Statistical Zero-Knowledge Proofs from Diophantine Equations*

Helger Lipmaa

Laboratory for Theoretical Computer Science
Department of Computer Science and Engineering
Helsinki University of Technology
P.O.Box 5400, FIN-02015 HUT, Espoo, Finland
helger@tcs.hut.fi

Abstract. A family (S_t) of sets is p-bounded Diophantine if S_t has a representing p-bounded polynomial $R_{S,t}$, s.t. $x \in S_t \iff (\exists y)[R_S(x;y) = 0]$. We say that (S_t) is unbounded Diophantine if additionally, $R_{S,t}$ is a fixed t-independent polynomial. We show that p-bounded (resp., unbounded) Diophantine set has a polynomial-size (resp., constant-size) statistical zero-knowledge proof system that a committed tuple x belongs to S. We describe efficient SZK proof systems for several cryptographically interesting sets. Finally, we show how to prove in SZK that an encrypted number belongs to S.

Keywords: Diophantine equations, integer commitment, statistical zero knowledge.

1 Introduction

A set S of ordered n-tuples of positive integers is called *Diophantine* if there is a representing polynomial $R_S(x;y)$ with integer coefficients such that a given n-tuple $x = (x_1, \dots, x_n)$ belongs to S iff there exists a tuple $y = (y_1, \dots, y_m)$ of integers *witnesses*) for which $R_S(x;y) = 0$: I.e., $x \in S \iff (\exists y)[R_S(x;y) = 0]$. Based on earlier work by Davis, Putnam and Robinson, in 1970 Matiyasevich [Mat70] showed that every recursively enumerable set is Diophantine (this is known as the Davis-Putnam-Robinson-Matiyasevich or the DPRM theorem), solving finally Hilbert's tenth problem.

We are interested in cryptographic applications of this result. For this, we look at families $\mathcal{S} = (S_t)$ of Diophantine sets and say that $\mathcal{S}$ is p-bounded Diophantine if there exists a uniform family $\mathcal{R}_{\mathcal{S}} = (R_{\mathcal{S},t})$ of p-computable polynomials, such that for every t, $x \in S_t$ iff there exists a witness y such that $R_{S,t}(x;y) = 0$. We say that $\mathcal{S}$ possesses a certifier $\mathcal{C}_{\mathcal{S}} = (C_{\mathcal{S},t})$ iff $R_{\mathcal{S},t}(x; C_{\mathcal{S},t}) = 0$ for

* Preliminary version, 05.11.01

every t. If $R_{\mathcal{S},t}$ is the same polynomial for all t then we say that $\mathcal{S}$ is unbounded Diophantine. In the latter case, we often identify $\mathcal{S}$ and $\bigcup_t S_t$.

Given a statistically hiding integer commitment scheme like [DF01] where it is possible to prove in statistical zero-knowledge (SZK) that two committed integers are in an additive or multiplicative relationship, one can also prove in SZK any polynomial relationship between a tuple of committed numbers. Thus, if $\mathcal{S}$ is p-bounded Diophantine then one can prove in SZK that committed tuple x belongs to S_t by proving that x together with another committed tuple y satisfies $R_{\mathcal{S},t}(x;y) = 0$. We call this membership proof a Diophantine membership proof (for $\mathcal{S}$).

When $\mathcal{S}$ is unbounded Diophantine, the resulting SZK proof for S has interaction length $\Theta(\log_2 T)$, where T is the a priori maximum of any input x_i or y_i in the concrete application. The situation changes when $\mathcal{S}$ is p-bounded Diophantine. In such a case, each S_t could have a designated proof system. This means that the resulting SZK proof has polynomial communication complexity; on the other hand, even some easy p-bounded Diophantine sets seem not to be unbounded. As an example, it is known that the Diophantine set $S = \{(x_1, x_2, x_1^{x_2}) : (x_1, x_2) \in \mathbb{Z}\}$ has a representing polynomial R_S in 14 witnesses y_i [JSWW76]. However, several witnesses y_i of R_S have superpolynomial length in $|x|$. On the other hand, the family $\mathcal{S} = (S_t)$, $S_t = \{(x_1, x_2, x_1^{x_2}) : x_1 \in \mathbb{Z} \wedge x_2 \in \mathbb{Z}_T\}$, $T = 2^t$, is p-bounded Diophantine; it even has a certifier!

We present a number of p-bounded families $\mathcal{S}$ with certifier that have cryptographic relevance; all such families have efficient SZK proof systems. We show that $\mathcal{S} = [0, \infty)$ is unbounded and present corresponding SZK proof. This proof bases on the result of Lagrange that every nonnegative integer is a sum of four squares and on the randomized algorithm of Rabin and Shallit [RS86] that finds these squares in $O(t^4)$ bit-operations. Note that efficient proof system for $[0, \infty)$ is crucially important for our entire framework, since in the definition of Diophantine sets one often requires the solutions to be positive. Moreover, our proof system for $[0, \infty)$ requires about 20% more communication than Boudot's membership proof [Bou00] for $[0, \infty)$; however, differently from Boudot's proof system, our proof system is perfectly complete.

Based on the proof for nonnegativity, we show for an arbitrary function $g : \mathbb{Z}^n \to \mathbb{Z}$ how to construct an efficient perfectly complete SZK proof system for proving that $x_n \neq g(x_1, \dots, x_{n-1})$, given that there exists an efficient SZK proof for proving that $x_n = g(x_1, \dots, x_{n-1})$. We derive an SZK proof system for $x_3 = \gcd(x_1, x_2)$. More examples will be given in Section 4.

In practice, it is often necessary to show that an *encrypted* (as opposed to a committed) value belongs to some set S. For many interesting sets S we are not aware of an *efficient* membership proof for S that would use only the corresponding public-key cryptosystem and no other primitives. Instead, we can build up a membership proof for encrypted numbers by using a SZK proof that a committed integer and an encrypted integer are equal (modulo the message space size), and then applying our SZK membership proof for S to the committed number. Sequential composition of these two proofs is naturally SZK. Finally, we construct an efficient SZK proof that discrete logarithm of encrypted value belongs to an interval, and show how to use it in the Damgård-Jurik voting scheme [DJ01] to achieve shorter proofs of vote correctness.

Road-map. Necessary preliminaries are given in Section 2. We will describe a Diophantine membership proof for $[0, \infty)$ in Section 3. Extension to general Diophantine equations is shown in Section 4. Section 5 presents protocols that allow to apply our proofs together with homomorphic cryptosystems. Finally, the appendix contains technical proofs of some theorems, a short description of the Rabin-Shallit algorithm and an discussion about proof systems for exponential relationship.

2 Preliminaries

Homomorphic encryption. A public-key cryptosystem $\mathit{\Pi}$ is a triple of efficient algorithms, $\mathit{\Pi} = (G, E, D)$, where G is the key generation algorithm, E is the encryption algorithm and D is the decryption algorithm. Throughout this paper, let t be the security parameter. Let $\mathcal{M}$ (resp., $\mathcal{C}$ and $\mathcal{R}$) denote the message space (resp., the ciphertext space and the randomness space), corresponding to a fixed value of t. We assume that all three sets $(\mathcal{M}, \mathcal{R}, \mathcal{C})$ are Abelian groups, with $\mathcal{C}$ written multiplicatively. We say that public-key cryptosystem $\mathit{\Pi} = (G, E, D)$ is *homomorphic* if $E_K(m_1 + m_2; r_1 + r_2) = E_K(m_1; r_1)E_K(m_2; r_2)$. Some example homomorphic cryptosystems are the Paillier cryptosystem [Pai99] and the Damgård-Jurik cryptosystem [DJ01]. Let $M := \lceil \log_2 |\mathcal{M}| \rceil$, $C := \lceil \log_2 |\mathcal{C}| \rceil$ and $R := \lceil \log_2 |\mathcal{R}| \rceil$. We will assume in our calculations that $M = R = 1024$ and $C = 2048$.

Proofs-of-knowledge. For a bit-string α and predicate $P(\cdot)$, $\mathrm{PK}_y(\alpha : y = P(\alpha))$ is a proof-of-knowledge between two parties that given a publicly known value y, the first party knows a value of α, such that the predicate $P(\alpha)$ is true. To simplify notation, we will always denote the values, knowledge of which has to be proven, by Greek letters. Additionally, we assume that the scope of such variables lies within one proof-of-knowledge. E.g., $\mathrm{PK}\,(c = E_K(m; \rho))$ is a proof that given a ciphertext c, plaintext m and a public key K, the prover knows a nonce ρ such that $c = E_K(m; \rho)$.

Most of the protocols in this paper are three-round interactive honest-verifier (statistical) zero-knowledge (abbreviated as HVZK or HVSZK, resp.) proof systems for proofs-of-knowledge of type $\mathrm{PK}\,(y = P(\alpha))$. One usually proves that such protocols are (1) Complete: That is, a honest verifier accepts a honest prover with probability $1 - \mathrm{neg}(t)$, where $\mathrm{neg}(t)$ is a negligible function in t; (2) Honest-verifier (statistical) zero-knowledge: Even without knowing α, one can generate a view of the protocol that has distribution, indistinguishable from (or statistically close to) the distribution of real views in the case when the verifier is honest; (3) Specially sound: Given two views of the protocol that begin with the same move but have different second moves, one can compute secret α. A proof system is called *perfectly complete* if a honest verifier always accepts a honest prover. An honest-verifier (statistical) zero-knowledge proof system can be made noninteractive by using the Fiat-Shamir heuristic [FS86] in the random-oracle model. For this, we introduce a random oracle $H : \{0,1\}^* \to \{0,1\}^{2t}$.

Damgård-Fujisaki integer commitment scheme. A (statistically hiding) integer commitment scheme Com allows a a participant P (a polynomial-time algorithm) to commit to any integer $m \in \mathbb{Z}$, so that (1) For any $m_1, m_2 \in \mathbb{Z}$, the distributions $Com_K(m_1)$ and $Com_K(m_2)$ are statistically close; and (2) It is intractable for P to find $m_2 \neq m_1$, such that $Com_K(m_1) = Com_K(m_2)$. The first integer commitment scheme that allowed an efficient proof system that committed integers are in multiplicative relationship was proposed by Fujisaki and Okamoto [FO97], but soundness proof of this scheme was later found to be flawed. This flaw has been only recently corrected by Damgård and Fujisaki, who proposed a new integer commitment scheme in [DF01]. Since the latter scheme is relatively new, we will give next give a longer description of it.

Let G be a suitable group. (We refer to [DF01] for the exact definition of suitable, but remark that G can be chosen as $\mathbb{Z}_n$ for RSA modulus $n = pq$, where $p \equiv q \equiv 3 \pmod 4$, $\gcd(p-1, q-1) = 2$, and the parts of $p-1$, $q-1$ with prime factors less than F are $O(t)$. One may choose B as $B \leftarrow \lceil \log_2 n \rceil + 1$. Then the security follows from the strong RSA assumption.) While the prover knows a reasonably close upper bound $2^B > \mathrm{ord}\, G$ to the order of G, he does *not* know the order itself. A large number F is chosen, such that it is still feasible to factor numbers that are smaller than F. Say, $F = O(t^{\log t})$ though in our calculations we will take $F = 2^{80}$. During the setup phase of Damgård-Fujisaki integer commitment scheme, P and V agree on a group G and a large integer F. Verifier V chooses a random F-rough element $h \in G$ and a random $x \in [0, 2^{B+t})$. Let $g = h^x$. V sends the public key $K = (g, h)$ to P and then proves in SZK that $g \in \langle h \rangle$. When committing to $m \in \mathbb{Z}$, P chooses a random $r \leftarrow [0, 2^{B+t})$ and sends $Com_K(m; r) := g^m h^r$ to V. To open commitment c, P sends (m, r, b), s.t. $c = g^m h^r b$ and $b^2 = 1$.

Protocol 1 SZK proof system for multiplicative relationship between committed numbers.

1. Prover P chooses a random $m_1 \leftarrow_R [0, 2^t FT)$, $r_1 \leftarrow_R [0, 2^{B+2t} F)$, $r_2 \leftarrow_R [0, 2^{B+2t} FT)$ and sends $c_5 \leftarrow g^{m_1} h^{r_1}$, $c_6 \leftarrow c_1^{m_1} h^{r_2}$ to V.
2. Verifier V generates a random $e \leftarrow_R [0, F)$ and sends it to P.
3. Prover sends $m_2 = m_1 + e\mu_2, r_3 \leftarrow r_1 + e\rho_2$ and $r_4 \leftarrow r_2 + e(\rho_3 - \mu_2 \rho_1)$ to V.
4. Verifier checks that $g^{m_2} h^{r_3} c_2^{-e} = c_5$ and $c_1^{m_2} h^{r_4} c_3^{-e} = c_6$.

One can build different SZK proof systems for different relationships between committed numbers μ_i. In all such proof systems, prover and verifier have to fix a priori upper bound T_i to every input μ_i. Proof system is guaranteed to be SZK only if $|\mu_i| < T_i$. In most of the protocols, proof complexity depends on $\log_2 T_i$, and hence it is beneficial to compute as precise values of T_i as feasible. At least it must be the case that $\log_2 T_i = t^{O(1)}$.

Let $\mathcal{C}_{Com}$ denote the commitment space of the used integer commitment scheme (in this concrete case, $\mathcal{C}_{Com} = G$) and let $C_{Com} := \lceil \log_2 |\mathcal{C}_{Com}| \rceil$, with security parameter understood from the context. We will assume in our calculations that $C_{Com} = 1024$.

Proof system for multiplicative relation. For their own integer commitment scheme, Damgård and Fujisaki [DF01] constructed an efficient proof system for $\mathrm{PK}\left(\left(\bigwedge_{i=1}^{3} c_1 = Com_K(\mu_1; \rho_1)\right) \wedge \mu_3 = \mu_1 \mu_2\right)$; that is, a proof that a committed integer μ_3 is product of another two committed integers. We will next give a description of this proof system. Let Com be the Damgård-Fujisaki integer commitment scheme, let t be the security parameter, let $K = (g, h)$ be the public key and let $\log_2 T = t^{O(1)}$. Then Protocol 1 is a complete, honest-verifier SZK, specially sound proof-of-knowledge for multiplicative relationship, assuming that $\mu_i \in [0, T)$ and $\rho_i \in [0, 2^{B+t})$:

Noninteractive version of this proof is (e, m_2, r_3, r_4), with $e = H(c_5, c_6)$, where V verifies that $e = H(g^{m_2} h^{r_3} c_2^{-e}, c_1^{m_2} h^{r_4} c_3^{-e}) \pmod{2^t}$. With parameters $C_{Com} = 1024$, $F = 2^{80}$, $T = 2^{1024}$, $t = 160$ and $B = 1024$, the noninteractive proof has length $4 \log_2 F + 2 \log_2 T + 6t + 2B = 320 + 2 \log_2 T + 480 + 2048 = 2848 + 2 \log_2 T$ bits or $356 + \frac{1}{4} \log_2 T$ bytes. When $T = 1024$ then this length is 1024 bits. When $\mu_1 = \mu_2$ (i.e., Protocol 1 is used to prove that μ_3 is a square), one can assume that $\mu_1 = \mu_2 \leq 2^{512}$ and the noninteractive proof is 484 bytes long. Note that only μ_2 has to be less than T, thus if it is known a priori that one of the arguments might be much greater than another one, it will make sense to use this argument as μ_1.

Boudot's membership proof for $[0, \infty)$**.** Boudot's proof system [Bou00] for PK $(c = Com_K(\mu; \rho) \wedge (\mu \geq 0))$ consists of several steps: First, represent μ as $\mu_1^2 + \mu_2$, where $\mu_1, \mu_2 = O(\sqrt{\mu})$. Since $\mu_1^2 \geq 0$, one is now only left to prove that $\mu_2 \geq 0$. Second, one can prove that $\mu_2 \geq -\theta$ for $\theta{:=}2^t F T^{1/2}$ by using the membership proof with tolerance by Chan, Frankel and Tsiounis [CFT98]. Now, one has proved that $\mu \geq -\theta$. Fourth, one can achieve zero tolerance by a priori multiplying μ with a suitably chosen constant 2^x such that $\theta < 2^{x/2}$. In this case, $2^{x/2}\mu > -2^{x/2}$ or $\mu_2 > -1$. When modified for Damgård-Fujisaki integer commitment system, this proof system has completeness error $\Theta(1/F)$; its noninteractive version is $1166 + \frac{1}{8}\left\lceil \log_2 T^{1/2} \right\rceil$ bytes long.

Algebraic complexity theory. A p-family over $\mathbb{Z}$ is a sequence $f = (f_t)$ of multivariate polynomials such that the number of variables as well as the degree of f_t are polynomially bounded (p-bounded) functions of n. Let $L(f_t)$ (resp., $L_*(f_t)$) denote the total complexity of f_t, that is, the minimum number of arithmetic operations $\{\cdot, +, -\}$ (resp., $\{\cdot\}$) sufficient to compute f_t from the input variables and constants in $\mathbb{Z}$ by a straight-line program. We call a p-family f *p-computable* iff the map $t \mapsto L(f_t)$ is p-bounded.

3 Proof that a Committed Number is Nonnegative

Before treating the general situation of an arbitrary (p-bounded) Diophantine family $\mathcal{S}$, we will give an efficient membership proof for $[0, \infty)$, i.e., that a committed number is nonnegative. There are a few good reasons for proceeding in this order. First, the approach we use in constructing this proof system has much in common with the general solution, and hence it serves as a motivational example. Second, for many $\mathcal{S}$, there is a more efficient representing polynomial when we consider only nonnegative solutions to this equation; for such an $\mathcal{S}$ it might make sense to use this polynomial and then to prove for every witness that it is nonnegative. Third, a proof system for $[0, \infty)$ is interesting in its own right, since it is used in many cryptographic protocols. (See [Bou00] for examples.)

The next theorem is crucial for our membership proof:

Theorem 1. *An integer μ can be represented as $\mu = \mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2$ with integer μ_i iff $\mu \geq 0$. Moreover, if $\mu \geq 0$ then the representation $(\mu_1, \mu_2, \mu_3, \mu_4)$ can be computed efficiently.*

Proof. If $\mu \geq 0$, such μ_i exist by a well-known result of Lagrange from 1770. Rabin and Shallit [RS86] proposed a probabilistic polynomial-time algorithm (described in Appendix B) for computing the representation. On the other hand, no negative number is a sum of four squares. $\square$

Protocol 2 SZK proof system for nonnegative integers.

1. Prover P represents μ as $\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2$, using the Rabin-Shallit algorithm. For $i \in [1,4]$, P chooses random $\rho_i \leftarrow_R [0, 2^{B+t})$ such that $\sum_i \rho_i = \rho$; P chooses random $m_{1i} \leftarrow_R [0, 2^t FT^{1/2})$, $r_{2i} \leftarrow_R [0, 2^{B+2t}F)$, $r_3 \leftarrow_R [0, 2^{B+2t}FT^{1/2})$, and lets $c_{1i} \leftarrow g^{\mu_i}h^{\rho_i}$, $c_{2i} \leftarrow g^{m_{1i}}h^{r_{2i}}$, $c_3 \leftarrow \prod_i c_{1i}^{m_{1i}} \cdot h^{r_3}$. Prover sends $((c_{1i}, c_{2i})_{i=1}^4, c_3)$ to V.
2. V generates a random $e \leftarrow_R [0, F)$ and sends it to P.
3. P computes $m_{2i} = m_{1i} + e\mu_i$, $r_{4i} \leftarrow r_{2i} + e\rho_i$, $i \in [1,4]$, and $r_5 \leftarrow r_3 + e\sum_i (1-\mu_i)\rho_i$. P sends $((m_{2i}, r_{4i})_{i=1}^4, r_5)$ to V.
4. V checks that (a) $g^{m_{2i}}h^{r_{4i}}c_{1i}^{-e} = c_{2i}$ for $i \in [1,4]$, and (b) $\prod_i c_{1i}^{m_{2i}} \cdot h^{r_5}c^{-e} = c_3$.

Briefly, during our proof system for $[0, \infty)$, prover first uses the Rabin-Shallit algorithm to represent μ as $\mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2$ (or a *witness*). After that, he proves to V in SZK that she knows such a representation. Complete proof system is given by the next theorem:

Theorem 2. *Let Com be the Damgård-Fujisaki integer commitment scheme, let t be the security parameter and let $T = n^{O(1)}$. Let $K = (g, h)$ be the public key. Protocol 2 is a perfectly complete, honest-verifier SZK and specially sound proof system for* $\mathrm{PK}\left(c = Com_K(\sum_{i=1}^4 \mu_i^2; \rho)\right)$*, or equivalently in the epistemic sense, for* $\mathrm{PK}\,(c = Com_K(\mu) \wedge \mu \geq 0)$*, if $\mu < T$.*

(Proof of this theorem is given in Appendix A.) Noninteractive version of this protocol is $((c_{1i})_{i=1}^4; e; (m_{2i}, r_{4i})_{i=1}^4, r_5)$, where the verifier checks that $e = H((c_{11})_{i=1}^4, (g^{m_{2i}}h^{r_{2i}}c_{1i}^{-e})_{i=1}^4, c^{-e}\prod_i c_{1i}^{m_{2i}} \cdot h^{r_5}) \pmod{2^t}$. The length of noninteractive proof system is $4C_{Com} + 2t + 4(B + 3t + 2\log_2 F + \frac{1}{2}\log_2 T) + B + 2t + \log_2 F + \frac{1}{2}\log_2 T = 4096 + 160 + 4 \cdot (1024 + 240 + 160) + 1024 + 160 + 80 + \frac{5}{2}\log_2 T = 11216 + \frac{5}{2}\log_2 T$ bits or $1402 + \frac{5}{16}\log_2 T$ bytes.

Hence, noninteractive version of Protocol 2 is $\approx 20\%$ longer than Boudot's proof system for the same problem. However, our proof system enjoys the property of perfect completeness, while Boudot's proof system has completeness error $\Theta(1/F)$. This result in interesting by itself, in particular since by a result of Vadhan, no complexity-preserving strong black-box transformation can eliminate completeness error [Vad00].

4 Membership Proofs from Diophantine Equations

The ideas used in Section 3 to build an efficient membership proof for $[0, \infty)$ can be generalized for proving in SZK that the committed tuple $\mu = (\mu_1, \ldots, \mu_m)$ belongs to many other (not necessary finite) sets $S \subset \mathbb{Z}^m$. For this we have to introduce a more complexity-theoretic flavor of Diophantine sets.

Definitions. Let $\mathcal{S} = (S_t)$ be a family of sets, such that the elements in S_t have length that is polynomial in t. We say that the family $\mathcal{S}$ is *p-bounded Diophantine* if there exists a uniform family $\mathcal{R}_S = (R_{S,t})$ of p-computable polynomials with integer coefficients, such that for every t, $x \in S_t$ iff there exists a *witness* y such that $R_{S,t}(x; y) = 0$. We say that the family $\mathcal{S}$ is *possesses a certifier $\mathcal{R}_S$* if there exists a family $\mathcal{R}_S = (R_{S,t})$ of p-bounded polynomials and a family of polynomial-time algorithms $\mathcal{C}_S = (C_{S,t})$, such that for every t, $x \in S_t$ iff $R_{S,t}(x; C_{S,t}) = 0$. If $R_{S,t}$ is a polynomial that does not depend on t, we say that *$\mathcal{S}$ has an unbounded Diophantine membership proof*. In the latter case one often (but not always) assumes that $\mathcal{S} = \bigcup_t S_t$ is a set itself.

Given a statistically hiding integer commitment scheme like [DF01] where one can prove in SZK that two committed integers are in an additive or in a multiplicative relationship, one can prove in SZK that a polynomial relationship holds between a number of committed integers, by using the methodology of [FO97]. Now, let $\mathcal{S} = (S_t)$ have a p-bounded Diophantine membership proof. By using an integer commitment scheme, one can then prove in SZK that he knows a y, s.t. $R_{\mathcal{S},t}(x; y) = 0$. Thus, such a proof is a valid proof system for PK $(c_i = Com_K(\mu_i; \rho_i) \wedge \cdots \wedge c_n = Com_K(\mu_n; \rho_n) \wedge (\mu_1, \ldots, \mu_n) \in S_t)$.

Discussion. The first example SZK proof system (for $\mathcal{S} = [0, \infty)$) was already given in Section 3. As seen from this example, communication complexity of such a proof system depends linearly on $L_*(S_t)$ and on values $\log_2 T_i$, where T_i is an a priori upper bound on input x_i. If the membership proof is unbounded then $L_*(S_t)$ is a constant and communication complexity is just a linear function of $\log_2 T_i$-s. (That is, on the input length, being hence optimal.) This explains why we are especially interested in unbounded Diophantine families $\mathcal{S}$. Surprisingly, as we will see later on, there are many cryptographically interesting unbounded Diophantine families $\mathcal{S}$.

Certifier is needed in situation where a party in a cryptographic protocol needs to prove that he has performed correct calculations over some data that were received from sources not controlled by him. On the other hand, when the prover can generate a committed number by himself and just has to prove that this number belongs to some correct set (e.g., is nonnegative, or is composite), certifier is not necessary.

Often, in the definition of Diophantine sets it is required that the witness must be nonnegative. However, if $\mathcal{S}$ has a representing polynomial $R_S(x; y)$ then $x \in S_t \iff (\exists y, y' \in [0, \infty))[R_{S,t}(x; y_1 - y'_1, \ldots, y_m - y'_m) = 0]$. One the other hand, if S has a nonnegative representing polynomial $R'_S(x; y)$ with nonnegative witnesses then it is represented by $R'_{S,t}(p_{11}^2 + \cdots + p_{14}^2, \ldots, p_{n1}^2 + \cdots + p_{n4}^2; q_{11}^2 + \cdots + q_{14}^2, \ldots, q_{m1}^2 + \cdots + q_{m4}^2)$.

$\mathcal{S}$	$R_{\mathcal{S}}$	$C_{\mathcal{S}}$
$[a, \infty)$	$y_1^2 + y_2^2 + y_3^2 + y_4^2 - x + a$	[RS86]
$[-\infty, b)$	$y_1^2 + y_2^2 + y_3^2 + y_4^2 + x - b$	[RS86]
$\{(x_1, x_2, x_3) : \gcd(x_1, x_2) \mid x_3\}$	$x_1 y_1 + x_2 y_2 - x_3$	Extended Euclidean
$\{(x_1, x_2) : x_2 \mid x_1\}$	$x_1 - x_2 y_1$	$y_1 \leftarrow x_1 / x_2$

Table 1. Some unbounded Diophantine sets $\mathcal{S}$ with representing polynomials and certifiers.

Next, let $\mathcal{S}_1$ and $\mathcal{S}_2$ be Diophantine sets with representing polynomials $R_{\mathcal{S}_i}$, s.t. $x \in \mathcal{S}_i \iff (\exists y)[R_{\mathcal{S}_i}(x; y) = 0]$. Then (1) $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ is Diophantine, with $R_{\mathcal{S}_1 \cup \mathcal{S}_2}(x; y, z) = R_{\mathcal{S}_1}(x; y) R_{\mathcal{S}_2}(x; z)$; and (2) $\mathcal{S} = \mathcal{S}_1 \cap \mathcal{S}_2$ is Diophantine, with $R_{\mathcal{S}_1 \cap \mathcal{S}_2}(x; y, z) = P_{\mathcal{S}_1}^2(x; y) + P_{\mathcal{S}_2}^2(x; z)$. Therefore, if $\mathcal{S}_1$ and $\mathcal{S}_2$ are (p-bounded) Diophantine then so are $\mathcal{S}_1 \cup \mathcal{S}_2$ and $\mathcal{S}_1 \cap \mathcal{S}_2$; thus the latter sets also have Diophantine membership proofs. However, described compositions add an extra multiplication per every union and two multiplications per every intersection, which is an undesirable overhead. A more efficient way is to use the methodology of Cramer, Damgård and Schoenmakers [CDS94] of composing several proofs-of-knowledge; when using their methods, we get Diophantine membership proofs for $\mathcal{S}_1 \cup \mathcal{S}_2$ and $\mathcal{S}_1 \cap \mathcal{S}_2$ that do not require extra multiplications.

The methodology of [CDS94] is is limited to composing sets by using the union and interaction but not complementing. (This is not surprising. More generally, the complement of a recursively enumerable set is not always recursively enumerable.) For example, one cannot derive from the framework of [CDS94] an efficient proof that a committed integer μ is *not* equal to some constant a, starting only from positive proofs. (Although efficient proof systems for $\mu \neq a$ exist [MS97].) An important feature of our approach is that we can implement nonmembership proofs. Reason for this lies in the flexibility that we have when choosing the sets $\mathcal{S}$. More precisely, efficient negative proofs are possible since some of these sets $\mathcal{S}$ (like $[a, \infty)$ and $(-\infty, b]$) are infinite, but their intersections are finite. We will present corresponding examples in a few paragraphs.

Examples. Some unbounded Diophantine sets $\mathcal{S}$ together with their representing polynomials and certifiers are depicted by Table 1. A few sets that are intersections of simpler sets are depicted by Table 2. Note that noninteractive version of SZK proof system for $\mu \in [a, b]$ has interaction length $2804 + \frac{5}{8}\log_2(b - a)$ bytes, since one can set $T \leftarrow b - a$. The set of composite numbers does not have a certifier unless factoring is easy. Other sets in this table have a certifier.

Let us look at more interesting examples. An efficient Diophantine membership proof system for PK ($c = Com_K(\mu) \wedge \mu \notin [a, b]$) can be based on the fact that $x \notin [a, b]$ iff $(\exists y)[R_{[b+1, \infty)}(x; y) = 0] \vee (\exists y)[R_{(-\infty, a-1]}(x; y) = 0]$.

$\mathcal{S}$	To prove that you know such x_i, show that ...
$[a, b]$ $(a \leq b)$	$x \in [a, \infty), x \in (-\infty, b]$
$\mathbb{Z} \setminus \{2^s : s \in \mathbb{Z}\}$	$x = y_1(2y_2 + 1) - x, y_2 > 0$
Set of composite numbers	$x = y_1 y_2, y_1 > 1, y_2 > 1$
Set of nonsquares	$x = y_1^2 + y_2, x = (y_1 + 1)^2 - y_3, y_2 > 0, y_3 > 0$
$\{(x_1, x_2, x_3) : x_1 = x_2 \pmod{x_3}\}$	$x_1 = x_2 + x_3 y_1, x_1 < x_3.$
$\{(x_1, x_2, x_3) : x_3 = \gcd(x_1, x_2)\}$	$\gcd(x_1, x_2) \mid x_3, x_3 \mid x_1, x_3 \mid x_2$

Table 2. Some more examples.

If $a = b$, this yields a Diophantine proof system that a committed number is not equal to some constant a. Moreover, one can prove that two committed numbers x_1 and x_2 are not equal, by proving that $(\exists y)[(x_1 = x_2 + y) \wedge (y_1 \neq 0)]$. This approach can be generalized to an arbitrary function g. Let $\mathcal{S} = \{(x_1, \ldots, x_n) : x_n = g(x_1, \ldots, x_{n-1})\}$, g a function, be a p-bounded (resp., unbounded) Diophantine set with representing polynomial $R_{\mathcal{S}} = (R_{\mathcal{S},t})$. Then $\mathcal{S}' = \{(x_1, \ldots, x_n) : x_n \neq g(x_1, \ldots, x_{n-1})\}$ is p-bounded (resp., unbounded) Diophantine: Namely, it suffices to show that $y = g(x_1, \ldots, x_n)$ and that $x_0 \neq y$.

Since the set of primes is recursively enumerable the DPRM theorem says that it is also Diophantine. Jones, Sato, Wada and Wiens [JSWW76] described a certain Diophantine equation system of 14 equations in 26 variables that has a positive integral solution iff one of parameters is a prime. Thus, one can prove that a committed number is prime by proving that he knows an integral solution to this equation system, and then proving that all solutions are positive. However, several witnesses y_i have superpolynomial length in $|x|$ and hence the set of primes is not known to have a Diophantine membership proof. Finally, a longer example for exponential relationship will be given in Appendix C

5 Applications to Encrypted Numbers

Proof that committed number = encrypted number. Let (G, E, D) be a homomorphic public-key cryptosystem with $\mathcal{M} = \mathbb{Z}_M$ and public key K_e. In cryptographic protocols, one often needs a zero-knowledge proof that an *encrypted* number belongs to some set S. For many cryptographically interesting sets $S \in \mathbb{Z}^n$, we are not aware of any efficient membership proofs for PK $((\bigwedge_i c = E_{K_e}(\mu_i; \rho_i)) \wedge (\mu_1, \ldots, \mu_n) \in S)$ that base solely on the security of the used encryption scheme. However, the next methodology enables to construct such proof systems, assuming that $S \subseteq \mathbb{Z}_M^n$ (an example S could be $S = \{(x_1, x_2, x_1^{x_3} : (x_1, x_2, x_1^{x_2}) \in \mathbb{Z}^3\})$ and there is an efficient proof system for PK $(c = Com_{K_c}(\mu;\rho) \wedge \mu \in S)$, where Com is an integer commitment scheme with key K_c:

1. For every i, P creates a random r_i and sends $c'_i = Com_{K_c}(\mu_i; r_i)$ to verifier.
2. For every i, P that PK $(c_i = E_{K_e}(\mu_i;\rho_i) \wedge c'_i = Com_{K_c}(\mu_i;\rho'_i))$.
3. Finally, P that PK $((\bigwedge_i c_i = Com_{K_c}(\mu_i;\rho)) \wedge (\mu_1, \ldots, \mu_n) \in S)$.

Note that $E_{K_e}(m + kM; r) = E_{K_e}(m; r)$ and therefore in the second step, P should prove that PK $(c_i = E_{K_e}(\mu_i \pmod M;\rho_i) \wedge c'_i = Com_{K_c}(\mu_i;\rho'_i))$. We will omit the "$(\bmod M)$" notation for the sake of simplicity.

Now, assuming that there is an efficient (Diophantine) proof system for PK $(c = Com_{K_c}(\mu;\rho) \wedge \mu \in S)$, we are only left to prove the next result.

Theorem 3. *Let Com be the Damgård-Fujisaki integer commitment scheme and let $\Pi = (G, E, D)$ be a homomorphic public-key cryptosystem. Let t be the security parameter and let $T = t^{O(1)}$. Let $\rho_2 \in [0, 2^{B+t})$. The next protocol is a complete, honest-verifier SZK, specially sound proof system for* PK $(c_1 = E_{K_e}(\mu;\rho_1) \wedge c_2 = Com_{K_c}(\mu;\rho_2))$*, given that $\mu < T$:*

1. *Prover generates $m_1 \leftarrow_R [0, 2^t FT)$, $r_1 \leftarrow_R \mathcal{R}$, $r_2 \leftarrow_R [0, 2^{B+2t}F)$, sets $c_3 \leftarrow E_{K_e}(m_1; r_1)$, $c_4 \leftarrow Com_{K_c}(m_1; r_2)$ and sends (c_3, c_4) to verifier.*
2. *Verifier generates $e \leftarrow_R [0, F)$ and sends e to Prover.*
3. *Prover sets $m_2 \leftarrow m_1 + e\mu$, $r_3 \leftarrow r_1 + e\rho_1$ and $r_4 \leftarrow r_2 + e\rho_2$ and sends (m_2, r_3, r_4) to verifier.*
4. *Verifier checks that $c_3 = E_{K_e}(m_2; r_3) \cdot c_1^{-e}$ and $c_4 = Com_{K_c}(m_2; r_4) \cdot c_2^{-e}$.*

Proof of this theorem is given in Appendix D.

As previously, let C denote the ciphertext space of Π and C_{Com} the commitment space of Com. Noninteractive version of the presented proof system has length $5t+2\log_2 F+B+\log_2 T+R = 5{\cdot}80+2{\cdot}80+1024+1024+1024 = 3632$ bits or 454 bytes.

Interval membership proof for an encrypted integer. As a concrete application, let us describe a proof system for PK $(c = E_K(\mu;\rho) \wedge \mu \in [a, b])$:

1. Prover P generates $r_1 \leftarrow_R [0, 2^{B+2t})$, $c_1 \leftarrow Com_{K_c}(\mu; r_1)$ and sends c_1 to verifier.
2. P proves to V that PK $(c = E_K(\mu;\rho) \wedge c_1 = Com_{K_c}(\mu;\rho_1))$.
3. P proves to V that PK $(c_1 = Com_{K_c}(\mu;\rho) \wedge \mu \in [a, b])$.

Noninteractive version of this proof system is $128+454+2784+\frac{5}{8}\log_2(b-a) = 3366 + \frac{5}{8}\log_2(b-a)$ bytes long.

As noted before, it must be the case that $[a, b] \subseteq \mathcal{M}$ for this proof to work. In particular, we cannot take $S = [0, \infty)$. Therefore, to construct a proof that an encrypted number μ does not belong to $[a, b] \subset \mathcal{M} = \mathbb{Z}_M$, it does not suffice to prove in step 3 that $\mu \in (-\infty, a-1] \vee \mu \in [b+1, \infty)$: Instead, one must prove that $\mu \in [0, a-1] \vee \mu \in [b+1, M-1]$.

Membership proof in exponents for encrypted number In several cryptographic protocols like electronic voting [DJ01], one needs membership proofs in exponents: That is, proofs of type PK $(c = E_{K_e}(n^\mu; \rho) \wedge \mu \in [0, b])$ for some n. We will give an efficient SZK proof-of-knowledge of a small discrete logarithm of committed number in the special case when n is a prime. Since n is a prime then $\log_n m \in [0, b]$ iff $m \mid n^b$ and $\mu > 0$. (Thus, we use two sets, $S_1 = \{x_1 : x_1 \mid n^b\}$ and $S_2 = \{x_1 : x_1 \geq 0\}$.) Using this observation and ideas from the previous sections of the current paper, we have established that one needs to describe a SZK proof that $c = E_{K_e}(\mu; \rho) \wedge c_2 = Com_{K_c}(\mu; \rho_2) \wedge c_3 = Com_{K_c}(\mu_3; \rho_3) \wedge c_4 = Com_{K_c}(n^b; \rho_4) \wedge \mu\mu_3 = n^b \wedge \mu > 0$. This can be done as follows:

1. Prover lets $r_1 \leftarrow_R [0, 2^{B+2t})$, $c_2 \leftarrow Com_{K_c}(\mu; r_1)$, $r_2 \leftarrow_R [0, 2^{B+2t})$, $c_3 \leftarrow Com_{K_c}(n^b/\mu; r_2)$, $c_4 \leftarrow Com_{K_c}(\mu_4; 0)$. She sends (c_2, c_3) to verifier who computes $c_4 \leftarrow Com_{K_c}(\mu_4; 0)$.
2. P proves to V that PK $(c = E_{K_e}(\mu; \rho_1) \wedge c_2 = Com_{K_c}(\mu; \rho_2))$.
3. P proves to V that PK $\left(\left(\bigwedge_{i=2}^{4} c_i = E_{K_e}(\mu_i; \rho_i)\right) \wedge (\mu_4 = \mu_3\mu_2)\right)$. (I.e., that $\mu_2 \mid n^b$.)
4. P proves to V that PK $(c_2 = E_{K_e}(\mu; \rho) \wedge \mu \geq 0)$.

Noninteractive version of this proof system has length $256 + 454 + 612 + 1402 + \frac{5}{16}\log_2(b-a) = 2724 + \frac{5}{16}\log_2(b-a)$ bytes. As an interesting sidenote, one could further shorten this proof by using the result of Legendre that if $n^\mu \neq 4^s(8k+7)$ for some s, k (for example, if n is a power of two) then n^μ can be represented as a sum of three squares.

Application to E-voting. Until now, the best (perfect) zero-knowledge proof system for the same problem seems to be due Damgård and Jurik [DJ01], who used membership proof in exponents to prove vote correctness with n being the (maximum) number of voters. While their membership proof in exponents did not require n to be a prime, its length was $\approx \lceil \log_2 V \rceil \cdot (6C + M + 3t + 4R)$, where V being the number of candidates to vote for. In the Damgård-Jurik proof system, the length of the interaction will be greater than in ours as soon as $V \geq 8$. However, our constant-size proof system is possible only since we restricted n to be prime.

Further Work

Efficient Diophantine membership proofs can be given for many interesting sets $S \subset \mathbb{Z}$. We did certainly not mention all cryptographically relevant sets S that have such proofs. Full version of this paper will also give more insight into the complexity-theoretic aspects of p-bounded Diophantine sets.

References

[Bou00] Fabrice Boudot. Efficient Proofs that a Committed Number Lies in an Interval. In Bart Preneel, editor, *Advances on Cryptology — EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 431–444, Bruges, Belgium, 14–18 May 2000. Springer-Verlag.

[CDS94] Ronald Cramer, Ivan Damgård, and Berry Schoenmakers. Proofs of Partial Knowledge and Simplified Design of Witness Hiding Protocols. In Yvo G. Desmedt, editor, *Advances in Cryptology—CRYPTO '94*, volume 839 of *Lecture Notes in Computer Science*, pages 174–187, Santa Barbara, USA, 21–25 August 1994. Springer-Verlag.

[CFT98] Agnes Chan, Yair Frankel, and Yiannis Tsiounis. Easy Come - Easy Go Divisible Cash. In Kaisa Nyberg, editor, *Advances on Cryptology — EUROCRYPT '98*, volume 1403 of *Lecture Notes in Computer Science*, pages 561–575, Helsinki, Finland, June 1998. Springer-Verlag.

[CM99] Jan Camenisch and Markus Michels. Proving in Zero-Knowledge that a Number Is the Product of Two Safe Primes. In Jacques Stern, editor, *Advances on Cryptology — EUROCRYPT '99*, volume 1592 of *Lecture Notes in Computer Science*, pages 107–122, Prague, Czech Republic, 2–6 May 1999. Springer-Verlag.

[Dav73] Martin Davis. Hilbert's Tenth Problem is Unsolvable. *American Mathematical Monthly*, 80(3):233–269, March 1973.

[DF01] Ivan Damgård and Eiichiro Fujisaki. An Integer Commitment Scheme Based on Groups with Hidden Order. Technical Report 064, IACR, 13 August 2001. Available at `http://eprint.iacr.org/2001/064/`.

[DJ01] Ivan Damgård and Mads Jurik. A Generalisation, a Simplification and Some Applications of Paillier's Probabilistic Public-Key System. In Kwangjo Kim, editor, *Public Key Cryptography '2001*, volume 1992 of *Lecture Notes in Computer Science*, pages 119–136, Cheju Island, Korea, 13–15 February 2001. Springer-Verlag.

[FO97] Eiichiro Fujisaki and Tatsuaki Okamoto. Statistical Zero Knowledge Protocols to Prove Modular Polynomial Relations. In Burton S. Kaliski, editor, *Advances on Cryptology — CRYPTO '97*, volume 1294 of *Lecture Notes in Computer Science*, pages 16–30, Santa Barbara, USA, 17–21 August 1997. Springer-Verlag.

[FS86] Amos Fiat and Adi Shamir. How to Prove Yourself: Practical Solutions to Identification and Signature Problems. In Andrew M. Odlyzko, editor, *Advances in Cryptology—CRYPTO '86*, volume 263 of *Lecture Notes in Computer Science*, pages 186–194, Santa Barbara, California, USA, 11–15 August 1986. Springer-Verlag, 1987.

[JSWW76] James P. Jones, Daihachiro Sato, Hideo Wada, and Douglas Wiens. Diophantine Representation of the Set of Prime Numbers. *American Mathematical Monthly*, 83(6):449–464, June–July 1976.

[Mat70] Yuri Matiyasevich. Enumerable Sets are Diophantine. *Soviet Math., Doklady*, 11:354–358, 1970. English translation.

[MS97] Markus Michels and Markus Stadler. Efficient Convertible Undeniable Signature Schemes. In *Proc. 4th Workshop on Selected Areas in Cryptography (SAC'97)*, pages 231–244, Ottawa, Canada, 1997.

[Pai99] Pascal Paillier. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In Jacques Stern, editor, *Advances on Cryptology — EUROCRYPT '99*, volume 1592 of *Lecture Notes in Computer Science*, pages 223–238, Prague, Czech Republic, 2–6 May 1999. Springer-Verlag.

[RS86] Michael O. Rabin and Jeffrey O. Shallit. Randomized Algorithms in Number Theory. *Communications in Pure and Applied Mathematics*, 39:239–256, 1986.

[Vad00] Salil P. Vadhan. On Transformation of Interactive Proofs that Preserve the Prover's Complexity. In *Proceedings of the Thirty-Second Annual ACM Symposium on the Theory of Computing*, pages 200–207, Portland, Oregon, USA, 21–23 May 2000. ACM Press.

A Proof of Theorem 2

This appendix provides the proof of Theorem 2.

Proof. Completeness: $g^{m_{2i}} h^{r_{4i}} c_{1i}^{-e} = g^{m_{1i}+e\mu_i} h^{r_{2i}+e\rho_i} g^{-e\mu_i} h^{-e\rho_i} = g^{m_{1i}} h^{r_{2i}} = c_{2i}$ and $\prod_i c_{1i}^{m_{2i}} \cdot h^{r_5} c^{-e} = \prod_i c_{1i}^{m_{1i}} \cdot \prod_i (g^{\mu_i} h^{\rho_i})^{e\mu_i} \cdot h^{r_3+e\sum_i (1-\mu_i)\rho_i} g^{-e\sum_i \mu_i^2} h^{-e\rho} = \prod_i c_{1i}^{m_{1i}} \cdot h^{r_3} = c_3$.

Honest-verifier SZK. The simulator acts as follows. For $i \in [1,4]$, generate $c_{1i} \leftarrow_R \mathcal{C}_{Com}$, $m_{2i} \leftarrow_R [0, 2^F T)$, $r_{4i} \leftarrow_R [0, 2^{B+2t} F)$. Generate $e \leftarrow_R [0, F)$, $r_5 \leftarrow_R [0, 2^{B+2t} FT)$. For $i \in [1,4]$, let $c_{2i} \leftarrow g^{m_{2i}} h^{r_{4i}} c_{1i}^{-e}$. Let $c_3 \leftarrow \prod_i c_{1i}^{m_{2i}} \cdot h^{r_5} c^{-e}$. The resulting view $((c_{1i}, c_{2i})_{i=1}^4, c_3; e; (m_{2i}, r_{4i})_{i=1}^4, r_5)$ is accepting and has distribution, statistically close to the distribution of views in a real execution.

Special soundness (from two accepting views, $((c_{1i}, c_{2i})_i, c_3; e; (m_{2i}, r_{4i})_i, r_5)$ and $((c_{1i}, c_{2i})_i, c_3; e'; (m'_{2i}, r'_{4i})_i, r'_5)$ with $e \neq e'$, one can efficiently find a pair $((\mu_i)_i, \rho)$, such that $c = Com_K(\sum \mu_i^2; \rho)$): Given such views, $g^{m_{2i}-m'_{2i}} h^{r_{4i}-r'_{4i}} = c_{1i}^{e-e'}$, for $i \in [1,4]$, and $\prod_i c_{1i}^{(m_{2i}-m'_{2i})} \cdot h^{r_5-r'_5} = c^{e-e'}$. We say that we have a bad case, if either $(e-e') \not| (m_{2i} - m'_{2i})$ or $(e-e') \not| (r_{4i} - r'_{4i})$ for some $i \in [1,4]$ or $(e-e') \not| (r_5 - r'_5)$. As in [DF01], we can argue that the bad case appears with a negligible probability if the group assumptions hold. Otherwise (when we do not have the bad case), let $\mu_i \leftarrow (m_{2i} - m'_{2i})/(e-e')$ and $\rho_i \leftarrow (r_{4i} - r'_{4i})/(e-e')$; then c_{1i} can be opened as $c_{1i} = g^{\mu_i} h^{\rho_i}$, for $i \in [1,4]$, and c can be opened as $c = \prod_i c_{1i}^{\mu_i} \cdot h^{(r_5-r'_5)/(e-e')} = ((g^{\sum_i \mu_i} h^{\sum_i \rho_i})^{\mu_i}) h^{(r_5-r'_5)/(e-e')} = g^{\sum_i \mu_i^2} h^{\sum_i \mu_i \rho_i + (r_5 - r'_5)/(e-e')}$. $\square$

B Rabin-Shallit Algorithm

For completeness, we will next give a short overview of the Rabin-Shallit algorithm [RS86] that takes $O((\log \mu)^4)$ bit-operations:

1. Write μ in the form $\mu = 2^s(2k+1)$, where $s, k \geq 0$.
2. If $s = 1$, then
   (a) Choose random $\mu_1, \mu_2 \leq \sqrt{\mu}$, with exactly one of μ_1, μ_2 even. Let $p \leftarrow \mu - \mu_1^2 - \mu_2^2$. Note that $p \equiv 1 \pmod 4$.
   (b) Hoping that p is prime, try to express $p = \mu_3^2 + \mu_4^2$ as follows: First, find a solution u to the equation $u^2 \equiv -1 \pmod p$. (This can be done in various efficient ways; for details see [RS86].) Now compute $\gcd(u + \mathrm{i}, p) = \mu_3 + \mu_4\mathrm{i}$ over the Gaussian integers. Again, this can be done efficiently. Check to see that $p = \mu_3^2 + \mu_4^2$. If not, p was not prime, so go back to step 2a.
   (c) Return $\mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2$ as a representation.
3. If s is odd but not 1, find a representation for $2(2k+1)$ and then multiply each term by the square t^2, where $t = 2^{(s-1)/2}$.
4. If s is even, find a representation $\mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2$ for $2(2k+1)$ by step 2. Then convert this to a representation for $(2k+1)$ as follows: Group μ_1, μ_2, μ_3, μ_4 so that $\mu_1 \equiv \mu_2 \pmod 2$ and $\mu_3 \equiv \mu_4 \pmod 2$. Then $(2k+1) = (\frac{1}{2}(\mu_1+\mu_2))^2 + (\frac{1}{2}(\mu_1-\mu_2))^2 + (\frac{1}{2}(\mu_3+\mu_4))^2 + (\frac{1}{2}(\mu_3-\mu_4))^2$. Now multiply by t^2, where $t = 2^{s/2}$.

C Longer Example: Proof Systems for Exponential Relationship

We will next describe a few proof systems for the exponential relationship, i.e., for $\mathrm{PK}\left(\left(\bigwedge_{i=1}^3 c_1 = Com_K(\mu_1; \rho_1)\right) \wedge \mu_1 = \mu_2^{\mu_3}\right)$. Since the set $\{(x_1, x_2, x_3) : x_1 = x_2^{x_3}\}$ is recursively enumerable, it is Diophantine. Matiyasevich was the first to describe an explicit representing polynomial for $\mathcal{S}$. Next, we will give a description of such a $R_{\mathcal{S}}$ due to [Dav73,JSWW76].

All nonnegative integral solutions $(x_n(a), y_n(a))$ of the Pell equation $x^2 - (a^2-1)y^2 = 1$ can be derived from the next recurrent identities. First, let $(x_0(a), y_0(a)) = (1, 0)$ and $(x_1(a), y_1(a)) = (a, 1)$. Second, let $x_{n+1}(a) = 2ax_n(a) - x_{n-1}(a)$ and $y_{n+1}(a) = 2ay_n(a) - y_{n-1}(a)$. Equivalently, $x_n(a) = \sum_{i=0,2|i}^n \binom{n}{i} a^{n-i}(a^2-1)^{i/2}$ and $y_n(a) = \sum_{i=1,2\nmid i}^n \binom{n}{i} a^{n-i}(a^2-1)^{(i-1)/2}$. Clearly, $a^n \leq x_n(a) \leq (2a)^n$.

Theorem 4 ([Dav73,JSWW76]). $m = n^k$ *iff the next* 9 *equations have a positive integral solution in remaining* 14 *arguments* $(a, c, d, e, f, g, h, l, r, u, w, x, y, z)$*:* $x^2 - (a^2-1)y^2 = 1$, $u^2 - (a^2-1)r^2y^4 = 1$,

$(x+cu)^2 - ((a+u^2(u^2-a))^2-1)(k+4(d-1)y)^2 = 1$, $y = k+e-1$, $(x-y(a-n)-m)^2 = (f-1)^2(2an-n^2-1)^2$, $m+g = 2an-n^2-1$, $w = n+h$, $w = k+l$, $a^2-(w^2-1)(w-2w+1)z^2 = 1$.

From one side, this representation of $\mathcal{S}$ seems to give a very efficient unbounded Diophantine proof system. Only 25 essentially different multiplications need to be performed. Computing witnesses is also "easy". Namely, if $m = n^k$ then $w \leftarrow \max(k,n)+1$, $a \leftarrow x_{w-1}(w)$, then $a > 1$. Set $h \leftarrow w-n$, $l \leftarrow w-k$, $x \leftarrow x_k(a)$, $y \leftarrow y_k(a)$, $y \leftarrow y_k(a)$, $u \leftarrow x_{2ky_k(a)}(a)$, $r \leftarrow y_{2ky_k(a)}(a)/y^2$, $c \leftarrow (x_k(a+u^2(u^2-a))-x)/u$, $d \leftarrow (y_k(a+u^2(u^2-a))-k)/(4y)+1$, $e \leftarrow y-k+1$. Finally let f be such that $f = 1 \pm (x-y(a-n)-m)/(2an-n^2-1)$. Therefore it might seem that to prove the exponential relationship by using Theorem 4, one only one need to use Protocol 1 25 times (and then to prove that some witnesses are positive).

However, there is a serious catch. The total amount of computation is superlinear in $\log_2 T$, where T is an a priori upper limit on the size of any multiplicands and any variables. And in this case, the largest multiplicand is $x+cu = x_k(a+u^2(u^2-a)) \geq (a+u^2(u^2-a))^k \geq u^{4k} = (x_{2ky_k(a)}(a))^{4k} \geq a^{8k^2y_k(a)} \geq a^{8k^2a^k} = (x_{w-1}(w))^{8k^2(x_{w-1}(w))^k} \geq w^{(w-1)8k^2w^{(w-1)k}}$, which is is definitely way and beyond the current computational power when $k \gg 50$.

On the other hand, $\mathcal{S} = (S_t)$ is clearly p-bounded Diophantine. Let $S_t = \{(x_1,x_2,x_3) : x_3 = {x_1}^{x_2} \wedge x_2 < 2^t\}$. Here, following [CM99,DJ01] and many other publications, one can represent x_2 as a binary number, $x_2 = \sum_{i=0}^{t} k_i 2^i$. After that, one can prove that $x_3 = \prod_i n_i$, where either $n_i = 1$ or $n_i = x_1^{2^i}$. In the previous proof system we had a constant number of multiplications, but inputs of superpolynomial size. In the current case we have $\Theta(\log_2 k)$ multiplications but the multiplicands are never bigger than x_3.

Finally, note that given a p-bounded (resp., unbounded) Diophantine proof system for exponential relationship, one can prove that $x_1 = \binom{x_2}{x_3}$ by showing that $y_1 = 2^{x_2+1}$, $(y_1+1)^{x_2} = y_2{y_1}^{x_3+1} + x_1{y_1}^{x_3} + y_3$, $y_3 = {y_1}^{x_3}$ and $x_1 < y_1$ for some nonnegative y_i. In this case, all intermediate results have length polynomial in $|x_1|$ and hence the proof system for binomial relationship would also be p-bounded (resp., unbounded).

D Proof for Theorem 3

Proof. Completeness. If prover is honest then $E_{K_e}(m_2;r_3)\cdot c_1^{-e} = E_{K_e}(m_2 - e\mu; r_3 - e\rho_1) = E_{K_e}(m_1;r_1) = c_3$ and $Com_{K_c}(m_2;r_4)\cdot c_2^{-e} = Com_{K_c}(m_2 - e\mu; r_4 - e\rho_2) = E_{K_e}(m_1;r_2) = c_4$.

Honest-verifier SZK. Simulator generates a random tuple $(e, m_2, r_3, r_4) \leftarrow [0,F) \times [0, 2^tFT) \times \mathcal{R} \times [0, 2^{B+2t})$ and sets $c_3 \leftarrow E_{K_e}(m_2;r_3)\cdot c_1^{-e}$, $c_4 \leftarrow$

$Com_{K_c}(m_2; r_4) \cdot c_2^{-e}$. Clearly, this view is an accepted view. Moreover, it has distribution that is statistically close to the distribution of real view.

Special soundness. Let the next two views be accepting: $(c_3, c_4; e; m_2, r_3, r_4)$ and $(c_3, c_4; e'; m_2', r_3', r_4')$ with $e \neq e'$. We know from [DF01] that then with an overwhelming probability $(e - e') \mid (m_2 - m_2')$. Therefore, $c_2 = Com_{K_c}(\mu; \rho_2)$ with $\mu = \frac{m_2 - m_2'}{e - e'}$. Similarly, $c_1 = E_{K_e}(\mu'; \rho_1)$, where $\mu' = \frac{m_2 - m_2'}{e - e'} \pmod{M}$. Hence, $\mu' = \mu \pmod{M}$. $\square$