Finding a Small Root of a Bivariate Integer Equation; Factoring with High Bits Known

Don Coppersmith

IBM Research
T.J. Watson Research Center
Yorktown Heights, NY 10598, USA

Abstract. We present a method to solve integer polynomial equations in two variables, provided that the solution is suitably bounded. As an application, we show how to find the factors of $N = PQ$ if we are given the high order $((1/4)\log_2 N)$ bits of P. This compares with Rivest and Shamir's requirement of $((1/3)\log_2 N)$ bits.

1 Introduction

We present a method to solve a polynomial equation $p(x, y) = 0$ over $\mathbb{Z}$, provided that the solution is suitably bounded: $|x| < X$ and $|y| < Y$, with X, Y depending on the coefficients and degree of p.

Our algorithm uses lattice basis methods [2]. It is similar in spirit to [1], which solved equations in one variable in ($\mathbb{Z}$ mod N), but the present algorithm requires a different analysis.

We require bounds X and Y on the absolute values of x and y in our solution. Suppose $p(x, y)$ has degree δ in each variable, and $p(x, y) = \sum_{ij} p_{ij} x^i y^j$. Define $D = \max_{ij} |p_{ij}| X^i Y^j$ as the largest possible term in $p(x, y)$ in the region of interest. Then we will find a bounded solution (x, y) (if it exists) provided that

$$XY < D^{2/(3\delta)} .$$

For fixed degree δ, the algorithm runs in time polynomial in $(\log D)$.

Similar methods can be applied to the multivariate case but are not assured of success; the proof breaks down at a critical point.

Our immediate application, and the framework in which the algorithm is described, is the problem of factoring an integer when we know the high order bits of its factors. If we know $N = PQ$ and we know the high order $(\frac{1}{4}\log_2 N)$ bits of P, then by solving the equation $(P_0 + x)(Q_0 + y) - N = 0$ over a suitable range of x and y we can find the factorization of N. By comparison, Rivest and Shamir [5] need about $(\frac{1}{3}\log_2 N)$ bits of P. This has applications to some RSA-based cryptographic schemes; see for example [7].

We give here a sketch of our algorithm. Define integer variables r_{gh} representing $x^g y^h$. Form the lattice of those values of $\{r_{gh}\}$ satisfying several polynomial relations $q_{ij}(x, y) = x^i y^j p(x, y) = 0$ under this interpretation. Claim that the lattice element s corresponding to our desired solution is relatively short (less

U. Maurer (Ed.): Advances in Cryptology - EUROCRYPT '96, LNCS 1070, pp. 178-189, 1996.

than the nth root of the determinant of a certain matrix). The expression of $\mathbf{s}$ in terms of a reduced basis of our lattice cannot involve the longest basis element (because $\mathbf{s}$ is short), so $\mathbf{s}$ is confined to the hyperplane spanned by the other basis elements. This gives a linear equation on $\{r_{gh} = x^g y^h\}$, which we interpret as a polynomial equation on x and y. We combine this with $p(x, y) = 0$ and solve for x_0 and y_0.

The remainder of the paper is organized as follows. In Section 2 the algorithm is developed, concentrating on the concrete case where $p(x, y)$ has degree 1 in each variable. Section 3 gives a brief discussion of linear lattice methods as applied to the nonlinear problem of solving polynomials. In Section 4 we extend the algorithm to other bivariate polynomials, and discuss the dependence on the size parameter D and degree δ of the polynomial p. We comment on possible extensions to three or more variables in Section 5. In Sections 6 and 7 we compare the current algorithm with previous ones. An application to Vanstone and Zuccherato's RSA variant is given in Section 8. The appendix proves a technical result on Toeplitz matrices.

2 Factoring with high order bits known

We present the algorithm in terms of the problem of factoring an integer when we know the high-order bits of one of the factors.

Suppose we know $N = PQ$, and suppose that for some $\epsilon > 0$ we know the high order $(\frac{1}{4} + \epsilon)(\log_2 N)$ bits of P. (We will dispense with the ϵ later.) By division we know the high order bits of Q as well.

$$\begin{aligned} P &= P_0 + x_0 \\ Q &= Q_0 + y_0 \\ |x_0| &< X = P_0/N^{(1/4)+\epsilon} \\ |y_0| &< Y = Q_0/N^{(1/4)+\epsilon} \\ p(x, y) &= (P_0 + x)(Q_0 + y) - N \\ &= (P_0 Q_0 - N) + Q_0 x + P_0 y + xy \\ p(x_0, y_0) &= PQ - N = 0 \end{aligned}$$

Here P_0 and Q_0 are known, while x_0 and y_0 are unknown, and x and y are dummy variables. $p(x, y)$ is an irreducible polynomial with integer coefficients, and its coefficients share no common factor.

We will relate the bounds X and Y to the quantity

$$D = \max\{|P_0 Q_0 - N|, Q_0 X, P_0 Y, XY\} \ .$$

This is the largest possible size of an individual term of $p(x, y)$ with bounded x and y. For our methods to work, we will require $(XY)^{3/2} < D$. In the case of a more general polynomial $p(x, y) = \sum_{ij} p_{ij} x^i y^j$, of degree δ separately in each variable, we would define

$$D = \max_{ij}\{|p_{ij}| X^i Y^j\}$$

and demand $(XY)^{3\delta/2} < D$. (The definitions of X, Y, D appear circular, but it's all right; the condition is equivalent to existence of indices $i, j \leq \delta$ such that $X^{(3\delta/2)-i}Y^{(3\delta/2)-j} < |p_{ij}|$, and the exponents $(3\delta/2)-i$ and $(3\delta/2)-j$ are strictly positive.)

We are trying to find a bounded pair of integers (x_0, y_0) solving $p(x_0, y_0) = 0$. We begin by selecting an integer $k > 1/(4\epsilon)$. For each pair of integers (i, j) with $0 \leq i < k$ and $0 \leq j < k$, form the polynomial $q_{ij}(x, y) = x^i y^j p(x, y)$. Obviously $q_{ij}(x_0, y_0) = 0$.

Form a matrix M_1 with $(k+1)^2$ rows, indexed by $\gamma(g, h) = (k+1)g + h$ with $0 \leq g, h < k+1$. M_1 has $(k+1)^2 + k^2$ columns, the left-hand columns indexed by $\gamma(g, h)$, and the right-hand columns indexed by $\beta(i, j) = (k+1)^2 + ki + j$ with $0 \leq i, j < k$. The left-hand block is a diagonal matrix whose $(\gamma(g, h), \gamma(g, h))$ entry is given by $X^{-g}Y^{-h}$. The $(\gamma(g, h), \beta(i, j))$ entry of the right-hand block is the coefficient of $x^g y^h$ in the polynomial $q_{ij}(x, y)$.

An explanation of M_1 is in order. The $\gamma(g, h)$ row corresponds to an integer unknown r_{gh} which represents $x_0^g y_0^h$. In the left-hand block, the diagonal entry $X^{-g}Y^{-h}$ will be used to bound $|r_{gh}|$ by approximately $X^g Y^h$. We will be concentrating on the sublattice in which the right-hand columns are zero; a zero in column $\beta(i, j)$ will correspond to the condition $q_{ij}(x_0, y_0) = 0$.

Perform elementary row operations on M_1 to produce a matrix M_2 whose right-hand block has the $k^2 \times k^2$ identity matrix on the bottom and the $(2k+1) \times k^2$ zero matrix on the top. We can do this because the coefficient of xy in p is 1, so that the right-hand block of M_1 contains an upper triangular matrix with 1 on the diagonal. (For a more general polynomial $p(x, y)$, we require that the coefficients of p share no nontrivial common factor; in other words, $p(x, y)$ does not factor as $k \times u(x, y)$ over $\mathbb{Z}$.)

The lattice formed by these top $2k+1$ rows of M_2 is the sublattice of the original lattice gotten by forcing to 0 all the right-hand columns. Call it M_3.

Consider the $(k+1)^2$-long row vector $\mathbf{r}$ whose $\gamma(g, h)$ entry is $x_0^g y_0^h$. The row vector $\mathbf{s}$ of length $(k+1)^2 + k^2$ given by $\mathbf{s} = \mathbf{r}M_1$ satisfies

$$\begin{aligned} s_{\gamma(g,h)} &= (x_0/X)^g (y_0/Y)^h \\ |s_{\gamma(g,h)}| &\leq 1 \\ s_{\beta(i,j)} &= q_{ij}(x_0, y_0) = 0 \\ |\mathbf{s}| &< k+1 \end{aligned}$$

Because its right-hand side is 0, $\mathbf{s}$ is one of the vectors in the lattice spanned by the rows of M_3. We will show that it is a "relatively short" vector in the lattice, which will enable us to confine it to a hyperplane, thus producing a linear equation relating its coefficients. This will translate directly to a polynomial equation on x_0 and y_0: $u(x_0, y_0) = 0$, where $u(x, y)$ is not a multiple of $p(x, y)$. We can then take the resultant of $u(x, y)$ with $p(x, y)$ to find a single polynomial equation $v(x) = 0$ satisfied by x_0, and solve this equation over $\mathbb{Z}$ to find x_0.

We proceed to estimate the sizes of the vectors in row lattice spanned by M_3, by estimating the determinant of a square submatrix of M_3. Define the diagonal matrix W of dimension $(k+1)^2 \times (k+1)^2$, with $(\gamma(g, h), \gamma(g, h))$ entry given

by X^gY^h. In the matrix WM_1, the left-hand block is the identity. In the $\beta(i,j)$ column of the right-hand side of WM_1, the largest element has absolute value X^iY^jD. That is, the element at $(\gamma(g,h),\beta(i,j))$ is

$$X^gY^hp_{g-i,h-j} = X^{i+a}Y^{j+b}p_{ab} = X^iY^j(X^aY^bp_{ab})$$

where $a = g - i$ and $b = h - j$. Further, the right-hand columns are "nearly orthogonal", because they are part of a Toeplitz array. (A formal statement and proof appear in the appendix.) Associated with this near orthogonality, there is a specific set of k^2 columns in the left-hand block. Whenever we delete these k^2 columns from a rectangular matrix M, we denote the resulting square matrix as $\hat{M}$. Deleting these columns from WM_1 leaves a matrix $W\hat{M}_1$ whose determinant satisfies

$$\begin{aligned} |\det(W\hat{M}_1)| &= \Omega(\textstyle\prod_{ij}(X^iY^jD)) \\ &= \Omega((XY)^{k^2(k-1)/2}D^{k^2}) \ , \end{aligned}$$

the constant implicit in Ω depending on k and the pattern of nonzero coefficients in the polynomial $p(x,y)$, as shown in the appendix. For the polynomial in our example, the constant is 1 (see appendix) so we drop the "Ω":

$$|\det(W\hat{M}_1)| = (XY)^{k^2(k-1)/2}D^{k^2} \ .$$

Since

$$\det(W) = \prod_{gh}(X^gY^h) = (XY)^{(k+1)^2k/2} \ ,$$

we can calculate

$$\begin{aligned} |\det(M_1)| &= |\det(WM_1)|/\det(W) \geq D^{k^2}(XY)^{\{k^2(k-1)/2\}-\{(k+1)^2k/2\}} \\ &= D^{k^2}(XY)^{-k(3k+1)/2} = \left(D^k(XY)^{-(3k+1)/2}\right)^k \ . \end{aligned}$$

Remark. Here is where we will use the fact that $(XY)^{3/2} < D$; in this example, it happens to be a consequence of the knowledge of $(\frac{1}{4}+\epsilon)(\log_2 N)$ bits.

From

$$\begin{aligned} D \geq |P_0Y| &= P_0Q_0/N^{1/4+\epsilon} \\ XY = (P_0/N^{1/4+\epsilon})(Q_0/N^{1/4+\epsilon}) &= P_0Q_0/N^{1/2+2\epsilon} \end{aligned}$$

we see

$$D^k(XY)^{-(3k+1)/2} \geq (P_0Q_0)^{-(k+1)/2}N^{(k/2)+(1/4)+(2k+1)\epsilon} \ .$$

Because $P_0Q_0 = N(1+o(N^{-1/4}))$ we can replace P_0Q_0 by N and incur negligible error:

$$D^k(XY)^{-(3k+1)/2} \geq N^{(-1/4)+(2k+1)\epsilon}(1+o(kN^{-1/4})) \ .$$

Recall $k > 1/(4\epsilon)$. Then

$$D^k(XY)^{-(3k+1)/2} > N^{(+1/4)+\epsilon}(1+o(kN^{-1/4})) > N^{1/4} \ ,$$

$$|\det(M_1)| > N^{k/4} \gg 1 \ .$$

The same estimate applies to $|\det(\hat{M}_2)|$ because M_2 was gotten from M_1 by elementary row operations. Also, because the lower right $k^2 \times k^2$ submatrix of $\hat{M}_2$ is the identity and the upper right submatrix is zero, the same estimate applies to the determinant of the upper left $(2k+1) \times (2k+1)$ submatrix of $\hat{M}_2$, namely the left-hand $(2k+1) \times (2k+1)$ square submatrix of $\hat{M}_3$. For the following discussion, we call that square submatrix L and its dimension $n = 2k+1$. So $|\det(L)| > N^{k/4}$.

We apply lattice basis reduction to the row basis of L, as prescribed in [2], to produce a reduced basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$. From the discussion in [2], the last element $\mathbf{b}_n$ of this reduced basis satisfies

$$|\mathbf{b}_n^*| \geq |\det(L)|^{1/n} 2^{-(n-1)/4} \ ,$$

where $\mathbf{b}_n^*$ denotes the component of $\mathbf{b}_n$ orthogonal to the subspace spanned by the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}$. As long as

$$k < \frac{1}{4} \log_2 N - 2 \log_2 \log_2 N - \Omega(1) \ ,$$

we will have that $|\mathbf{b}_n^*| > |\mathbf{s}|$:

$$\begin{aligned} |\mathbf{b}_n^*| &\geq |\det(L)|^{1/n} 2^{-(n-1)/4} > N^{(k/4)(1/(2k+1))} 2^{-2k/4} \\ &\approx N^{(1/8)-(1/(16k))} 2^{-k/2} > k+1 > |\mathbf{s}| \ . \end{aligned}$$

Assume this inequality holds. For any row vector $\mathbf{t}$ in the lattice spanned by L, if $\mathbf{t}$ is not in the lattice spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}$, then its expression as an integer combination of the $\mathbf{b}_i$ involves $\mathbf{b}_n$ nontrivially. Thus we have $|\mathbf{t}| \geq |\mathbf{b}_n^*| > k+1 > |\mathbf{s}|$. Looked at the other way, for any $\mathbf{t}$ in the lattice spanned by L, if $|\mathbf{t}| \leq |\mathbf{s}|$, then $\mathbf{t}$ is in the lattice spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}$.

Consider $\mathbf{s}$ itself. The $n = 2k+1$ entries of $\mathbf{s}$ corresponding to those columns in the left-hand side that remain when we transform M to $\hat{M}$, form a row vector $\hat{\mathbf{s}}$ in the lattice spanned by L (since the right-hand elements of $\mathbf{s}$ are 0). Also, $|\hat{\mathbf{s}}| \leq |\mathbf{s}|$. Thus $\hat{\mathbf{s}}$ is in the lattice spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}$.

Membership in this subspace gives us a linear relation on the coefficients $r_{gh} = x_0^g y_0^h$ expressing $\mathbf{s}$ as a linear combination of the rows of M_1. This relation is linearly independent of the k^2 relations given by the polynomials $q_{ij}(x, y)$ which determined that $\mathbf{s}$ had right-hand side 0 and thus was in the lattice of M_3 to start with. So this relation translates to a polynomial relation $u(x_0, y_0) = 0$ where $u(x, y)$ is not a polynomial multiple of $p(x, y)$.

Take the resultant of $p(x, y)$ and $u(x, y)$ with respect to y. Because $p(x, y)$ is irreducible and $u(x, y)$ is not a polynomial multiple of $p(x, y)$, we have that $\text{Resultant}_y(p(x, y), u(x, y)) = v(x)$ is a nontrivial integer polynomial $v(x)$ in one variable x satisfied by x_0: $v(x_0) = 0$. Since $u(x, y)$ has degree at most k in each variable and $p(x, y)$ has degree 1 in each variable, $v(x)$ has degree at most $2k$. Solve $v(x) = 0$ over $\mathbb{Z}$ to find a small number of candidates for x_0, namely those integer solutions satisfying the bound $|x_0| < X$.

Each candidate value of x_0 can be substituted into p to get an equation $p(x_0, y) = 0$ which we can solve for y over $\mathbb{Z}$, and select those integer solutions satisfying the bound $|y_0| < Y$.

Thus we have proven:

Theorem 1. *If we know an integer $N = PQ$ and we know the high order $(1/4 + \epsilon)(\log_2 N)$ bits of P, with $\epsilon > 2/(\log_2 N)$, then in time polynomial in $\log N$ and $1/\epsilon$ we can discover P and Q.*

Proof. The condition on ϵ insures that k can be chosen to satisfy

$$\frac{1}{4\epsilon} < k < \frac{\log_2 N}{4} - O(\log \log N) \ .$$

The complexity is due to invocation of lattice basis reduction on a matrix of size $2k + 1 \approx 1/(2\epsilon)$, whose elements are integers with bit length bounded by a polynomial in $\log N$. (We have to transform our rational matrices to integer matrices by multiplying by some integer.) □

Corollary 2. *If we know an integer $N = PQ$ and we know the high order $(1/4)(\log_2 N)$ bits of P, then in time polynomial in $\log N$ we can discover P and Q.*

Proof. Set $\epsilon = 4/\log_2 N$ and do exhaustive search on $O(1)$ unknown high order bits of x_0 (or middle bits of P). □

3 Discussion on lattice methods

The lattice basis reduction method is inherently linear. If we want to relate several unknowns by a polynomial equation $p(x, y) = \sum_{gh} p_{gh} x^g y^h = 0$, one natural approach is to replace each monomial $x^g y^h$ by a new independent variable r_{gh}, and let p become a linear relation among several independent bounded variables: $\sum_{gh} p_{gh} r_{gh} = 0$, $|r_{gh}| \leq X^g Y^h = R_{gh}$. In order to get results, we would need the desired vector to be among the shortest of the lattice; we would need its length to be smaller than the root of the determinant of the appropriate square matrix. This requirement translates to (approximately) $\prod R_{gh} \leq D$, and this would imply severe restrictions on X and Y, namely $(XY)^{(\delta+1)^2\delta/2} \leq D$, where $p(x, y)$ has degree δ in each variable. This is essentially the technique used by [6] in the modular setting.

In the present paper we have extended this approach by using several polynomials q_{ij} but reusing the same independent bounded variables r_{gh}. We are able to amortize the cost of the several variables over several equations. This accounts for the success of the present method in increasing the feasible sizes of X and Y. Specifically, each unknown r_{gh} contributes a factor of $X^{-g}Y^{-h}$ to $\det(M_1)$, while each equation $q_{ij}(x, y) = 0$ contributes a factor of $X^i Y^j D$. In order for our techniques to work, we require $\det(M_1) \gg 1$, which yields a bound on X and Y in terms of D. Because we have several polynomial equations, each contributing positively to the determinant, this bound is relatively mild; we can tolerate larger ranges X, Y on our variables (in terms of D) than with other methods, namely $(XY)^{3\delta/2} < D$, as we will see in the next section.

4 Other bivariate polynomials

Similar techniques can be applied successfully to other polynomial equations besides the given $p(x, y) = 0$. They are not guaranteed for polynomials of more than two variables; see Section 5.

Even in the case of two variables, the present technique is sensitive to the form of the polynomial $p(x, y)$. For an arbitrary quadratic polynomial we could not tolerate ranges X, Y for the unknowns x, y quite as large as we did here. In the case that we used for our example, although $p(x, y)$ has degree two, it does not have terms involving x^2 or y^2, only xy.

We sketch here how the bounds X and Y depend on D and the form of p. When we estimated $|\det(M_1)|$ by dividing the estimate of $|\det(WM_1)|$ by $\det(W)$, there was considerable cancellation in the powers of X and Y. The term $|\det(WM_1)|$ had a factor X^iY^jD for each pair (i, j) with $0 \le i, j < k$; these pairs represented the monomials x^iy^j by which $p(x, y)$ was multiplied. The term $\det(W)$ had a factor X^gY^h for each pair (g, h) with $0 \le g, h < k + 1$; these pairs represented the monomials x^gy^h appearing in these products $x^iy^jp(x, y)$. The range on g exceeds the range on i by 1 because p has degree 1 in x. If p had an x^2 term, we would have needed to enlarge the range of g. The powers of X and Y appearing in the ratio $|\det(WM_1)|/\det(W)$ arise from the pairs (g, h) outside the range of (i, j), namely

$$\{(g, h)|g = k, 0 \le h < k\} \cup \{(g, h)|h = k, 0 \le g \le k\} \ .$$

The inclusion of an x^2 term would have enlarged that region by another layer. This would have led directly to a stricter requirement on the sizes of X and Y.

If our polynomial $p(x, y)$ has degree δ in x and τ in y, then we can tolerate ranges X and Y satisfying

$$X^{\delta+(\tau/2)}Y^{\tau+(\delta/2)} < D$$

by using polynomials $q_{ij}(x, y)$ in the range $0 \le i, j < k$. More generally, for any positive value of the parameter α we can tolerate X and Y with

$$X^{\delta+(\alpha\tau/2)}Y^{\tau+(\delta/(2\alpha))} < D$$

by allowing $0 \le i < k\alpha$ and $0 \le j < k$.

If $p(x, y)$ has total degree δ then we can tolerate about

$$D > (XY)^{\delta}$$

by allowing pairs (i, j) with $i \ge 0$, $j \ge 0$, and $i + j < k$. This is better than the previous approach if p is a general bivariate polynomial of total degree δ, but worse if (like the current example) it is really of degree $\delta/2$ separately in each variable.

We summarize these results:

Theorem 3. *Let $p(x,y) = \sum_{ij} p_{ij} x^i y^j$ be a bivariate polynomial over $\mathbb{Z}$ of degree δ in x and τ in y. Assume p is irreducible over $\mathbb{Z}$. Let X and Y be bounds on desired solutions $|x_0|$ and $|y_0|$. Define $D = \max_{ij} |p_{ij}| X^i Y^j$. Choose $\alpha > 0$. Assume*

$$X^{\delta+(\alpha\tau/2)} Y^{\tau+(\delta/(2\alpha))} < D \times 2^{-3(\delta^2+\tau^2)-2} .$$

In time polynomial in δ, τ and $\log_2 D$, our algorithm will produce all integer pairs (x_0, y_0) with $|x_0| < X$, $|y_0| < Y$, and $p(x_0, y_0) = 0$.

Let $p(x,y)$ be as before, but with total degree δ. Assume

$$(XY)^{\delta} < D \times 2^{-6\delta^2-2} .$$

In time polynomial in δ and $\log_2 D$, our algorithm will produce all integer pairs (x_0, y_0) with $|x_0| < X$, $|y_0| < Y$, and $p(x_0, y_0) = 0$.

Proof. The proof will be given in the full paper, but is quite similar to that of Theorem 1 and Corollary 2. □

5 More variables

Suppose we have a polynomial $p(x,y,z)$ in three variables. We can mimic the present approach. If the ranges X, Y, Z are small enough, we will end up with a polynomial relation $u(x,y,z)$, not a multiple of p, satisfied by (x_0, y_0, z_0). Then the resultant of p and u with respect to z will give a polynomial $v(x,y)$ in two variables. We can then try to solve v by the current methods. But the degree of v will be quite high, so that the ranges X and Y which can be tolerated will be quite small. This approach will be unsatisfactory in general.

We still have a heuristic procedure which *might* work for a given multivariate polynomial. We are guaranteed to find a space of codimension 1 (a hyperplane) containing all the short vectors of the lattice M_3. But we might easily find a space of larger codimension. (There is a good possibility that for many basis vectors $\mathbf{b}_i$, the orthogonal component $|\mathbf{b}_i^*|$ exceeds our known upper bound on $|\mathbf{s}|$, and each one increases the codimension of the space which contains all the short vectors.) We develop several polynomial equations $u_i(x,y,z)$ satisfying $u_i(x_0, y_0, z_0) = 0$; the number of such equations is equal to the codimension of this space. We can then take resultants and gcd's of the various u_i and p and hope to produce a single polynomial equation in a single variable $v(x_0) = 0$, which we solve over $\mathbb{Z}$. This is only a heuristic approach, which might or might not work for a given polynomial p. (Even if we can generate several equations, they may not be independent.)

There must be limits to the success of this approach in general. Manders and Adleman [3] show that finding suitably bounded solutions to $p_N(x,y,z) = x^2 - yN - z = 0$ is NP-hard. Nonetheless the approach might work for a particular polynomial, and it is worth trying.

6 Comparison with the univariate modular algorithm

In a companion paper [1], the author applies a very similar algorithm for the solution of a univariate modular polynomial.

Two differences between the algorithms are worth noticing. In the modular case, the size X of the acceptable solutions x_0 was related to the modulus N. In the present integer case, there is no such natural measure as N, and we needed to develop a bound in terms of D.

A second difference is that in the modular case, we were able to define polynomials $q_{ij}(x) = x^i p(x)^j$ and assert that $q_{ij}(x_0) = 0 \pmod{N^j}$; the extra information ($\mathrm{mod} N^j$ versus $\mathrm{mod}\ N$) improved our bound on X from $N^{1/(2\delta-1)}$ to $N^{1/\delta}$. In the present integer case, using the polynomial equations $q_{ijk}(x,y) = x^i y^j p(x,y)^k = 0$ would not help, because (for the appropriate ranges of indices i, j, k) the integer linear combinations of the polynomials q_{ij} are exactly the same as those of the polynomials q_{ijk}. For example, with the given $p(x,y)$, $\delta = 1$, and setting $k = 4$, the integer linear combinations of

$$q_{ij}(x,y) = x^i y^j p(x,y), \quad 0 \le i,j < 4$$

are the same as the integer linear combinations of

$$p, xp, x^2p, x^3p, yp, y^2p, y^3p, p^2, xp^2, x^2p^2, yp^2, y^2p^2, p^3, xp^3, yp^3, p^4 \ ,$$

so that we end up defining the same matrix M_1 (up to elementary column operations). (If p is not monic, we appear to gain something from the high coefficient of p, but we actually lose a corresponding amount in the proof, so that using powers of p still neither helps nor hurts us.) This much is also true in the modular case; however, there we gain the extra advantage of working $\mathrm{mod}\ N^j$ as opposed to working $\mathrm{mod} N$, and here in the integer case we can derive no such advantage.

7 Comparison with previous work

Rivest and Shamir [5] solve the problem of factorization if given $(\log_2 N)/3$ bits rather than $(\log_2 N)/4$ bits as we do. They too use lattice methods, but only one polynomial $q_{00}(x,y) = p(x,y)$.

Vallée et al. [6] employ a method similar to [5] in the case of modular polynomials, again using only one polynomial.

Maurer [4] uses a different approach, related to factoring algorithms based on smooth integers, to ask $(\epsilon \log_2 N)$ yes/no oracle questions and determine the factorization of N with failure probability $O\left(N^{-\epsilon/2}\right)$.

8 Application to RSA variant

Vanstone and Zuccherato [7] propose an identity-based variant of RSA in which the user's modulus N is related to his identity. For example, the high order bits of N may be the user's name encoded in ASCII.

Unfortunately, the modulus N is generated in such a way that somewhat more than the high order $((1/4)\log_2 N)$ bits of P are revealed to the public. This enables the present attack to discover the factorization of each modulus and break the scheme.

9 Acknowledgements

Matt Franklin and Mike Reiter's Crypto 95 rump session talk opened this whole line of investigation; subsequent discussions were also useful. The author gratefully acknowledges enlightening discussions with Andrew Odlyzko. Barry Trager helped the experimental effort by coding up in Axiom an implementation of the Lenstra Lenstra Lovasz lattice basis reduction algorithm. The Eurocrypt program committee made suggestions which improved the presentation of the paper.

References

1. D. Coppersmith, "Finding a Small Root of a Univariate Modular Equation," Proceedings of Eurocrypt 96.
2. A. K. Lenstra, H. W. Lenstra and L. Lovasz, "Factoring Polynomials with Integer Coefficients," Matematische Annalen **261** (1982), 513–534.
3. K. Manders and L. Adleman, "NP-complete decision problems for binary quadratics," *J. Comput. System Sci.* **16**, 168-184.
4. U. M. Maurer, "Factoring with an Oracle," Advances in Cryptology – EUROCRYPT '92, Springer LNCS **658** (1993) 429–436.
5. R. L. Rivest and A. Shamir, "Efficient factoring based on partial information," Advances in Cryptology – EUROCRYPT '85, Springer LNCS **219** (1986) 31–34.
6. B. Vallée, M. Girault and P. Toffin, "How to Guess ℓ-th Roots Modulo n by Reducing Lattice Bases," Proceedings of AAECC 6, Springer LNCS **357** (1988) 427–442.
7. S. A. Vanstone and R. J. Zuccherato, "Short RSA Keys and Their Generation," Journal of Cryptology **8** number 2 (Spring 1995) 101–114.

10 Appendix

In this appendix we give a proof of the technical result needed in Section 2: that several columns of the matrix WM_1 are "nearly orthogonal". A modification of this proof would apply to any Toeplitz matrix.

We develop the corresponding result for a general bivariate polynomial. Let M_4 be the right-hand block of WM_1. Let $p(x, y)$ have degree δ in each variable separately. Define indexing functions γ and μ for M_4: The rows of M_4 are indexed by $\gamma(g, h) = (k+\delta)g + h$ for $0 \leq g, h < k+\delta$, and the columns by $\mu(i, j) = ki + j$ for $0 \leq i, j < k$. Define $\tilde{p}(x, y) = p(Xx, Yy)$, so that $\tilde{p}_{ab} = X^a Y^b p_{ab}$ and $\max_{ab} |p_{ab}| = D$.

Lemma 4. *There is a $k^2 \times k^2$ submatrix of M_4 whose determinant has absolute value at least*

$$D^{k^2} 2^{-6k^2\delta^2 - 2k^2} .$$

If the largest coefficient of $\tilde{p}$ is one of $\tilde{p}_{00}$, $\tilde{p}_{0\delta}$, $\tilde{p}_{\delta 0}$ or $\tilde{p}_{\delta\delta}$, then the absolute value of the determinant is exactly D^{k^2}.

Proof. Select indices (a, b) so that $D = |\tilde{p}_{ab}|$ is the largest coefficient of $\tilde{p}$. Select indices (c, d) to maximize the quantity

$$8^{(c-a)^2+(d-b)^2} |\tilde{p}_{cd}| .$$

Select the rows

$$\gamma(c + i, d + j), 0 \le i, j < k$$

of M_4 to create the desired submatrix $\tilde{M}$. The rows and columns of $\tilde{M}$ are indexed by $\mu(i, j) = ki + j$. The matrix element $\tilde{M}_{\mu(g,h),\mu(i,j)}$ is the coefficient of $x^{c+g} y^{d+h}$ in $x^i y^j \tilde{p}(x, y)$, namely

$$\tilde{M}_{\mu(g,h),\mu(i,j)} = \tilde{p}_{g-i+c,h-j+d}$$

Multiply the $\mu(g, h)$ row of $\tilde{M}$ by $8^{2(c-a)g+2(d-b)h}$, and multiply the $\mu(i, j)$ column by $8^{-2(c-a)i-2(d-b)j}$, to create a new matrix M' with the same determinant. Its typical element is

$$M'_{\mu(g,h),\mu(i,j)} = \tilde{p}_{g-i+c,h-j+d} 8^{2(c-a)(g-i)+2(d-b)(h-j)}.$$

From maximality of (c, d) we find

$$|\tilde{p}_{g-i+c,h-j+d}| 8^{(g-i+c-a)^2+(h-j+d-b)^2} \le |\tilde{p}_{cd}| 8^{(c-a)^2+(d-b)^2},$$

from which

$$|\tilde{p}_{g-i+c,h-j+d}| 8^{2(g-i)(c-a)+2(h-j)(d-b)} \le |\tilde{p}_{cd}| 8^{-(g-i)^2-(h-j)^2}.$$

Thus each diagonal entry of M' is $\tilde{p}_{cd}$, and each off-diagonal entry is bounded in absolute value by $|\tilde{p}_{cd}| 8^{-(g-i)^2-(h-j)^2}$. This implies that M' is diagonally dominant, because the absolute values of the off-diagonal entries in its $\mu(i, j)$ row sum to at most

$$|\tilde{p}_{cd}| \times \sum_{(g,h)\neq(i,j)} 8^{-(g-i)^2-(h-j)^2} = |\tilde{p}_{cd}| \times \sum_{(a,b)\neq(0,0)} 8^{-a^2-b^2}$$
$$= |\tilde{p}_{cd}| \times \left[-1 + \sum_{(a,b)} 8^{-a^2-b^2}\right] = |\tilde{p}_{cd}| \times \left[-1 + (\sum_a 8^{-a^2})^2\right] < \tfrac{3}{4}|\tilde{p}_{cd}|$$

Each eigenvalue of M' is within $\frac{3}{4}|\tilde{p}_{cd}|$ of $\tilde{p}_{cd}$, and so exceeds $\frac{1}{4}|\tilde{p}_{cd}|$ in absolute value. By choice of (c, d) we know

$$8^{(c-a)^2+(d-b)^2} |\tilde{p}_{cd}| \ge 8^0 |\tilde{p}_{ab}| = D$$

$$|\tilde{p}_{cd}| \ge 8^{-2\delta^2} D$$

$$|\det(M')| \geq \left(\frac{1}{4}|\tilde{p}_{cd}|\right)^{k^2} \geq \left(\frac{1}{4}8^{-2\delta^2}D\right)^{k^2} = D^{k^2}2^{-6k^2\delta^2-2k^2}.$$

For the second claim of the lemma: If the largest coefficient of $\tilde{p}$ is either $\tilde{p}_{00}$ or $\tilde{p}_{\delta\delta}$, set $(c,d) = (a,b)$ and notice that $\tilde{M}$ is an (upper or lower) triangular matrix whose diagonal entries have absolute value D. If the largest coefficient is either $\tilde{p}_{0\delta}$ or $\tilde{p}_{\delta 0}$, redefine the indexing function on columns as $\mu(i,j) = ki+(k-1-j)$ so that again $\tilde{M}$ is a triangular matrix whose diagonal entries have absolute value D. Similar results hold if (a,b) is any corner of the Newton polygon associated with $\tilde{p}$.

For the particular case $p(x,y) = (P_0 + x)(Q_0 + y) - N$, we have $\delta = 1$, and the only non-zero coefficients of $\tilde{p}$ are $\tilde{p}_{00}$, $\tilde{p}_{0\delta}$, $\tilde{p}_{\delta 0}$ and $\tilde{p}_{\delta\delta}$; thus the second claim must hold. □

The lemma gives a $k^2 \times k^2$ submatrix $\tilde{M}$ of M_4, where M_4 is the right-hand $(k+\delta)^2 \times k^2$ block of WM_1. To apply the lemma, we need to find k^2 column indices in the left-hand block of WM_1 whose deletion leaves a $(k+\delta)^2 \times (k+\delta)^2$ submatrix $W\hat{M}_1$ of WM_1 with $|\det(W\hat{M}_1)| = |\det(\tilde{M})|$. We simply delete those columns whose indices match those of the k^2 rows accepted in $\tilde{M}$. Recall that the left-hand side of WM_1 is the identity matrix, so each remaining left-hand column has a single 1 among the 0's, and expansion by minors gives $|\det(W\hat{M}_1)| = |\det(\tilde{M})|$ as desired.