure instances.

• **Decoding algorithms for Algebraic Geometric codes:** The algorithmic knowledge of TANC will be used to accelerate decoding algorithms, be they the classical one (up to half to the minimum distance), or new ones which decode many more errors.

### 2.2. Exploratory topics

As described in the name of our project, we aim to provide robust primitives for asymmetric cryptography. In recent years, we have made several attempts at applying our knowledge to real life protocols. We also aim to promote the use of curve-based cryptography in new environments, such as ad hoc networks. We will also try to promote the use of AG codes, which are the coding-theoretic analogue of elliptic curves in cryptology.

### 3. Scientific Foundations

#### 3.1. General overview

Once considered beautiful but useless, arithmetic has proven a spectacular success in the creation of a new paradigm in cryptography. Classical cryptography was mainly concerned with symmetric techniques: two principals wishing to communicate secretly had to share a common secret beforehand, and this same secret was used both for encrypting the message and for decrypting it. This mode of communication is efficient enough when traffic is low, or when the principals can meet prior to communication.

However, modern networks are simply too large for the classical paradigm to remain efficient any longer. Hence the need for cryptography without first contact. In theory, this is easy. Find two algorithms $E$ and $D$ that are reciprocal (i.e., $D(E(m)) = m$) and such that the knowledge of $E$ does not help in computing $D$. Then $E$ is dubbed a public key available to anyone, and $D$ is the secret key, reserved to a user. When Alice wants to send an email message $m$ to Bob, she uses his public key $E$ to send him the encrypted message $E(m)$, which he can decrypt with the secret key $D$: we have thus achieved secret communication without a common secret key. (Of course, everything has to be presented in the modern language of complexity theory: $E$ and $D$ must be computable in polynomial time, while finding $D$ from $E$ alone without some secret knowledge should be possible only in, say, exponential time.) Though simplified and somewhat idealized, this is the heart of asymmetric cryptology. Modern cryptography provides not only secure communication channels but also solutions to the signature problem, as well as some solutions for identifying all parties in protocols, thus enabling products to be usable on the Internet (such as ssh and ssl/tls).

Now, where do difficult problems come from? Mostly from arithmetic, where we find problems such as the integer factorization problem and the discrete logarithm problem. It appears to be important to vary the groups which act as settings for concrete instances of the abstract difficult problems, since this provides some biodiversity which is key to resisting crypto-analytic attacks. The groups proposed include finite fields, modular integers, algebraic curves, and class groups. All of these now form cryptographic primitives that need to be assembled in protocols, and finally in commercial products.

Our activity is concerned with the beginning of this process: we are interested in difficult problems arising in computational number theory and the efficient construction of these primitives. TANC concentrates on modular arithmetic, finite fields and algebraic curves.
We have a strong, well-known reputation for breaking records, whatever the subject is: constructing systems or breaking them. We have world-record computations in areas including primality proving, class polynomials, modular equations, computing cardinalities of algebraic curves, and discrete logarithms. This means writing programs and putting in all the work needed to make them run for weeks or months. An important part of our task is now to transform record-breaking programs into programs to solve everyday cryptographic problems for current parameter sizes.

Certificates are another of our major concerns. By certificates, we mean efficiently verifiable proofs of the properties of the objects we build. While these certificates might be difficult to build, they are easy to check (by customers, for example). The traditional example is that certificates for primality of prime numbers, which were introduced by Pratt in 1974. We know how to construct certificates for the important properties of elliptic curves, with the aim of establishing what we call an identity card for a curve (including its cardinality together with the proof of its factorization, its group structure with proven generators, and its discriminant with its proven factorization, and the class number of the associated order). The theory is ready for this, and the algorithms are not out of reach. This approach must be extended to other curves; the theory is almost ready in several cases, but algorithms are still to be found. This is one of the main problems facing TANC.

The mathematics used in cryptology is becoming more and more complex (for example, consider the recent algorithms using $p$-adic cohomology). The new, more mathematically complex algorithms cannot live if we do not implement them. For implementations, we need more and more evolved algorithmic primitives; currently, these may be available in very rare mathematical systems such as MAGMA. Once our algorithms work in MAGMA, it is customary to rewrite them in C or C++ to gain speed. Along the same lines, some of our C programs developed for our research (an old version of ECPP, some parts of discrete log computations, cardinality of curves) are now included in the MAGMA system, as a result of our collaboration with the Sydney group.

3.2. Algebraic curves over finite fields

One of the most common cryptographic protocols is Diffie–Hellman Key Exchange, which enables Alice and Bob to exchange secret information over an insecure channel. Given a publicly known cyclic group $G$ of generator $g$, Alice sends $g^a$ for a random $a$ to Bob, and Bob responds with a random $g^b$. Both Alice and Bob can now compute $g^{ab}$, and this is henceforth their common secret. Of course, this a schematic presentation; real-life protocols based on this need more security properties. Being unable to recover $a$ from $g^a$ (the discrete log problem – DLP) is a fundamental to the security of the scheme, and groups for which the DLP is difficult must be favored. Therefore, the choice of group $G$ is crucial, and TANC concentrates on groups derived from algebraic curves. These groups offer a very interesting alternative to finite fields: the DLP in a finite field can be broken by subexponential algorithms, while exponential time is required for an elliptic curve over the same field. Smaller keys can therefore be used in curve-based cryptosystems; this is very interesting from the point of view of limited-power devices.

In order to build a cryptosystem on an algebraic curve over a finite field, one needs to efficiently compute the group law (and hence have a nice representation for elements of the Jacobian of the curve). Next, one must compute the cardinality of the Jacobian, so that we can find generators of the group. Once the curve is built, one needs to test its security, for example by determining the hardness of the DLP in its Jacobian.

3.2.1. Effective group laws

The curves that interest us are typically defined over a finite field $\mathbb{GF}(p^n)$, where $p$ is the (prime) characteristic of the field. The points of an elliptic curve $E$ (of equation $y^2 = x^3 + ax + b$, say) form an abelian group, that was thoroughly studied during the preceding millennium. Adding two points is usually done using the so-called chord-and-tangent formula. When dealing with a genus $g$ curve (the elliptic case being $g = 1$), the associated group is the Jacobian (set of $g$-tuples of points modulo an equivalence relation), an object of dimension $g$. Points are replaced by polynomial ideals. This requires the help of tools from effective commutative algebra, such as Gröbner bases or Hermite normal forms.
The great catalog of usable curves is now complete, as a result of the work of TANC, notably in two ACI (CRYPTOCOURBES and CRYPTOLOGIE P-ADIQUE) that are now finished.

### 3.2.2. Cardinality

Once the group law is tractable, one has to find means of computing the cardinality of the group, which is not an easy task in general. Of course, this has to be done as fast as possible, if changing the group very frequently in applications is imperative.

Two parameters enter the scene: the genus $g$ of the curve, and the characteristic $p$ of the underlying finite field. When $g = 1$ and $p$ is large, the only current known algorithm for computing the number of points of $E/\text{GF}(p)$ is that of Schoof–Elkies–Atkin. Thanks to the works of the project, world-wide implementations are able to build cryptographically strong curves in less than one minute on a standard PC. Recent improvements were made by F. Morain and P. Gaudry (CACA0), see [42]. The current record of SEA was established by F. Morain in 2007 for a prime $p$ of 2500 decimal digits (again compared to 500dd back in 1995), using the work in [3] and in [9], in which a new approach to the eigenvalue computation is described and proven.

When $p$ is small (one of the most interesting cases for hardware implementation in smart cards being $p = 2$) the best current methods use $p$-adic numbers, following the breakthrough of T. Satoh with a method working for $p \geq 5$. The first version of this algorithm for $p = 2$ was proposed independently by M. Fouquet, P. Gaudry and R. Harley and by B. Skjernaa. J.-F. Mestre has designed the currently fastest algorithm using the arithmetico-geometric mean (AGM) approach. Developed by R. Harley and P. Gaudry, it led to new world records. Then, P. Gaudry combined this method with other approaches to make it competitive for cryptographic sizes [41].

When $g > 1$ and $p$ is large, polynomial time algorithms exist, but their implementation is not an easy task. P. Gaudry and É. Schost have modified the best existing algorithm so as to make it more efficient. They were able to build the first random cryptographically strong genus 2 curves defined over a large prime field [43]. To get one step further, one needs to use genus 2 analogues of modular equations. After a theoretical study [44], they are now investigating the practical use of these equations.

When $p = 2$, $p$-adic algorithms led to striking new results. First, the AGM approach extends to the case $g = 2$ and is competitive in practice (only three times slower than in the case $g = 1$). In another direction, Kedlaya has introduced a new approach, based on the Monsky–Washnitzer cohomology. His algorithm works originally when $p > 2$. P. Gaudry and N. Gürel implemented this algorithm and extended it to superelliptic curves, which had the effect of adding these curves to the list of those usable in cryptography.

Closing the gap between small and large characteristic leads to pushing the $p$-adic methods as far as possible. In this spirit, P. Gaudry and N. Gürel have adapted Kedlaya’s algorithm and exhibited a linear complexity in $p$, making it possible to reach a characteristic of around 1000 (see [39]). For larger $p$’s, one can use the Cartier–Manin operator. Recently, A. Bostan, P. Gaudry and É. Schost have found a much faster algorithm than currently known ones [26]. Primes $p$ around $10^9$ are now doable.

### 3.2.3. Computing isogenies

The core of the Schoof–Elkies–Atkin (SEA) algorithm that computes the cardinality of elliptic curves over finite fields consists in using the theory of isogenies to find small factors of division polynomials. SEA is still the method of choice for the large characteristic case, but no longer for small characteristics.

Isogenies are also a tool for understanding the difficulty of the Discrete Log problem among classes of elliptic curves [52]. Recently, there appeared suggestions to use isogenies in a cryptographic context, replacing the multiplication on curves by the use of such morphisms [63], [60].

Algorithms for computing isogenies are very well known and used in the large characteristic case. When the characteristic is small, three algorithms exist: two due to Couveignes [29], [30], [56], and one due to Lercier [55].
3.2.4. The discrete logarithm problem

The discrete logarithm problem is one of the major difficult problems that allow us to build secure cryptosystems. It has essentially been proved equivalent to the computational Diffie–Hellman problem, which is closer to the actual security of many systems. For an arbitrary group of prime order $N$, it can be solved by a generic, exponential algorithm using $\Theta(\sqrt{N})$ group operations. For elliptic curves, set aside some rare and easily avoidable instances, no faster algorithms are known.

In higher genus curves, the algorithms with the best complexity create relations as smooth principal divisors on the curve and use linear algebra to deduce discrete logarithms, similarly to the quadratic sieve for factoring. The first such algorithm for high genus hyperelliptic curves with a heuristic complexity analysis is given in [23], and A. Enge has developed the first algorithm with a proven subexponential run time of $L(1/2)$ in [35]. Generalisations to further groups suggested for cryptography, in particular ideal class groups of imaginary quadratic number fields, are obtained by A. Enge and P. Gaudry in [5] [34]. Proofs for arbitrary curves of large genus are given by J.-M. Couveignes [28] and F. Heß [49].

The existence of subexponential algorithms shows that high genus curves are less secure than, say, elliptic ones in cryptography. By analyzing the same algorithms differently, concrete recommendations for key lengths can be obtained, an approach introduced by P. Gaudry in [40] and pursued in [45]. It turns out that elliptic curves and hyperelliptic curves of genus 2 are not affected, while the key lengths have to be increased in higher genus, for instance by 12 % in genus 3.

Using similar algorithms to those analyzed in [5], C. Diem has shown in [31] that non-hyperelliptic curves (of genus at least 3) are even less secure than hyperelliptic ones of the same genus. This effectively leaves elliptic and low genus hyperelliptic curves as potential sources for public-key cryptosystems.

3.2.5. Pairings on algebraic curves

Algebraic curves have first been used in cryptography as a source for groups in which the discrete logarithm problem should be harder than in the multiplicative group of a finite field. Totally new applications stem from the use of structures proper to algebraic curves, the Tate and Weil pairings. These are bilinear maps that associate to two group elements, at least one of which is defined in an extension field, a root of unity in the same extension field. Among the first new cryptographic primitives were a tripartite Diffie–Hellman key exchange [53] and identity based encryption [61]. Subsequently, the number of articles concerned with pairings has exploded, and a specialised series of conferences has been inaugurated with Pairings 2007 in Tokyo, A. Enge being a member of the programme committees in 2007 and 2008.

One of the most challenging problems related to pairing based cryptography is to find suitable curves, that are hidden like needles in a hay stack. Supersingular elliptic curves yield a rather limited supply of doubtful security. Using its expertise on complex multiplication, the TANC team has published one of the first two algorithms for finding pairing friendly ordinary curves for arbitrary field extension degrees in [33], the other one being developed in [24].

3.3. Complex multiplication

3.3.1. Genus 1

Despite the achievements described above, random curves are sometimes difficult to use, since their cardinality is not easy to compute or useful instances are too rare to occur (curves for pairings for instance). In some cases, curves with special properties can be used. For instance curves with complex multiplication (in brief CM), whose cardinalities are easy to compute. For example, the elliptic curve defined over $GF(p)$ of equation $y^2 = x^3 + x$ has cardinality $p + 1 - 2u$, when $p = u^2 + v^2$, and computing $u$ is easy.

The CM theory for genus 1 is well known and dates back to the middle of the nineteenth century (Kronecker, Weber, etc.). Its algorithmic part is also well understood, and recently more work was done, largely by TANC. Twenty years ago, this theory was applied by Atkin to the primality proving of arbitrary integers, yielding the ECPP algorithm developed ever since by F. Morain. Though the decision problem $\text{ISPRIME}$ was shown to be
in $P$ (by the 2002 work of Agrawal, Kayal, Saxena), practical primality proving of large random numbers is still done only with ECPP.

These CM curves enabled A. Enge, R. Dupont and F. Morain to give an algorithm for building good curves that can be used in identity based cryptosystems [33].

CM curves are defined by algebraic integers, whose minimal polynomials have to be computed exactly, the coefficients being exact integers. The fastest algorithm to perform these computations requires a floating point evaluation of the roots of the polynomial to a high precision. F. Morain on the one hand and A. Enge (together with R. Schertz) on the other, have developed the use of new class invariants that characterize CM curves. The union of these two families is currently the best that can be achieved in the field (see [7]). Later, F. Morain and A. Enge have designed a fast method for the computation of the roots of this polynomial over a finite field using Galois theory [36]. These invariants, together with this new algorithm, are incorporated in the working version of the program ECPP.

F. Morain analyzed a fast variant of ECPP, called fastECPP, which led him to gain one order of magnitude in the complexity of the problem (see [12] [58]), reaching heuristically $O((\log N)^{4+\epsilon})$, compared to $O((\log N)^{5+\epsilon})$ for the basic version. By comparison, the best proven version of AKS [54] has complexity $O((\log N)^{6+\epsilon})$ and has not been implemented so far; the best randomized version [25] reaches the same $O((\log N)^{4+\epsilon})$ bound but suffers from memory problems and is not competitive yet. F. Morain implemented fastECPP and was able to prove the primality of 10,000 decimal digit numbers [12], as opposed to 5,000 for the basic (historical) version. Continuously improving this algorithm, this led to new records in primality proving, some of which obtained with his co-authors J. Franke, T. Kleinjung and T. Wirth [38] who developed their own programs.

F. Morain set the current world record to 20,562 decimal digits early June 2006, as opposed to 15,071 two years before. This record was made possible by using an updated MPI-based implementation of the algorithm and its distribution process on a cluster of 64-bit bi-processors (AMD Opteron(tm) Processor 250 at 2.39 GHz). In 2007, another large number was proven to be prime, namely $(2^{42737} + 1)/3$ with 12,865 decimal digits.

In his thesis, R. Dupont has investigated the complexity of the evaluation of some modular functions and forms (such as the elliptic modular function $j$ or the Dedekind eta function for example). High precision evaluation of such functions is at the core of algorithms to compute class polynomials (used in complex multiplication) or modular polynomials (used in the SEA elliptic curve point counting algorithm).

Exploiting the deep connection between the arithmetic-geometric mean (AGM) and a special kind of modular forms known as theta constants, he devised an algorithm based on Newton iterations and the AGM that has quasi-optimal linear complexity. In order to certify the correctness of the result to a specified precision, a fine analysis of the algorithm and its complexity was necessary [14].

Using similar techniques, he has given a proven algorithm for the evaluation of the logarithm of complex numbers with quasi-optimal time complexity.

### 3.3.2. Genus 2

The theory of Complex Multiplication also exists for non-elliptic curves, but is more intricate, and only recently can we dream to use them. Some of the recent results occurred as the work of R. Dupont (former member of TANC) in his thesis.

R. Dupont has worked on adapting his algorithm to genus 2, which induces great theoretical and technical difficulties. He has studied a generalization of the AGM known as Borchardt sequences, has proven the convergence of these sequences in a general setting, and has determined the set of limits such sequences have in genus 2. He has then developed an algorithm for the fast evaluation of theta constants in genus 2, and as a byproduct obtains an algorithm to compute the Riemann matrix of a given hyperelliptic curve: given the equation of such a curve, it computes a lattice $L$ such that the Jacobian of the curve is isomorphic to $\mathbb{C}/L$. These algorithms are both quasi-linear, and have been implemented (in C, using the multiprecision package GMP – see [http://gmplib.org/](http://gmplib.org/)).
Using these implementations, R. Dupont has began computing modular polynomials for groups of the form \( \Gamma_0(p) \) in genus 2 (these polynomials link the genus 2 \( j \)-invariants of \( p \)-isogenous curves). He computed the modular polynomials for \( p = 2 \), which had never been done before, and did some partial computations for \( p = 3 \) (results are available at [http://www.lix.polytechnique.fr/Labo/Regis.Dupont](http://www.lix.polytechnique.fr/Labo/Regis.Dupont)). He also studied more theoretically the main ingredient used in his algorithms in genus 2, a procedure known as Borchardt sequences. In particular, he proved a theorem that parametrizes the set of all possible limits of Borchardt sequences starting with a fixed 4-tuple.

### 3.4. Algebraic Geometry codes

There are many other applications of algorithmic methods on algebraic curves besides simple cryptography. Daniel Augot plans to develop a new activity around algebraic geometry (AG) codes, a very powerful family of codes that often beat records for their parameters: they often offer the best correction capacity. The main topic of research is to accelerate the decoding algorithms of these codes, which have a slightly expensive cost [50]. A reference implementation would be of major interest, to help people compare AG codes with Reed–Solomon codes.

Guruswami and Suday have obtained a breakthrough [48] for decoding AG codes with many errors. Still, there is no implementation available yet, even for the most simple AG codes (which are the Hermitian codes). In this domain too, an objective is to produce a publicly available reference implementation.

### 4. Application Domains

#### 4.1. Telecommunications

Clearly, our main field of applications is telecommunications. We participate in the protection of information. We are proficient on a theoretical level, and ready to develop applications using modern cryptologic techniques, with a main focus on elliptic curve cryptography and codes based on algebraic curves. One potential application is cryptosystems in environments with limited resources as smart cards, mobile phones, and ad hoc networks. For coding, we envisage developing algebraic codes for the erasure channel or distributed storage.

### 5. Software

#### 5.1. ECPP

F. Morain has been continuously improving his primality proving algorithm called ECPP, originally developed in the early ’90s. Binaries for version 6.4.5 have been available since 2001 on his web page. Proving the primality of a 512 bit number requires less than a second on an average PC. His personal record is about 20,000 decimal digits, with the fast version he started developing in 2003. All of the code is written in C, and based on the GMP package.

#### 5.2. mpc

The MPC library, developed in C by A. Enge in collaboration with Ph. Théveny and P. Zimmermann, implements the basic operations on complex numbers in arbitrary precision, which can be tuned to the bit. This library is based on the multiprecision libraries GMP and MPFR. Each operation has a precise semantics, in such a way that the results do not depend on the underlying architecture. Several rounding modes are available. This software, licensed under the GNU Lesser General Public License (LGPL), can be downloaded freely from the URL [http://www.multiprecision.org/mpc/](http://www.multiprecision.org/mpc/).
The library currently benefits from an INRIA Opération de développement logiciel. The latest version (0.8) was released in November 2009. A Debian package has been available (in the unstable distribution) since October 2008. The perl wrapper Math::MPC (http://search.cpan.org/~sisyphus/Math-MPC/) has been available on CPAN since version 0.4.6.

The mpc library is used in our team to build curves with complex multiplication and to compute modular polynomials (cf. Section 6.1), and it is de facto incorporated in the ECPP program. It is used by the Magma Computational Algebra System (http://magma.maths.usyd.edu.au/magma/) and by Trip (http://www.imcce.fr/Equipes/ASD/trip/trip.php), a symbolic-numeric system for celestial mechanics developed at Institut de Mécanique Céleste et de Calcul des Éphémérides.

5.3. mpfrcx

The mpfrcx library, developed in C by A. Enge, implements the arithmetic of univariate polynomials with floating point coefficients of arbitrary precision, be they real (mpfr) or complex (mpc). Version 0.2 was published in May 2009, and is available at http://www.multiprecision.org/mpfrcx. Advanced asymptotically fast algorithms have been implemented, such as Karatsuba and Toom–Cook multiplication, various flavours of the FFT and division with remainder by Newton iteration. Special algorithms for symbolic computation, such as fast multievaluation, are also available.

Publishing mpfrcx is part of an ongoing effort to make A. Enge’s program for building elliptic curves with complex multiplication available. This program is a very important building block for cryptographic purposes as well as for primality proving (fastECPP).

5.4. TIFA

In late 2005, we hired J. Milan as ingénieur associé to help us developing and cleaning our programs. He first spent some time making a tour of publicly available implementations of the IEEE P-1363 cryptography standards. Following this study, it did not appear worthwhile to develop our own framework when others are approaching maturity and almost complete. He therefore switched to one of our other themes, namely writing integer factorization software for which the results can be guaranteed.

However, besides this quite daunting task, we have a more pragmatic, twofold-interest in fast factorization implementations for small numbers.

- Our first motivation is directly related to the ANR CADO project [22] we are involved in, together with other teams such as the INRIA project-team CACAO. The objective of the CADO project is to implement an optimized and distributed implementation of the Number Field Sieve (NFS), asymptotically the fastest integer factorization algorithm currently known. This algorithm needs to factor a lot of much smaller integers (about 80 bits for current factorization records). Since a recursive application of the NFS would be totally inefficient in practice, there is indeed a need for routines better suited to factor this wealth of smaller by-products.
- Our second motivation lies in our long-term commitment to produce identity cards for elliptic curves, in order to select curves with the properties needed for cryptographic use. An identity card requires the knowledge of the factorization of the order of the curve (about 200 bits for cryptographic use).

Hence, J. Milan began the development of the so-called TIFA library (short for Tools for Integer FActorization) in 2006. TIFA is made up of a base library written in C99 using the GMP library, together with stand-alone factorization programs and a basic benchmarking framework to assess the performance of the relative algorithms.

TIFA has been continuously improved during the last few years. As of November 2009, TIFA includes the following algorithms:

- CFRAC (Continued FRACtion factorization [59])
- ECM (Elliptic Curve Method)
• Fermat (McKee’s “fast” variant of Fermat’s algorithm [57])
• SIQS (Self-Initializing Quadratic Sieve [27])
• SQUFOF (SQUare FOrm Factorization [47])

In early 2009, disappointing comparisons to other factorization tools (such as the ones provided by PARI/GP) prompted J. Milan to undertake a major rewrite of his SIQS implementation. Together with other optimizations throughout the code base, this effort led to dramatic improvements, making TIFA’s SIQS more than twice as fast as PARI/GP’s version. TIFA’s SQUFOF and SIQS are now amongst the fastest available implementations. For tiny numbers (say between 100 to 160 bits), TIFA’s SIQS may even be the fastest.

J. Milan still plans to maintain and improve the library, particularly its ECM implementation which is, now, the only significant part of the software which is really behind the competition.

So far, TIFA has been kept internal to the TANC team and CADO project. Recently, we have received several requests from the community asking for access to this library. Consequently, we are in the process of making it public under an open source license (most probably the Lesser General Public License version 2.1 or higher). We plan to have it available before the end of the year, or at worst, in early 2010.

5.5. FAAST

The FAAST library is developed in C++ by L. De Feo and makes use of the NTL library. It implements the algorithms presented in [19], plus other algorithms needed by the author for its research on explicit isogenies.

Version 0.2.0, released on July 11 2009, is available at http://www.lix.polytechnique.fr/Labo/Luca.De-Feo/FAAST/. The source code is distributed under the General Public License version 2 or higher.

FAAST is a very efficient library for lattices of extensions of finite fields. Our aim is to add support for arbitrary finite fields, making it an essential building block for efficient computer algebra systems.

6. New Results

6.1. Algebraic curves over finite fields

6.1.1. Cardinality

Participants: Andreas Enge, François Morain.

A crucial ingredient for these records was A. Enge’s new algorithm [15] for computing modular equations of index greater than 2000. The algorithm computes bivariate modular polynomials by an evaluation and interpolation approach and relies on the ability to rapidly evaluate modular functions in complex floating point arguments. It has a quasi-linear complexity with respect to its output size, so that the performance of the algorithm is limited only by the size of the result: we have in fact been able to compute modular polynomials of degree larger than 10000 and of size 16 GB by a parallelised implementation of the algorithm, that uses mpc and mpfrx for the arithmetic of complex numbers and of polynomials with floating point coefficients, see Sections 5.2 and 5.3. For the point counting algorithm, the polynomials of prime level up to 6000 have been used. They occupy a disk space of close to 1 TB. Despite this progress, computing modular polynomials remains the stumbling block for new point counting records. Clearly, to circumvent the memory problems, one would need an algorithm that directly obtains the polynomial specialised in one variable.

We plan to make our new implementation available as an extension to the NTL library.

6.1.2. Isogenies

Participants: François Morain, Luca De Feo, Benjamin Smith.
Together with A. Bostan, B. Salvy (from projet ALGO), and É. Schost, F. Morain gave quasi-linear algorithms for computing the explicit form of a strict isogeny between two elliptic curves, another important block in the SEA algorithm [3]. This article contains a survey of previous methods, all applicable in the large characteristic case. Joux and Lercier have recently announced a $p$-adic approach for computing isogenies in all characteristic with the same complexity and based on our work.

For the small case, the old algorithms of Couveignes and Lercier were studied from scratch, and Lercier’s algorithm reimplemented in NTL by F. Morain, as a benchmark for other methods. In 2009 L. De Feo and É. Schost gave new asymptotically fast algorithms for arithmetics in Artin–Schreier towers [19]. The algorithms have been packaged in the C++ library FAAST and served as a basis for a new efficient implementation of Couveignes’ algorithm. Integration with F. Morain’s implementation of SEA is in progress.

In 2009, B. Smith gave new constructions of families of isogenies of Jacobians of high-genus curves; the existence of these families is remarkable. An article exhibiting twelve families of higher-genus hyperelliptic curves was submitted to the proceedings of AGCT 12 [20], and an article describing six infinite series of families (each giving isogenies in arbitrarily high dimension) is in preparation.

6.1.3. Discrete logarithms on curves

Participants: Andreas Enge, Jean-François Biasse, Benjamin Smith.

Jean-François Biasse has worked on an implementation of a subexponential algorithm to solve instances of the discrete logarithm problem on hyperelliptic curves of genus 8, in order to study the efficiency of a cryptosystem proposed by Edlyn Teske. This cryptosystem relies on the facility of solving this problem, as well as the difficulty of solving the discrete logarithm problem on an elliptic curve. This work was presented in October 2008 at the “Journée Nationales du Calcul Formel” in Luminy.

An extended version of B. Smith’s 2008 work on polynomial-time reductions of discrete logarithm problem instances from a large class of hyperelliptic curves of genus 3 to non-hyperelliptic curves of genus 3 (where Diem’s algorithm [31] can solve the discrete logarithm problem in time $\tilde{O}(q)$, a significant improvement over the previous best known $\tilde{O}(q^{4/3})$ algorithm for solving hyperelliptic genus 3 discrete logarithms due to P. Gaudry, E. Thomé, N. Thériault, and C. Diem [45]) has now appeared in the Journal of Cryptology [17].

6.1.4. Pairings

Participants: Andreas Enge, Benjamin Smith.

B. Smith has recently published joint work with S. Galbraith (Auckland), J. Pujolas (Lleida), and C. Ritzenthaler (Luminy), giving explicit constructions of distortion maps for supersingular genus 2 curves [16]. This enables practical pairing-based cryptography based on genus 2 curves.

6.2. Complex multiplication

Participants: Andreas Enge, François Morain.

A. Enge has been able to analyse precisely the complexity of class polynomial computations via complex floating point approximations [4]. Using techniques from fast symbolic computation, namely multievaluation of polynomials, and results from R. Dupont’s PhD thesis [32], he has obtained two algorithms which are quasi-linear (up to logarithmic factors) in the output size. The second algorithm has been used for a record computation of a class polynomial of degree 100,000, the largest coefficient of which has almost 250,000 bits. The implementation is based on GMP, mpfr, mpc and mpfrcx (see Section 5); the only limiting factor for going further has become the memory requirements of the final result.

Alternative algorithms use $p$-adic approximations or the Chinese remainder theorem to compute class polynomials over the integers. A. Enge and his coauthors have presented an optimized algorithm based on Chinese remaindering in [2] and improved the number theoretic bounds underlying the complexity analysis. They have shown that all three different approaches have a quasi-linear complexity, while the the floating point algorithm appeared to be the fastest one in practice.
Inspired by [2], A. Sutherland has come up with a new implementation of the Chinese remainder based algorithm that has led to new record computations [62]. Unlike the other algorithms, this approach does not need to hold the complete polynomial in main memory, but essentially only one coefficient at a time, which enables it to go much further. The main bottleneck is currently an extension of the algorithm to class invariants, which is work in progress by A. Enge.

Morai and Enge have contributed to the study of generalized Weber functions, enabling a partial classification for some cases [21].

6.3. Algebraic codes

Participants: Daniel Augot, Morgan Barbier.

6.3.1. List decoding of Reed–Solomon codes

This is a new activity of the TANC project-team, whose aim is to accelerate decoding algorithms of Reed–Solomon codes (with the Guruswami–Sudan algorithm), and of Algebraic Geometric codes. With Alexander Zeh, Daniel has found a relation between so-called key equations, which are the standard tool for decoding algebraic codes, and the new interpolation based algorithms [13]. The connection is established, and the next step is to use efficient algorithms, that are used for key equations, in the context of the Guruswami–Sudan algorithm.

6.3.2. List decoding of Algebraic Geometry codes

This is also a new activity for the TANC project team, started with the arrival of Guillaume Quintin, a new PhD student, supervised by Daniel Augot and Grégoire Lecerf (from the university of Versailles Saint-Quentin). These AG codes are a generalization of Reed–Solomon codes. Quintin did a first implementation of the factorisation step in MAGMA, to understand the algorithms, and the needed material. He is starting to rewrite the algorithm within the MATHEMAGIX framework.

6.3.3. List decoding of binary codes

Another new topic that begins with the arrival of Morgan Barbier is to study list decoding algorithms for codes defined over small alphabets. It was a challenging open problem until the publication of Wu [64], which achieves a high decoding radius for BCH codes, which are subfield subcodes of Reed–Solomon codes. This opens a new field of applications of these algorithms, and we have in mind to apply Wu’s algorithm for steganography, using the ideas of Fontaine and Galand [37]. They used Reed–Solomon codes, and it seems very natural to use the same ideas with BCH codes. Implementing Wu’s algorithm and applying it to steganography is the plan of Barbier’s thesis.

If the number of errors in a received word is less than the code’s error correction capacity, the decoding algorithm is guaranteed to return a single codeword. This property led to the term unique decoding, which has been (and still mostly is) the standard decoding method. However, in the last decade much attention has been given to so-called list decoding methods which can correct far more errors, at the expense of losing the uniqueness of the decoded word.

While the concept of list decoding code dates back from the fifties, the first interesting algorithm only appeared in 1990 when Madhu Sudan introduced a list decoding algorithm for Reed–Solomon codes able to correct up to $1 - \frac{2R}{n}$ errors, where $R = \frac{k}{n}$ is the code rate. Building upon this work, Sudaq and his student Venkatesan Guruswami then designed an improvement to Sudan’s algorithm correcting $1 - \sqrt{R}$ errors. Since then, a few other algorithms were proposed but the Guruswami–Sudan is still considered as the reference for list decoding.

As previously mentioned, list decoding trades the unicity of the corrected codeword for larger correction capabilities. Needless to say, if more errors are allowed, the list of returned codeword candidates will be larger. An important bound in list decoding is that of Johnson. Basically, if $n_e$, the number of errors allowed, is less than the Johnson bound $J_q(n, d)$, the size of the candidate list will grow polynomially with $n_e$. 
For a linear code $C$ defined over $\mathbb{F}_q$, of length $n$, dimension $k$ and minimal distance $d$, the Johnson bound is given by

$$J_q(n, d) = n \frac{q - 1}{q} \left( 1 - \sqrt{1 - \frac{q}{q - 1} \frac{d}{n}} \right).$$

Traditionally, we distinguish the binary codes, defined over $\mathbb{F}_2$, from the general case. For binary codes, the Johnson bound takes the simpler form

$$J_2(n, d) = \frac{n}{2} - \frac{n}{2} \sqrt{1 - \frac{2d}{n}}.$$

In the general case, provided $q/(q - 1) \approx 1$, we approximate $J_q(n, d)$ by

$$J(n, d) = n - n \sqrt{1 - \frac{d}{n}}.$$

The Johnson bound in the binary case is more interesting, since we are able to correct more errors for a given length and distance than in the general case.

6.3.4. Homomorphic encryption

Gentry’s breakthrough paper [46] has realized fully homomorphic encryption, albeit in a quite theoretical way. The properties of such schemes are that operations on the ciphertexts correspond to the same operations on the plaintext. This enables powerful applications, including querying encrypted databases. But Gentry’s scheme, although widely publicised, appears to be quite unpractical, since it implies huge ciphertexts.

Daniel Augot, Ludovic Perret, and Frederik Armknecht have devised a code-based encryption scheme based on evaluation codes, which has been given a particular instance with $q$-ary Reed–Muller codes. Although our scheme is secret-key, it still enables the desirable applications envisioned by Gentry, and is much more efficient with respect to ciphertext size and computational complexity of encryption operations. A paper has been submitted to the Eurocrypt 2010 conference [18].

6.4. Number fields

Participant: Jean-François Biasse.

Jean-François Biasse has made practical improvements to the sieving-based algorithm of Jacobson [51] for computing the group structure of the ideal class group of an imaginary quadratic number field. These improvements, based on the use of large prime variants combined with proper structured gaussian elimination, led to the computation of the structure of a class group corresponding to a number field with a 110-digit discriminant (whereas older techniques were limited to 90-digit discriminants). This work has been accepted for publication in the journal Advances In Mathematics of Communications. Biasse is currently working with Jacobson to adapt those techniques to the real quadratic extensions. Biasse also defined a class of number fields for which the the ideal class group, the regulator, and a system of fundamental units of the maximal order can be computed in subexponential time $L(1/3, O(1))$ (whereas the best known general algorithms have complexity $L(1/2, O(1)))$. This class of number fields is analogous to the class of $C_{ab}$ curves described by Enge and Gaudry [6]. This work has been submitted to Mathematics of Computation.

7. Contracts and Grants with Industry

7.1. Gemplus

This corresponds to É. Brier’s thesis on the use of (hyper-)elliptic curves in cryptology.
7.2. Industrial ANR

PACE: Pairings and advances in cryptology for e-cash, since 2007; with France Télécom R&D, Gemalto, NXP Semiconductors, Cryptolog International, École normale supérieure Paris and Université Caen

8. Other Grants and Activities

8.1. Network of excellence

Together with the SECRET project at INRIA Rocquencourt, the project TANC has taken part in ECRYPT, a NoE in the Information Society Technologies theme of the 6th European Framework Programme (FP6).

8.2. ANR

CADO (since 2006-09-01): two meetings (18-19/01/07 in Nancy for the kickoff and 21-21/06/07 in Paris).

9. Dissemination

9.1. Programme committees

Daniel Augot was in the programme committee of WCC2009 (http://wcc2009.org/).
Daniel Augot was in the programme committee of the Journées C2 at Fréjus (http://www-salsa.lip6.fr/~bettale/C2/).
Daniel Augot was in the programme committee of CESAR 2009 (http://www.rennes.supelec.fr/CESAR/).

9.2. Teaching

François Morain was in charge of half of the second year course Algorithmes et Programmation: du séquentiel au distribué (together with J.-M. Steyaert). He gives a cryptology course in Majeure 2. He is president of the Département d’Informatique. He has been representing École polytechnique in the Commission des Études du Master MPRI, since its creation in 2004.

At École polytechnique, A. Enge has taught computer science labs for the second year course Algorithmes et Programmation: du séquentiel au distribué. He has developed the practical module for the master-level cryptology course, centered around securing a network application in the Java cryptography framework JCE.
Daniel Augot gave some lectures on algebraic coding theory in the MPRI Master 2. He also taught a course in information theory at École polytechnique.

B. Smith taught the module on elliptic curve cryptography and pairings in the MPRI Master 2 course Cryptologie.

9.3. Seminars and talks


L. De Feo presented [19] at ISSAC ’09 in Seoul; it was awarded the SIGSAM Distinguished Student Author Award. L. De Feo gave an invited talk on Isogeny computation in small characteristic at the Workshop on Elliptic Curve Cryptography in Calgary and a contributed talk on the same subject at the Journées Arithmétiques in Saint-Étienne. He gave invited talks on Fast arithmetics in Artin–Schreier towers at the Canadian Mathematical Society Winter meeting in Ottawa and at the Rencontres ArithmÂtique de l’Informatique MathÂmatique in Lyon.
Jean-François Biasse gave a talk on Improvements in the computation of ideal class group in imaginary quadratic number fields at the conference CHILE2009, Frutillar, Chile, March 2009.

9.4. Vulgarization and Summer schools
B. Smith gave three lectures on advanced topics in elliptic curves at the ECrypt II Winter School on Mathematical Foundations in Cryptography in February 2009 in Lausanne.

9.5. Editorship
A. Enge has been an editor of Designs, Codes and Cryptography since 2004.
D. Augot is a guest editor, with Jean-Charles Faugère and Ludovic Perret of a special issue of the Journal of Symbolic Computation, on Gröbner Bases Techniques in Cryptography and Coding Theory.

9.6. Awards
L. De Feo has received the SIGSAM Distinguished Student Paper Award for [19] at ISSAC '09 in Seoul.

9.7. Thesis committees
D. Augot was a referee for Eleonara Guerinni’s thesis in Trento (27/04/2009).
D. Augot was in the defense committee of Roberto Speicys Cardoso.

9.8. Research administration
A. Enge is correspondent for European affairs of INRIA Saclay–Île-de-France (formerly INRIA Futurs) since 2006 and correspondent for international affairs since 2007.
F. Morain represents INRIA in the Conseil d’UFR 929 Maths Université Paris 6 since September 2005.

10. Bibliography

Major publications by the team in recent years


Year Publications

Articles in International Peer-Reviewed Journal


International Peer-Reviewed Conference/Proceedings


Other Publications


References in notes


