Montgomery's method of polynomial selection for the number field sieve

Nicholas Coxon

MONTGOMERY'S METHOD OF POLYNOMIAL SELECTION FOR THE NUMBER FIELD SIEVE

NICHOLAS COXON

Abstract. The number field sieve is the most efficient known algorithm for factoring large integers that are free of small prime factors. For the polynomial selection stage of the algorithm, Montgomery proposed a method of generating polynomials which relies on the construction of small modular geometric progressions. Montgomery's method is analysed in this paper and the existence of suitable geometric progressions is considered.

1. Introduction

In this paper, N denotes a positive integer that is destined to be factored. When N is large and free of small factors, the most efficient publicly known algorithm for determining its factors is the number field sieve [20]. Such N include RSA [28] moduli, for which numerous record factorisations have been achieved with the number field sieve, including the current 768-bit record [17].

The number field sieve is comprised of several stages, commonly referred to as polynomial selection, sieving, filtering, linear algebra and square root computation. The polynomial selection stage requires the selection of coprime irreducible polynomials $f_1, f_2 \in \mathbb{Z}[x]$ that have a common root modulo N. After polynomial selection, sieving is used to identify coprime integer pairs (a, b) such that the prime factors of $f_i(a/b)b^{\deg f_i}$ are below some bound y_i for $i = 1, 2$. Obtaining sufficiently many pairs with this property, called *relations*, is the most time consuming stage of the number field sieve, with the time taken greatly influenced by the choice of polynomials [24, 25].

Let $\Psi(x, y)$ denote the number of positive integers less than x that are free of prime factors greater than y. Canfield, Erdős and Pomerance [6] showed that for any $\varepsilon > 0$, $\Psi(x, x^{1/u}) = xu^{-u(1+o(1))}$ for $u \to \infty$, uniformly in the region $x \geq u^{u(1+\varepsilon)}$. It follows, heuristically, that in the polynomial selection stage of the number field sieve, the polynomials f_1 and f_2 should be chosen to minimise the size of the values $f_1(a/b)b^{\deg f_1}$ and $f_2(a/b)b^{\deg f_2}$ over the pairs (a, b) considered in the sieve stage. Thus, it is necessary for the polynomials to have small coefficients. As a result, the degrees of f_1 and f_2 should not be too small. However, the degrees should not be too large either, since $f_i(a/b)b^{\deg f_i}$ is a homogeneous polynomial of degree $\deg f_i$ in a and b. In practice, low-degree polynomials are used. For example, the two largest factorisations of RSA moduli [17, 4] both used a sextic polynomial together with a linear polynomial. To quantify the coefficient size of a polynomial, the skewed

Date: December 18, 2014.

2010 *Mathematics Subject Classification.* Primary 11Y05, 11Y16.

Key words and phrases. Integer factorisation, number field sieve, polynomial selection, Montgomery's method.

2-norm $\|.\|_{2,s}$ is used. The norm is defined as follows: if $f = \sum_{i=0}^{d} a_i x^i$ is a degree d polynomial with real coefficients, then

$$\|f\|_{2,s} = \sqrt{\sum_{i=0}^{d} \left(a_i s^{i-\frac{d}{2}}\right)^2} \quad \text{for all } s > 0.$$

The parameter s captures the shape of the sieve region, which is modelled by a rectangular region $[-A, A] \times (0, B]$ or an elliptic region

$$\left\{(x, y) \in \mathbb{R}^2 \mid 0 < y \leq B\sqrt{1 - (x/A)^2}\right\}$$

such that $A/B = s$. In practice, the polynomial selection stage proceeds by first generating many "raw" polynomial pairs with small coefficients. Then various methods of optimisation [25, 3, 2] are used to improve the quality of the raw pairs by taking into account additional factors that influence a pair's yield of relations, such as the presence of real roots and roots modulo small primes [24, 25].

The methods of polynomial selection used in all recent record factorisations [24, 25, 15, 16] produce polynomials f_1 and f_2 such that one polynomial is linear. However, it is expected that a significant advantage is gained by using two nonlinear polynomials [9, Section 6.2.7] (see also [27, Section 4] for practical considerations relating to sieving). Montgomery [22, 23] provided a method for generating two nonlinear polynomials with small coefficients. This paper extends and sharpens Montgomery's original analysis of the method.

2. Montgomery's method

A *geometric progression* of *length* ℓ and *ratio* r modulo N is an integer vector $[c_{\ell-1}, \ldots, c_0]$ such that $c_i \equiv c_0 r^i \pmod{N}$ for $i = 0, \ldots, \ell - 1$. Square brackets are used to distinguish geometric progressions from regular vectors, which are denoted with round brackets. Montgomery [22, 23] showed that a length $2d - 1$ geometric progression $\boldsymbol{c} = [c_{2d-2}, \ldots, c_0]$ modulo N such that $\gcd(c_0, c_1, \ldots, c_{d-2}, N) = 1$ and

$$C = C(c_{2d-2}, \ldots, c_0) = \begin{pmatrix} c_{2d-2} & c_{2d-3} & \cdots & c_{d-1} \\ c_{2d-3} & c_{2d-4} & \cdots & c_{d-2} \\ \vdots & \vdots & & \vdots \\ c_{d-1} & c_{d-2} & \cdots & c_0 \end{pmatrix}$$

has full rank can be used to construct polynomials $f_1, f_2 \in \mathbb{Z}[x]$ of maximum degree d that have a common root modulo N. Once a suitable geometric progression $\boldsymbol{c}$ has been found, Montgomery's method proceeds by computing a basis $\{(a_{1,d}, \ldots, a_{1,0})^T, (a_{2,d}, \ldots, a_{2,0})^T\}$, where A^T denotes the transpose of a matrix A, for the free $\mathbb{Z}$-module that is the set of integer vectors in the kernel of the matrix

$$\partial C = \begin{pmatrix} c_{2d-2} & c_{2d-3} & \cdots & c_{d-2} \\ c_{2d-3} & c_{2d-4} & \cdots & c_{d-3} \\ \vdots & \vdots & & \vdots \\ c_d & c_{d-1} & \cdots & c_0 \end{pmatrix}. \tag{2.1}$$

The basis yields polynomials $f_1 = \sum_{i=0}^{d} a_{1,i} x^i$ and $f_2 = \sum_{i=0}^{d} a_{2,i} x^i$. If r is the ratio of the geometric progression modulo N, then $(r^d, r^{d-1}, \ldots, 1)$ is a linear combination of the row vectors of ∂C modulo N since $\gcd(c_0, c_1, \ldots, c_{d-2}, N) = 1$.

Thus, r is a root of f_1 and f_2 modulo N. Denote by $\widehat{\partial C}$ the submatrix of ∂C obtained by deleting its first column. Then $\widehat{\partial C}$ has full rank since it is equal to the submatrix of C obtained by deleting its first row. Consequently, the kernel of ∂C is 2-dimensional. Moreover, and at least one of f_1 and f_2 has degree equal to d, otherwise $(a_{1,d-1}, \ldots, a_{1,0})^T$ and $(a_{2,d-1}, \ldots, a_{2,0})^T$ are linearly independent and in the 1-dimensional kernel of $\widehat{\partial C}$, which is absurd. Finally, as was observed by Montgomery [22] and which is shown to hold in this paper, the polynomials f_1 and f_2 are coprime since C is nonsingular.

To ensure that the norms $\|f_1\|_{2,s}$ and $\|f_2\|_{2,s}$ are small for some $s > 0$, the basis is chosen such that

$$\left\{ \left(a_{1,d}s^d, a_{1,d-1}s^{d-1}, \ldots, a_{1,0}\right)^T, \left(a_{2,d}s^d, a_{2,d-1}s^{d-1}, \ldots, a_{2,0}\right)^T \right\} \tag{2.2}$$

is Lagrange-reduced (see [26, p. 41]). As a result,

$$s^{\frac{\deg f_1 + \deg f_2}{2} - d} \|f_1\|_{2,s} \|f_2\|_{2,s} \leq \frac{\gamma_2}{N^{d-2}} \|\boldsymbol{c}\|_{2,s^{-1}}^{d-1} \tag{2.3}$$

(see [8, Section 3.2]), where $\gamma_2 = 2/\sqrt{3}$ is Hermite's constant for dimension two, and the vector norm $\|.\|_{2,s}$ is defined as follows: if $\boldsymbol{v} = (v_n, v_{n-1}, \ldots, v_0)$ is a real $(n+1)$-dimensional vector, then

$$\|\boldsymbol{v}\|_{2,s} = \sqrt{\sum_{i=0}^{n} \left(v_i s^{i - \frac{n}{2}}\right)^2} \quad \text{for all } s > 0.$$

Consequently, it is a requirement of Montgomery's method that $\|\boldsymbol{c}\|_{2,s^{-1}}$ is small. Therein lies the difficulty of the method, as the basis (2.2) is readily computed in polynomial time (see [8, Section 3.1.2]). The problem of constructing small geometric progressions has been addressed by several authors [22, 31, 27, 18, 8].

It is natural to consider the existence of small geometric progressions. Montgomery [22] showed that if there exist two degree d polynomials that have small coefficients and a common root r modulo N, then there exists a small length $2d-1$ geometric progression with ratio r modulo N. Montgomery's proof is constructive and is generalised to two polynomials f_1 and f_2 of maximum degree d in this paper. Furthermore, it is shown that if the polynomials are coprime, then the geometric progression given by the construction is the unique vector $(c_{2d-1}, \ldots, c_0)$, up to scalar multiple, such that the coefficient vectors of f_1 and f_2 are in the kernel of the matrix ∂C defined in (2.1). As a result, the analysis of the construction contributes to the analysis of Montgomery's method.

This paper is organised as follows: the definitions and some properties of the Sylvester matrix, the Bezout matrix and the resultant are reviewed in the next section; the generalisation of Montgomery's geometric progression construction is presented and analysed in Section 4; and a full analysis of Montgomery's method is provided in Section 5.

3. The Sylvester matrix, the Bezout matrix and the resultant

Matrices with the property that each of their rows contain the coefficients of some polynomial are frequently encountered in this paper. The Sylvester and Bezout matrices are constructed in this manner. Consequently, compact notation for such matrices is defined before introducing the protagonists of this section.

For $m \geq 1$ polynomials $f_1, \ldots, f_m$ and any integer $n \geq \max_{1 \leq i \leq m} \deg f_i$, denote by $(f_1, \ldots, f_m)_n$ the $m \times (n+1)$ matrix $(a_{i,j})_{i=1,\ldots,m;j=1,\ldots,n+1}$ where $a_{i,j}$ is the coefficient of x^{n+1-j} in f_i. When $n = \max_{1 \leq i \leq m} \deg f_i$, the subscript n is drop, giving the notation $(f_1, \ldots, f_m)$. The parameter n is viewed as the formal degree of the polynomials $f_1, \ldots, f_m$. For example, (f) is the vector of coefficients of f, while $(f)_n$ for some $n \geq \deg f$ is the vector of coefficients of f when view as a polynomial of formal degree n. Define $(\ )_n$ to be the $0 \times n$ empty matrix.

3.1. **The Sylvester matrix and the resultant.** Let $\mathbb{A}$ be an integral domain. Then the *Sylvester matrix* of non-constant polynomials $f_1, f_2 \in \mathbb{A}[x]$ is the matrix

$$\mathrm{Syl}(f_1, f_2) = \left(x^{\deg f_2 - 1} f_1, \ldots, f_1, x^{\deg f_1 - 1} f_2, \ldots, f_2\right).$$

The determinant of $\mathrm{Syl}(f_1, f_2)$ is called the *resultant* of f_1 and f_2, and is denoted $\mathrm{Res}(f_1, f_2)$. The resultant of f_1 and f_2 is zero if and only if the polynomials have a nontrivial gcd over the field of fractions of $\mathbb{A}$.

For non-constant polynomials $f_1, f_2 \in \mathbb{Z}[x]$ and all real numbers $s > 0$, define $\theta_s(f_1, f_2)$ to be the angle (in $[0, \pi]$) between the row vectors of $(f_1(sx), f_2(sx))$. The following lemma provides upper and lower bounds on the resultant of a pair of number field sieve polynomials (see [8, Section 2.1.2] for a proof):

Lemma 3.1. *Suppose that $f_1, f_2 \in \mathbb{Z}[x]$ are non-constant, coprime and have a common root modulo N. Then*

$$N \leq |\mathrm{Res}(f_1, f_2)| \leq |\sin \theta_s(f_1, f_2)|^{\min\{\deg f_1, \deg f_2\}} \, \|f_1\|_{2,s}^{\deg f_2} \, \|f_2\|_{2,s}^{\deg f_1}$$

for all $s > 0$.

As a consequence of Lemma 3.1, a pair of number field sieve polynomials in considered to have optimal resultant if it is equal to N, and optimal coefficient size if $\|f_1\|_{2,s}^{\deg f_2} \, \|f_2\|_{2,s}^{\deg f_1}$ is $O(N)$. Thus, inequality (2.3) implies that a pair of degree d polynomials generated by Montgomery's method has optimal coefficient size if $\|\boldsymbol{c}\|_{2,s^{-1}} = O(N^{1-1/d})$.

3.2. **The Bezout matrix.** Let $\mathbb{A}$ be an integral domain and $f_1, f_2 \in \mathbb{A}[x]$ be non-constant. Write $f_i = \sum_{j=0}^{d} a_{i,j} x^j$ for $i = 1, 2$ such that $d = \max\{\deg f_1, \deg f_2\}$ and the coefficients $a_{i,j}$ are elements of $\mathbb{A}$. Define polynomials

$$p_{i+1} = \left(\sum_{j=0}^{i} a_{2,d-i+j} x^j\right) f_1 - \left(\sum_{j=0}^{i} a_{1,d-i+j} x^j\right) f_2 \quad \text{for } i = 0, \ldots, d-1. \tag{3.1}$$

Then the *Bezout matrix*, or *Bezoutian*, of f_1 and f_2 is the matrix $\mathrm{Bez}(f_1, f_2) = (p_1, \ldots, p_d)_{d-1}$. Denote by $\mathrm{lc}(f)$ the leading coefficient of a polynomial f. Similar to the Sylvester matrix, the determinant of the Bezout matrix of f_1 and f_2 is related to their resultant (see [29, Section 2]):

$$\det \mathrm{Bez}(f_1, f_2) = (-1)^{\frac{d(d+1)}{2}} \, \mathrm{lc}(f_1)^{d - \deg f_2} \left((-1)^d \, \mathrm{lc}(f_2)\right)^{d - \deg f_1} \mathrm{Res}(f_1, f_2). \tag{3.2}$$

A *Hankel matrix* over $\mathbb{A}$ is a square matrix $H = (h_{i,j})_{i=1,\ldots,n;j=1,\ldots,n}$ such that $h_{i,j} = h_{i+j-1}$ for some $h_1, \ldots, h_{2n-1} \in \mathbb{A}$. Lander [19] showed that the inverse of a Bezout matrix is a Hankel matrix and, conversely, that the inverse of a Hankel matrix is the Bezout matrix of two polynomials. For an $m \times n$ matrix $H = (h_{i,j})_{i=1,\ldots,m;j=1,\ldots,n}$ such that $h_{i,j} = h_{i+j-1}$ for some $h_1, \ldots, h_{m+n-1} \in \mathbb{A}$, define $\partial^k H = (h_{i+j-1})_{i=1,\ldots,m+k;j=1,\ldots,n-k}$ for $k = 0, \ldots, m-1$. Let ∂H denote

the matrix $\partial^1 H$. Define $\widehat{\partial H} = (h_{i+j})_{i=1,\ldots,m+1;j=1,\ldots,n-2}$. The following result of Heinig and Rost [14, Theorem 4.2] expresses the inverse of a real Hankel matrix H as the Bezout matrix of two polynomials obtained from the kernel of ∂H:

Lemma 3.2. *Let $H = (h_{i+j-1})_{i=1,\ldots,d;j=1,\ldots,d}$ be a real nonsingular Hankel matrix. If $f_1, f_2 \in \mathbb{R}[x]$ such that $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis of the kernel of ∂H, then*

$$H^{-1} = -\frac{1}{\det \psi} \operatorname{Bez}(f_1, f_2) \quad \textit{where} \quad \psi = \begin{pmatrix} h_d & \ldots & h_{2d-1} & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix} (f_1, f_2)_d^T .$$

The formulation of the Bezout matrix presented in this paper is different to that used by Heinig and Rost. Thus, it is necessary to convert between the two formulations:

Proof of Lemma 3.2. Let $H = (h_{i+j-1})_{i=1,\ldots,d;j=1,\ldots,d}$ be a real nonsingular Hankel matrix. Suppose that $f_1, f_2 \in \mathbb{R}[x]$ such that $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis of the kernel of ∂H. Define ψ as in the statement of the theorem. The matrix $\widehat{\partial H}$ is equal to the submatrix of H obtained by deleting its first row. Thus, $\widehat{\partial H}$ has full rank. It follows that $\max\{\deg f_1, \deg f_2\} = d$, otherwise $(f_1)_{d-1}^T$ and $(f_2)_{d-1}^T$ would be linearly independent and in the kernel of $\widehat{\partial H}$. Interchanging f_1 and f_2 changes $\operatorname{Bez}(f_1, f_2)$ and $\det \psi$ by a factor of -1. Therefore, assume without loss of generality that $\deg f_1 = d$. Write $f_i = \sum_{j=0}^{d} a_{i,j} x^j$ for $i = 1, 2$ such that the coefficients $a_{i,j}$ are real numbers. Define $g_i = \sum_{j=0}^{d} a_{i,d-j} x^j$ for $i = 1, 2$. Then Theorem 4.2 of Heinig and Rost [14] implies that

$$H^{-1} = -\frac{1}{\det \psi} \mathrm{B}(g_1, g_2) \tag{3.3}$$

where the matrix $\mathrm{B}(g_1, g_2)$, called the (Hankel) Bezoutian of g_1 and g_2 by Heinig and Rost [14, Section 2.1] (and many other authors), is defined as follows: $\mathrm{B}(g_1, g_2) = (b_{i,j})_{i=1,\ldots,d;j=1,\ldots,d}$ such that $b_{i,j}$ is the coefficient of $x^{i-1}y^{i-1}$ in the polynomial

$$b(x, y) = \frac{g_1(x)g_2(y) - g_2(x)g_1(y)}{x - y}.$$

Expanding the numerator of $b(x, y)$ shows that

$$\begin{aligned} b(x, y) &= \sum_{k=0}^{d} a_{2,d-k} \left(\sum_{i=0}^{d} a_{1,d-i} \frac{x^i y^k - x^k y^i}{x - y} \right) \\ &= \sum_{k=0}^{d} a_{2,d-k} \left(\sum_{i=k+1}^{d} a_{1,d-i} \frac{x^{i-k} - y^{i-k}}{x - y} x^k y^k - \sum_{i=0}^{k-1} a_{1,d-i} \frac{x^{k-i} - y^{k-i}}{x - y} x^i y^i \right). \end{aligned}$$

Therefore,

$$\mathrm{B}(g_1, g_2) = \sum_{k=0}^{d} a_{2,d-k} \begin{pmatrix} & & -a_{1,d} & & & \\ & \iddots & \vdots & & & \\ -a_{1,d} & \ldots & -a_{d-k+1} & a_{d-k-1} & \ldots & a_0 \\ & & & \vdots & \iddots & \\ & & & a_0 & & \end{pmatrix},$$

where the omitted entries are zeros. Let $k \in \mathbb{Z}$ such that $0 \leq k \leq d$, and $p_1, \ldots, p_d \in \mathbb{R}[x]$ such that $\operatorname{Bez}(f_1, x^{d-k}) = (p_1, \ldots, p_d)$. Then

$$p_{i+1} = -x^{d-k} \sum_{j=0}^{i} a_{1,d-i+j} x^j \quad \text{for } i = 0, \ldots, k-1,$$

and

$$p_{i+1} = x^{i-k} f_1 - x^{k-d} \sum_{j=0}^{i} a_{1,d-i+j} x^j = x^{i-k} \sum_{j=0}^{d-i-1} a_j x^j \quad \text{for } i = k, \ldots, d-1.$$

Hence,

$$\operatorname{B}(g_1, g_2) = \sum_{k=0}^{d} a_{2,d-k} \operatorname{Bez}(f_1, x^{d-k}) = \operatorname{Bez}(f_1, f_2).$$

Combining this equation with (3.3) completes the proof. □

4. Existence of geometric progressions

The *integer kernel* of an integer matrix A is the free $\mathbb{Z}$-module consisting of all integer column vectors $\boldsymbol{x}$ such that $A\boldsymbol{x} = \boldsymbol{0}$. Montgomery [22] showed that if two coprime degree $d \geq 2$ integer polynomials f_1 and f_2 have a common root modulo N, then a vector $\boldsymbol{c} \in \mathbb{Z}^{2d-1}$ such that $\boldsymbol{c}^T$ spans the integer kernel of the matrix

$$\left(x^{d-2} f_1, \ldots, f_1, x^{d-2} f_2, \ldots, f_2\right) \tag{4.1}$$

is a geometric progression modulo N. Moreover, $\|\boldsymbol{c}\|_{2,1} = O(\|f_1\|_{2,1}^{d-1} \|f_2\|_{2,1}^{d-1})$ since the entries of $\boldsymbol{c}$ are, up to a constant, order $2d-2$ minors of the matrix in (4.1). Koo, Jo and Kwon [18, Theorem 2] generalise Montgomery's result by providing a construction which, for $k \in \{1, \ldots, d-1\}$, uses $j = \lceil (d-1)/k \rceil + 1$ degree d polynomials $f_1, \ldots, f_j$ with a common root modulo N to construct a length $d+k-1$ geometric progression $\boldsymbol{c}$ such that $\|\boldsymbol{c}\|_{2,1} = O(\max_{1 \leq i \leq j} \|f_i\|_{2,1}^{d+k-1})$. In this section, another generalisation is presented, with two polynomials f_1 and f_2 of maximum degree d used to construct geometric progressions of lengths $\deg f_1 + \deg f_2 - 1, \ldots, 2d-1$. The geometric progressions are shown to have small size whenever f_1 and f_2 have small coefficients. Therefore, the existence of small geometric progressions for Montgomery's method, and relaxations of the method that employ shorter progressions [31, 27, 18, 8], is established under the assumption that good nonlinear polynomial pairs exist.

Let $\mathbb{A}$ be an integral domain. For $f_1, f_2 \in \mathbb{A}[x]$ such that $2 \leq \deg f_2 \leq \deg f_1$, and $t \in \{\deg f_2, \ldots, \deg f_1\}$, define the $(\deg f_1 + t - 2) \times (\deg f_1 + t - 1)$ matrix

$$\operatorname{S}_t(f_1, f_2) = \left(x^{t-2} f_1, \ldots, f_1, x^{\deg f_1 - 2} f_2, \ldots, f_2\right).$$

Define signed minors $\operatorname{M}_{t,1}(f_1, f_2), \ldots, \operatorname{M}_{t,\deg f_1 + t - 1}(f_1, f_2)$ of the matrix $\operatorname{S}_t(f_1, f_2)$ as follows: for $i = 1, \ldots, \deg f_1 + t - 1$, $\operatorname{M}_{t,i}(f_1, f_2)$ is equal to $(-1)^{1+i}$ times the determinant of the submatrix of $\operatorname{S}_t(f_1, f_2)$ obtained by deleting its ith column. Define

$$\boldsymbol{c}_t(f_1, f_2) = (\operatorname{M}_{t,1}(f_1, f_2), \ldots, \operatorname{M}_{t,\deg f_1 + t - 1}(f_1, f_2))$$

for $t = \deg f_2, \ldots, \deg f_1$. When f_1 and f_2 are clear from the context, S_t is used to denote $\operatorname{S}_t(f_1, f_2)$, $\operatorname{M}_{t,i}$ is used to denote $\operatorname{M}_{t,i}(f_1, f_2)$, and $\boldsymbol{c}_t$ is used to denote

$\boldsymbol{c}_t(f_1, f_2)$. The ith entry of the vector $\mathrm{S}_t(f_1, f_2) \cdot \boldsymbol{c}_t(f_1, f_2)^T$ is equal to the determinant of the matrix obtained by appending the ith row of $\mathrm{S}_t(f_1, f_2)$ to its top, and thus is equal to zero. Therefore,

$$\mathrm{S}_t(f_1, f_2) \cdot \boldsymbol{c}_t(f_1, f_2)^T = \mathbf{0}_{\deg f_1 + t - 2} \quad \text{for } t = \deg f_2, \ldots, \deg f_1, \tag{4.2}$$

where $\mathbf{0}_n$ denotes the n-dimensional column vector of zeros for $n = 1, 2, \ldots$. If A is an $m \times n$ integer matrix such that $m \leq n$, define $\Delta(A)$ to be the greatest common divisor of all $m \times m$ minors of A, with $\Delta(A) = 0$ if all such minors are zero. Then $\Delta(A)$ is nonzero if and only if A has full rank.

Let $f_1, f_2 \in \mathbb{Z}[x]$ be coprime degree $d \geq 2$ polynomials that have a common root modulo N. Then $\mathrm{S}_d(f_1, f_2)$ is the matrix in (4.1). Therefore, $\mathrm{S}_d(f_1, f_2)$ is the submatrix of $\mathrm{Syl}(f_1, f_2)$ obtained by first deleting its first and $(d+1)$th rows, giving a matrix whose first column contains zeros, then deleting the first column of the resulting matrix. Thus, $\mathrm{S}_d(f_1, f_2)$ has full rank, otherwise $\mathrm{Syl}(f_1, f_2)$ is singular. Therefore, $\boldsymbol{c}_d(f_1, f_2)^T$ is nonzero and in the integer kernel of $\mathrm{S}_d(f_1, f_2)$. Consequently, Montgomery's result implies $\boldsymbol{c}_d(f_1, f_2)/\Delta(\boldsymbol{c}_d(f_1, f_2))$ is a geometric progression modulo N. More generally, if $2 \leq \deg f_2 \leq \deg f_1$, then the vectors $\boldsymbol{c}_t(f_1, f_2)$ for $t = \deg f_2, \ldots, \deg f_1$ are geometric progressions modulo N:

Theorem 4.1. *Let f_1 and f_2 be coprime integer polynomials such that $2 \leq \deg f_2 \leq \deg f_1$, f_1 and f_2 have a common root $r \in \mathbb{Z}$ modulo N, and $\mathrm{lc}(f_1)\,\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2))$ is relatively prime to N. Then, for $t \in \{\deg f_2, \ldots, \deg f_1\}$, the vector*

$$\boldsymbol{c}_t = \boldsymbol{c}_t(f_1, f_2) = (c_{t, \deg f_1 + t - 2}, \ldots, c_{t,0})$$

satisfies the following properties:

1. *$\boldsymbol{c}_t$ is a nonzero geometric progression with ratio r modulo N;*
2. *$\gcd(c_{t,0}, N) = 1$;*
3. *the matrix $C_t = (c_{t, \deg f_1 + t - i - j})_{i=1,\ldots,t; j=1,\ldots,\deg f_1}$ has full rank;*
4. *the vectors $(f_1)^T_{\deg f_1}$ and $(f_2)^T_{\deg f_1}$ are in the kernel of ∂C_t; and*
5. *the inequalities*

$$\|\boldsymbol{c}_t\|_{2, s^{-1}} \leq \left(|\sin \theta_s(f_1, f_2)|\, \|f_1\|_{2,s}\right)^{t-1} \left(s^{\frac{\deg f_2 - t}{2}}\, \|f_2\|_{2,s}\right)^{\deg f_1 - 1}$$

and

$$\|\boldsymbol{c}_t\|_{2, s^{-1}} \geq s^{\frac{t - \deg f_1}{2}} \left|s^{\deg f_2}\, \mathrm{lc}(f_2)\right|^{\frac{\deg f_1}{t} - 1} \left|\mathrm{lc}(f_1)^{t - \deg f_2}\, \mathrm{Res}(f_1, f_2)\right|^{1 - \frac{1}{t}}$$

hold for all $s > 0$.

Theorem 4.1 is proved as a series of lemmas and corollaries in the next section, with some of the properties presented as part of more general statements. The requirement in the theorem that $\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2))$ and N are relatively prime has not been encountered so far in this paper, nor is it usual to require a pair of nonlinear number field sieve polynomials to satisfy it. In the next section, it is shown that

$$\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2)) \leq \left(|\sin \theta_s(f_1, f_2)|\, \|f_1\|_{2,s}\right)^{\deg f_2 - 1} \|f_2\|_{2,s}^{\deg f_1 - 1} \quad \text{for all } s > 0$$

(see Lemma 4.3 and Remark 4.4). Therefore, if f_1 and f_2 are a pair of number field sieve polynomials that are close the attaining the lower bound from Lemma 3.1, then $\gcd(\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2)), N)$ is expected to equal one in practice, or a factorisation of N is possibly obtained. However, as the requirement is new, its meaning is discussed before proceeding to the proof of the theorem.

Let $\mathbb{A}$ be an integral domain, $f_1, f_2 \in \mathbb{A}[x]$ such that $2 \leq \deg f_2 \leq \deg f_1$, and $d_i = \deg f_i$ for $i = 1, 2$. Then the *first subresultant* of f_1 and f_2 is the polynomial

$$\mathrm{Sres}_1(f_1, f_2) = (-1)^{d_1+d_2-1} \left(\mathrm{M}_{d_2,d_1+d_2-2}(f_1, f_2)x - \mathrm{M}_{d_2,d_1+d_2-1}(f_1, f_2)\right).$$

Recall that $\mathrm{Res}(f_1, f_2) = 0$ if and only if $\deg \gcd(f_1, f_2) \geq 1$, where the gcd is computed over the field of fractions of $\mathbb{A}$. This statement is refined by considering the first subresultant of the polynomials: if $\mathrm{Res}(f_1, f_2) = 0$, then $\deg \gcd(f_1, f_2) = 1$ if and only if $\mathrm{Sres}_1(f_1, f_2) \neq 0$ (see, for instance, [21, Section 7.7.1] or [10]).

Suppose now that f_1 and f_2 are coprime integer polynomials. Write $f_1 = \sum_{i=0}^{d_1} a_{1,i}x^i$ and $f_2 = \sum_{i=0}^{d_1} a_{2,i}x^i$ such that the coefficients $a_{i,j}$ are integers. Then using the Laplace expansion (see [1, Section 33]) to compute the determinant of $\mathrm{Syl}(f_1, f_2)$ with respect to rows d_2 and $d_1 + d_2$ shows that

$$\mathrm{Res}(f_1, f_2) = (-1)^{d_2-1} \sum_{i=1}^{d_1} (a_{1,i}a_{2,0} - a_{1,0}a_{2,i}) \, \mathrm{M}_{d_2,d_1+d_2-i}(f_1, f_2).$$

Thus, $\Delta(\mathrm{S}_{d_2}(f_1, f_2))$ divides $\mathrm{Res}(f_1, f_2)$ and $\mathrm{Sres}_1(f_1, f_2)$. Therefore, if p is a prime that divides $\Delta(\mathrm{S}_{d_2}(f_1, f_2))$ and does not divide $\mathrm{lc}(f_1)\,\mathrm{lc}(f_2)$, then the reductions of f_1 and f_2 modulo p have a common factor of degree greater than one. Conversely, if the reductions of f_1 and f_2 modulo a prime p are nonzero and have a common divisor of degree $w \geq 2$, then Gomez et al. [12] showed that p^w divides $\mathrm{Res}(f_1, f_2)$. Modifying their proof shows that p^{w-1} divides the minors $\mathrm{M}_{\deg f_2,i}$, and thus p^{w-1} divides $\Delta(\mathrm{S}_{d_2}(f_1, f_2))$. Hence, if p is a prime that does not divide $\mathrm{lc}(f_1)\,\mathrm{lc}(f_2)$, then p divides $\Delta(\mathrm{S}_{d_2}(f_1, f_2))$ if and only if the reductions of f_1 and f_2 modulo p have a common factor of degree greater than one. If f_1 and f_2 are a pair of number field sieve polynomials, then $\gcd(\mathrm{lc}(f_1)\,\mathrm{lc}(f_2), N)$ is expected to equal one in practice. Therefore, the requirement that $\gcd(\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2)), N) = 1$ denies the possibility that f_1 and f_2 have a factor of degree greater than one modulo some factor of N.

4.1. **Proof of Theorem 4.1.** In this section, f_1 and f_2 are integer polynomials such that $2 \leq \deg f_2 \leq \deg f_1$. Furthermore, let $d = \deg f_1$,

$$\boldsymbol{c}_t = \boldsymbol{c}_t(f_1, f_2) = (c_{t,d+t-2}, \ldots, c_{t,0}) \quad \text{and} \quad C_t = (c_{t,d+t-i-j})_{i=1,\ldots,t;j=1,\ldots,d}$$

for $t = \deg f_2, \ldots, d$. The first lemma of this section proves Property (1) and Property (2) of Theorem 4.1:

Lemma 4.2. *If* $\gcd(\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2)), N) = 1$ *and there exists an integer* r *that is a root of* f_1 *and* f_2 *modulo* N*, then* $\boldsymbol{c}_t$ *is a nonzero geometric progression with ratio* r *modulo* N *and* $\gcd(c_{t,0}, N) = \gcd(\mathrm{lc}(f_1)^{t-\deg f_2}, N)$ *for* $t = \deg f_2, \ldots, d$.

Proof. Suppose that $\gcd(\Delta(\mathrm{S}_{\deg f_2}(f_1, f_2)), N) = 1$ and $r \in \mathbb{Z}$ is a root of f_1 and f_2 modulo N. Let $d_2 = \deg f_2$. The lemma is proved in two steps: first, it is shown that $\boldsymbol{c}_{d_2}$ is a geometric progression with ratio r modulo N; and second, it is shown that if $d_2 < d$ and $\boldsymbol{c}_{t-1}$ is a geometric progression with ratio r modulo N for some $t \in \{d_2 + 1, \ldots, d\}$, then so too is $\boldsymbol{c}_t$.

Let U be a $(d + d_2 - 1) \times (d + d_2 - 1)$ unimodular matrix such that $S_{d_2}U$ is in Hermite normal form (as defined by Cohen [7, Definition 2.4.2]). Performing elementary row operations on the columns of S_{d_2} does not change $\Delta(\mathrm{S}_{d_2})$, which is nonzero since $\gcd(\Delta(\mathrm{S}_{d_2}), N) = 1$. Thus,

$$\mathrm{S}_{d_2}U = \begin{pmatrix} \mathbf{0}_{d+d_2-2} & H \end{pmatrix}$$

for some $(d+d_2-2)\times(d+d_2-2)$ matrix H such that $\det H = \pm\Delta(\mathrm{S}_{d_2})$.

The first column vector of U is a basis for the integer kernel of S_{d_2} (see [7, Proposition 2.4.9]). Thus, (4.2) implies that $\boldsymbol{c}_{d_2}^T$ is equal to $\pm\Delta(\boldsymbol{c}_{d_2})$ times the first column vector of U. The definition of $\boldsymbol{c}_{d_2}$ implies that $\Delta(\boldsymbol{c}_{d_2}) = \Delta(\mathrm{S}_{d_2})$. Therefore, $\boldsymbol{c}_{d_2}$ is nonzero and $\boldsymbol{c}_{d_2}U^{-T} = \pm\left(\Delta(\mathrm{S}_{d_2}), 0, \ldots, 0\right)$.

Let $\boldsymbol{r} = (r^{d+d_2-2}, r^{d+d_2-3}, \ldots, 1)$ and $\boldsymbol{r}U^{-T} = (r_1, \ldots, r_{d+d_2-1})$. Then $\mathrm{S}_{d_2}\boldsymbol{r}^T \equiv \mathbf{0}_{d+d_2-2} \pmod N$ since r is a root of f_1 and f_2 modulo N. Consequently,

$$H(r_2, \ldots, r_{d+d_2-1})^T \equiv (\mathrm{S}_{d_2}U)\left(\boldsymbol{r}U^{-T}\right)^T \equiv \mathrm{S}_{d_2}\boldsymbol{r}^T \equiv \mathbf{0}_{d+d_2-2} \pmod N$$

Therefore, $(r_2, \ldots, r_{d+d_2-1}) \equiv \mathbf{0}_{d+d_2-2}^T \pmod N$ since H is invertible modulo N. It follows that $\gcd(r_1, N) = 1$ since $\Delta(\boldsymbol{r}U^{-T}) = \Delta(\boldsymbol{r}) = 1$. Hence,

$$\boldsymbol{c}_{d_2}U^{-T} \equiv \pm\frac{\Delta(\mathrm{S}_{d_2})}{r_1}\boldsymbol{r}U^{-T} \pmod N.$$

Multiplying both sides of this equation on the right by U^T shows that $\boldsymbol{c}_{d_2}$ is a geometric progression with ratio r modulo N and $\gcd(c_{d_2,0}, N) = 1$.

Suppose that $d_2 < d$ and $\boldsymbol{c}_{t-1}$ for some $t \in \{d_2+1, \ldots, d\}$ is a nonzero geometric progression with ratio r modulo N and $\gcd(c_{t-1,0}, N) = \gcd(\mathrm{lc}(f_1)^{t-1-d_2}, N)$. Let $f_1 = \sum_{i=0}^{d} a_{1,i}x^i$ such that $a_{1,d}, \ldots, a_{1,0} \in \mathbb{Z}$. Then

$$\mathrm{S}_t(f_1, f_2) = \left(\begin{array}{c|cccccc} a_{1,d} & a_{1,d-1} & \ldots & a_{1,0} & 0 & \ldots & 0 \\ \hline 0 & & & & & & \\ \vdots & & & \mathrm{S}_{t-1}(f_1, f_2) & & & \\ 0 & & & & & & \end{array}\right). \tag{4.3}$$

Therefore, deleting the ith column of $\mathrm{S}_t(f_1, f_2)$ and computing the determinant of the resulting matrix along its first row shows that

$$\mathrm{M}_{t,i} = \begin{cases} \sum_{i=1}^{d} a_{1,d-i}\mathrm{M}_{t-1,i} & \text{if } i = 1, \\ -a_{1,d}\mathrm{M}_{t-1,i-1} & \text{if } i \in \{2, \ldots, d+t-1\}. \end{cases} \tag{4.4}$$

It follows that $\boldsymbol{c}_t$ is nonzero since $a_{1,d}$ and $\boldsymbol{c}_{t-1}$ are nonzero. Furthermore, $c_{t,0} = -a_{1,d}c_{t-1,0}$ and thus $\gcd(c_{t,0}, N) = \gcd(\mathrm{lc}(f_1)^{t-d_2}, N)$.

By assumption, $\boldsymbol{c}_{t-1}$ is a geometric progression with ratio r modulo N. Thus, $\mathrm{M}_{t-1,i}r \equiv \mathrm{M}_{t-1,i-1}$ for $i = 2, \ldots, d+t-1$. Therefore, (4.4) implies that

$$\mathrm{M}_{t,i}r \equiv -a_{1,d}\mathrm{M}_{t-1,i-1}r \equiv -a_{1,d}\mathrm{M}_{t-1,i-2} \equiv \mathrm{M}_{t,i-1} \pmod N$$

for $i = 3, \ldots, d+t-1$. The first entry of $\mathrm{S}_{t-1}\boldsymbol{c}_{t-1}^T$ is equal to $\sum_{i=0}^{d} a_{1,d-i}\mathrm{M}_{t-1,i+1}$. Consequently, (4.2) implies that $\sum_{i=0}^{d} a_{1,d-i}\mathrm{M}_{t-1,i+1} = 0$. Thus, (4.4) implies that

$$\mathrm{M}_{t,2}r \equiv -a_{1,d}\mathrm{M}_{t-1,1}r \equiv \sum_{i=1}^{d} a_{1,d-i}\mathrm{M}_{t-1,i+1}r \equiv \sum_{i=1}^{d} a_{1,d-i}\mathrm{M}_{t-1,i} \equiv \mathrm{M}_{t,1} \pmod N.$$

Hence, $\boldsymbol{c}_t$ is a geometric progression with ratio r modulo N. □

Define the *volume* of a real matrix A, denoted $\mathrm{vol}\, A$, to be $\sqrt{\det AA^T}$. For a matrix A over a commutative ring, define $\mathrm{vol}^2 A = \det AA^T$. If A is the $0 \times n$ empty matrix, then $\mathrm{vol}\, A = 1$ and $\mathrm{vol}^2 A = 1$. If A is an $m \times n$ matrix such that $m \le n$, then $\mathrm{vol}\, A$ is nonzero if and only if A has full rank. The volume function is multiplicative in the following sense: if A is an $m \times m$ matrix and B is an $m \times n$ matrix, then $\mathrm{vol}\, AB = \mathrm{vol}\, A \cdot \mathrm{vol}\, B$ and $\mathrm{vol}^2 AB = \mathrm{vol}^2 A \cdot \mathrm{vol}^2 B$.

The following lemma provides the upper bound on $\|\boldsymbol{c}_t\|_{2,s^{-1}}$ in Theorem 4.1. In particular, the lemma shows that if $|\sin\theta_s(f_1,f_2)|\,\|f_1\|_{2,s}\,\|f_2\|_{2,s} = O(N^{1/d})$ for some $s > 0$, i.e., the polynomials are close the attaining the lower bound from Lemma 3.1, then $\|\boldsymbol{c}_d\|_{2,s^{-1}} = O(N^{1-1/d})$.

Lemma 4.3. *The inequality*

$$\|\boldsymbol{c}_t\|_{2,s^{-1}} \leq \left(|\sin\theta_s(f_1,f_2)|\,\|f_1\|_{2,s}\right)^{t-1}\left(s^{(\deg f_2 - t)/2}\,\|f_2\|_{2,s}\right)^{d-1} \tag{4.5}$$

holds for $t = \deg f_2, \dots, d$ *and all* $s > 0$.

Proof. If $\mathrm{S}_{\deg f_2}(f_1,f_2)$ does not have full rank, then the recurrence relation (4.4) implies that $\boldsymbol{c}_t$ is the zero vector for $t = \deg f_2, \dots, d$. Thus, if $\mathrm{S}_{\deg f_2}(f_1,f_2)$ does not have full rank, then the lemma holds trivially. Therefore, assume that $\mathrm{S}_{\deg f_2}(f_1,f_2)$ has full rank. Then the recurrence relation (4.3) implies that $\mathrm{S}_t(f_1,f_2)$ has full rank for $t = \deg f_2, \dots, d$.

Let $t \in \{\deg f_2, \dots, d\}$, s be a positive real number and

$$S = s^{-\frac{d+t-2}{2}} \cdot \operatorname{diag}\left(s^{d+t-2}, s^{d+t-3}, \dots, 1\right).$$

Then the Binet–Cauchy formula (see [1, Section 36]) implies that

$$\operatorname{vol}^2\left(\mathrm{S}_t(f_1,f_2)S\right) = \sum_{i=1}^{d+t-1}\left(\mathrm{M}_{t,i}\frac{\det S}{s^{(d+t-2i)/2}}\right)^2 = \sum_{i=1}^{d+t-1}\frac{\mathrm{M}_{t,i}^2}{s^{d+t-2i}} = \|\boldsymbol{c}_t\|_{2,s^{-1}}^2. \tag{4.6}$$

To complete the proof, Fischer's inequality [11] is used to derive an upper bound on $\operatorname{vol}^2(\mathrm{S}_t(f_1,f_2)S)$ in a manner similar to the proof of [8, Lemma 2.2].

Define matrices $A_1, \dots, A_t$ as follows: A_1 is the $0 \times (d+t-1)$ empty matrix if $t = d$; $A_1 = (x^{d-2}f_2, \dots, x^{t-1}f_2)_{d+t-2}$ if $t \neq d$; and $A_i = (x^{t-i}f_1, x^{t-i}f_2)_{d+t-2}$ for $i = 2, \dots, t$. For $i = 1, \dots, t$, let B_i be the $(d-t+2(i-1)) \times (d+t-1)$ matrix obtained by arranging the matrices $A_1, \dots, A_i$ consecutively beneath each other. Then B_t is obtained from S_t by permuting its rows. Thus, $\operatorname{vol}^2(B_tS) = \operatorname{vol}^2(\mathrm{S}_t(f_1,f_2)S)$. Moreover, as $\mathrm{S}_t(f_1,f_2)$ has full rank and $s \neq 0$, $(A_iS)(A_iS)^T$ and $(B_iS)(B_iS)^T$ are positive definite Hermitian for $i = 1, \dots, t$. Therefore, Fischer's inequality implies that $\operatorname{vol}^2(B_iS) \leq \operatorname{vol}^2(A_iS) \cdot \operatorname{vol}^2(B_{i-1}S)$ for $i = 2, \dots, t$. Hence,

$$\operatorname{vol}^2(B_tS) \leq \operatorname{vol}^2(A_1S) \cdot \operatorname{vol}^2(A_2S) \cdots \operatorname{vol}^2(A_tS). \tag{4.7}$$

For $i = 2, \dots, d$,

$$\left\|(x^{d-i}f_2)_{d+t-2}S\right\|_2 = s^{\frac{d+\deg f_2 - t}{2} - (i-1)}\,\|f_2\|_{2,s}.$$

Thus, if $t \neq d$, then Hadamard's determinant theorem [13] implies that

$$\operatorname{vol}^2(A_1S) \leq \prod_{i=2}^{d-t+1} s^{d-t-\deg f_2 - 2(i-1)}\,\|f_2\|_{2,s}^2 = \left(s^{(\deg f_2 - 1)}\,\|f_2\|_{2,s}^2\right)^{d-t}. \tag{4.8}$$

This inequality also holds trivially if $t = d$. The angle between the row vectors of A_iS is $\theta_s(f_1,f_2)$ and $\left\|(x^{t-i}f_1)_{d+t-2}S\right\|_2 = s^{(t/2)-(i-1)}\,\|f_1\|_{2,s}$ for $i = 2, \dots, t$. Therefore, by computing $(A_iS)^T(A_iS)$ or by viewing $\operatorname{vol} A_iS$ as the area of the parallelogram generated by the row vectors of A_iS, it follows that

$$\operatorname{vol}(A_iS) = s^{t - \frac{d-\deg f_2}{2} - 2(i-1)}\,\|f_1\|_{2,s}\,\|f_2\|_{2,s}\,|\sin\theta_s(f_1,f_2)| \quad \text{for } i = 2, \dots, t.$$

Combining this equation with (4.7) and (4.8) yields the inequality

$$\mathrm{vol}^2(B_t S) \leq \left(|\sin\theta_s(f_1,f_2)|\,\|f_1\|_{2,s}\right)^{2(t-1)} \left(s^{(\deg f_2 - t)/2}\,\|f_2\|_{2,s}\right)^{2(d-1)}.$$

Combining this inequality with (4.6) (recalling that $\mathrm{vol}^2(B_t S) = \mathrm{vol}^2(\mathrm{S}_t S)$) and computing roots yields (4.5). Hence, as t and s were chosen arbitrarily, it follows that (4.5) holds for $t = \deg f_2, \ldots, d$ and all $s > 0$. □

Remark 4.4. The last equality of (4.6) holds for $t = \deg f_2, \ldots, d$ and all $s > 0$. Thus, the inequalities

$$\|\boldsymbol{c}_t\|_{2,s^{-1}}^2 \geq \Delta(\mathrm{S}_t(f_1,f_2))^2 \left(s^{d+t-2} + s^{d+t-4} + \ldots + s^{-(d+t-2)}\right) \geq \Delta(\mathrm{S}_t(f_1,f_2))^2$$

hold for $t = \deg f_2, \ldots, d$ and all $s > 0$.

Property (4) of Theorem 4.1 is now proved. Combining this property with Property (1) of theorem shows that the coefficient vectors of each pair of nonlinear number field sieve polynomials appear in the kernel of a matrix of the form (2.1) for some nonzero length $2d-1$ geometric progression $[c_{2d-2}, \ldots, c_0]$, where d is the maximum degree of the polynomials.

Lemma 4.5. *The vectors $(f_1)_d^T$ and $(f_2)_d^T$ belong to the kernel of ∂C_t for $t = \deg f_2, \ldots, d$.*

Proof. The ith entry of $\partial C_t \cdot (f_1)_d^T$ is equal to the ith entry of $\mathrm{S}_t(f_1,f_2)\cdot \boldsymbol{c}_t^T$ for $i = 1, \ldots, t-1$. Similarly, the ith entry of $\partial C_t \cdot (f_2)_d^T$ is equal to the $(d+i-1)$th entry of $\mathrm{S}_t(f_1,f_2)\cdot \boldsymbol{c}_t^T$ for $i = 1, \ldots, t-1$. Thus, (4.2) implies that $(f_1)_d^T$ and $(f_2)_d^T$ are in the kernel of ∂C_t for $t = \deg f_2, \ldots, d$. □

Denote by $\mathrm{adj}\, A$ the adjoint (or adjugate) of a square matrix A. The following lemma, from which the remaining properties of Theorem 4.1 are deduced, shows that the adjoint of the Bezout matrix has entries which, up to sign, are minors of the Sylvester matrix:

Lemma 4.6. *Let $g_1 = \sum_{i=0}^d u_i x^i$ and $g_2 = \sum_{i=0}^d v_i x^i$ such that $u_0, \ldots, u_d$ and $v_0, \ldots, v_d$ are algebraically independent indeterminates over $\mathbb{Z}$. Then*

$$\mathrm{adj}\,\mathrm{Bez}(g_1,g_2) = (-1)^{d(d-1)/2}\left(\mathrm{M}_{d,i+j-1}(g_1,g_2)\right)_{i=1,\ldots,d;j=1,\ldots,d}. \tag{4.9}$$

Proof. Let $g_1 = \sum_{i=0}^d u_i x^i$ and $g_2 = \sum_{i=0}^d v_i x^i$ such that $u_0, \ldots, u_d$ and $v_0, \ldots, v_d$ are algebraically independent indeterminates over $\mathbb{Z}$. Then the algebraic independence of the coefficients implies that $\mathrm{Res}(g_1,g_2) \in \mathbb{Z}[u_0, \ldots, u_d, v_0, \ldots, v_d]$ is nonzero. Therefore, equation (3.2) implies that it is sufficient to show that

$$\mathrm{Bez}(g_1,g_2)\cdot\left(\mathrm{M}_{d,i+j-1}(g_1,g_2)\right)_{\substack{i=1,\ldots,d\\ j=1,\ldots,d}} = (-1)^d\,\mathrm{Res}(g_1,g_2)\cdot\mathrm{Id}_d, \tag{4.10}$$

where Id_n denotes the $n \times n$ identity matrix for all integers $n \geq 1$.

Following (3.1), define

$$p_{i+1} = \left(\sum_{k=0}^{i} v_{d-i+k}x^k\right) g_1 - \left(\sum_{k=0}^{i} u_{d-i+k}x^k\right) g_2 \quad \text{for } i = 0, \ldots, d-1. \tag{4.11}$$

Then $\mathrm{Bez}(g_1,g_2) = (p_1, \ldots, p_d)$. Define

$$\mathrm{H}_{i,j} = \left(x^{d-j}p_i, x^{d-2}g_1, \ldots, g_1, x^{d-2}g_2, \ldots, g_2\right) \quad \text{for } 1 \leq i,j \leq d. \tag{4.12}$$

The matrices $\mathrm{H}_{i,j}$ are square of order $2d-1$ since $\deg g_i = d$ for $i = 1, 2$, and

$$p_i = (u_{d-i}v_d - v_{d-i}u_d)x^{d-1} + \text{lower order terms} \quad \text{for } i = 1, \ldots, d.$$

Expanding the determinant of each matrix $\mathrm{H}_{i,j}$ along its first row shows that

$$\mathrm{Bez}(g_1, g_2) \cdot (\mathrm{M}_{d,i+j-1}(g_1, g_2))_{\substack{i=1,\ldots,d \\ j=1,\ldots,d}} = (\det \mathrm{H}_{i,j})_{\substack{i=1,\ldots,d \\ j=1,\ldots,d}}.$$

It follows from (4.11) that

$$x^{d-j}p_i = \left(\sum_{k=d-j}^{d+i-j-1} v_{k-i+j+1}x^k \right) g_1 - \left(\sum_{k=d-j}^{d+i-j-1} u_{k-i+j+1}x^k \right) g_2 \tag{4.13}$$

for $1 \leq i, j \leq d$. Therefore, for indices i and j such that $1 \leq i < j \leq d$, the determinant $\det \mathrm{H}_{i,j}$ is zero since the polynomial $x^{d-j}p_i$ is a linear combination of the polynomials $x^{d-2}g_1, \ldots, g_1$ and $x^{d-2}g_2, \ldots, g_2$. Similarly, for $1 \leq j < i \leq d$,

$$\begin{aligned} x^{d-j}p_i &= x^{i-j-1}\left(g_2 - \sum_{k=0}^{d-i} v_k x^k \right) g_1 - x^{i-j-1}\left(g_1 - \sum_{k=0}^{d-i} u_k x^k \right) g_2 \\ &= -\left(\sum_{k=i-j-1}^{d-j-1} v_{k-i+j+1}x^k \right) g_1 + \left(\sum_{k=i-j-1}^{d-j-1} u_{k-i+j+1}x^k \right) g_2, \end{aligned}$$

Thus, $\det \mathrm{H}_{i,j} = 0$ for $1 \leq j < i \leq d$. Consequently,

$$\mathrm{Bez}(g_1, g_2) \cdot (\mathrm{M}_{d,i+j-1}(g_1, g_2))_{\substack{i=1,\ldots,d \\ j=1,\ldots,d}} = \mathrm{diag}\left(\det \mathrm{H}_{1,1}, \ldots, \det \mathrm{H}_{d,d}\right). \tag{4.14}$$

It is now shown that

$$\det \mathrm{H}_{i,i} = (-1)^d \operatorname{Res}(g_1, g_2) \quad \text{for } i = 1, \ldots, d. \tag{4.15}$$

The special case $i = d$ has been proved, up to sign, by Sederberg, Goldman and Du [29, Proposition 2.3]. Their arguments are modified to obtain (4.15). In particular, the proof proceeds by computing the determinants of the following matrices two ways:

$$\overline{\mathrm{H}}_i = \left(x^{d-i}p_i, x^{d-2}g_1, \ldots, g_1, x^{d-1}g_2, \ldots, g_2\right) \quad \text{for } i = 1, \ldots, d. \tag{4.16}$$

Substituting $i = j$ into (4.13) shows that

$$x^{d-i}p_i = \left(\sum_{k=d-i}^{d-1} v_{k+1}x^k \right) g_1 - \left(\sum_{k=d-i}^{d-1} u_{k+1}x^k \right) g_2 \quad \text{for } i = 1, \ldots, d.$$

Therefore, by performing elementary row operations on $\overline{\mathrm{H}}_i$, it follows that

$$\det \overline{\mathrm{H}}_i = \det\left(v_d x^{d-1}g_1, x^{d-2}g_1, \ldots, g_1, x^{d-1}g_2, \ldots, g_2\right) = v_d \operatorname{Res}(g_1, g_2) \tag{4.17}$$

for $i = 1, \ldots, d$. The first column vector of $\overline{\mathrm{H}}_i$ contains v_d in the $(d+1)$th coordinate and zeros elsewhere. Furthermore, the submatrix of $\overline{\mathrm{H}}_i$ obtained by deleting its first column and $(d+1)$th row is equal to $\mathrm{H}_{i,i}$. Therefore, expanding the determinant of $\overline{\mathrm{H}}_i$ along its first column shows that

$$\det \overline{\mathrm{H}}_i = (-1)^d v_d \det \mathrm{H}_{i,i} \quad \text{for } i = 1, \ldots, d. \tag{4.18}$$

Hence, combining (4.17) and (4.18) implies that (4.15) holds. Then combining (4.14) and (4.15) implies that (4.10) holds. □

Property (3) of Theorem 4.1 is now deduced from Lemma 4.6 by specialising the coefficients of the generic polynomials g_1 and g_2:

Corollary 4.7. *The adjoint of* $\operatorname{Bez}(f_1, f_2)$ *is*

$$\operatorname{adj}\operatorname{Bez}(f_1, f_2) = (-1)^{d(d-1)/2} C_d. \tag{4.19}$$

Consequently, if f_1 *and* f_2 *are coprime, then* C_t *has full rank for* $t = \deg f_2, \ldots, d$.

Proof. Let $\mathbb{A} = \mathbb{Z}[u_0, \ldots, u_d, v_0, \ldots, v_d]$ such that $u_0, \ldots, u_d$ and $v_0, \ldots, v_d$ are algebraically independent indeterminates over $\mathbb{Z}$. Set $g_1 = \sum_{i=0}^{d} u_i x^i \in \mathbb{A}[x]$ and $g_2 = \sum_{i=0}^{d} v_i x^i \in \mathbb{A}[x]$. Then (4.9) holds. Define the evaluation homomorphism $\varphi : \mathbb{A} \to \mathbb{Z}$ by $u_i \mapsto a_{1,i}$ and $v_i \mapsto a_{2,i}$ for $i = 0, \ldots, d$. Extend φ entry-wise to matrices. As $\deg f_1 = d$, it holds that $\varphi(\operatorname{adj}\operatorname{Bez}(g_1, g_2)) = \operatorname{adj}\operatorname{Bez}(f_1, f_2)$ and $\varphi(\mathrm{M}_{d,t}(g_1, g_2)) = \mathrm{M}_{d,t}(f_1, f_2)$ for $i = 1, \ldots, 2d-1$. Therefore, the φ-image of the each side of (4.9) is equal to its respective side of (4.19).

Suppose now that f_1 and f_2 are coprime. Then (3.2) implies that $\operatorname{Bez}(f_1, f_2)$ is nonsingular. Thus, (4.19) implies that C_d is nonsingular. If $t \in \mathbb{Z}$ such that $\deg f_2 \le t \le d$, then the recurrence relation (4.3) implies that the submatrix of C_d formed by its last t rows is equal to $(-\operatorname{lc}(f_1))^{d-t} C_t$. As $\operatorname{lc}(f_1)$ is nonzero, it follows that C_t has full rank for $t = \deg f_2, \ldots, d$. □

All that remains in the proof of Theorem 4.1 is to establish the lower bound on $\|\boldsymbol{c}_t\|_{2,s^{-1}}$ stated in Property (5) of the theorem. The remainder of this section is dedicated to the proof of this property, which proceeds as follows: first, the volume of $(\partial^k C_t)S$, where S is an arbitrary nonsingular matrix, is computed; then, for an appropriate choice of S, the volume of $(\partial^k C_t)S$ is bounded above by a power of $\|\boldsymbol{c}_t\|_{2,s^{-1}}$, providing a lower bound on $\|\boldsymbol{c}_t\|_{2,s^{-1}}$; and finally, by examining a special case of this bound, the lower bound stated in Property (5) is proved.

Let A be an $m \times n$ matrix. For all subsets $I \subseteq \{1, \ldots, m\}$ and $J \subseteq \{1, \ldots, n\}$, define $A_{I,J}$ to be the $|I| \times |J|$ submatrix of A formed by the intersection of the rows that have indices in I with the columns that have indices in J. If $m = n$, and $\{I, I'\}$ and $\{J, J'\}$ are partitions of $\{1, \ldots, n\}$ such that $|I| = |J|$, then Jacobi (see [1, Section 42] or [5]) showed that

$$\det(\operatorname{adj} A)_{I,J} = (-1)^{\sum_{i' \in I'} i' + \sum_{j' \in J'} j'} (\det A)^{|I|-1} \det\left(A^T\right)_{I',J'}. \tag{4.20}$$

The following technical lemma is proved by repeatedly applying this identity:

Lemma 4.8. *Suppose that* A *and* S *are* $n \times n$ *matrices such that* $n \ge 2$ *and* S *is invertible. Then, for each partition* $\{I, I'\}$ *of* $\{1, \ldots, n\}$,

$$\operatorname{vol}^2\left((\operatorname{adj} A)_{I,\{1,\ldots,n\}} S\right) = (\det A)^{2(|I|-1)} (\det S)^2 \operatorname{vol}^2\left(\left(A^T\right)_{I',\{1,\ldots,n\}} S^{-T}\right).$$

Proof. Suppose that A and S are $n \times n$ matrices such that $n \ge 2$ and S is invertible. Let $\{I, I'\}$ be a partition of $\{1, \ldots, n\}$. Set $B = \operatorname{adj}(S)A$, $\mathcal{J} = \{J \subseteq \{1, \ldots, n\} \mid |J| = |I|\}$ and $\mathcal{J}' = \{\{1, \ldots, n\} \setminus J \mid J \in \mathcal{J}\}$. Then the Binet–Cauchy formula implies that

$$\operatorname{vol}^2(\operatorname{adj} B)_{I,\{1,\ldots,n\}} = \sum_{J \in \mathcal{J}} \left(\det(\operatorname{adj} B)_{I,J}\right)^2.$$

Using (4.20) to compute each summand on the right hand side shows that

$$\operatorname{vol}^2(\operatorname{adj} B)_{I,\{1,\ldots,n\}} = (\det B)^{2(|I|-1)} \sum_{J' \in \mathcal{J}'} \left(\det\left(B^T\right)_{I',J'}\right)^2.$$

Using the Binet–Cauchy formula to compute the sum on the right hand side yields

$$\operatorname{vol}^2 (\operatorname{adj} B)_{I,\{1,\dots,n\}} = (\det B)^{2(|I|-1)} \operatorname{vol}^2 \left(B^T\right)_{I',\{1,\dots,n\}}. \tag{4.21}$$

If X and Y are $n \times n$ matrices, then $(XY)_{K,\{1,\dots,n\}} = X_{K,\{1,\dots,n\}}Y$ for all $K \subseteq \{1,\dots,n\}$. It follows that

$$\operatorname{vol}^2 (\operatorname{adj} B)_{I,\{1,\dots,n\}} = (\det S)^{2(n-2)|I|} \operatorname{vol}^2 \left((\operatorname{adj} A)_{I,\{1,\dots,n\}} S\right)$$

and

$$\operatorname{vol}^2 \left(B^T\right)_{I',\{1,\dots,n\}} = (\det S)^{2(n-|I|)} \operatorname{vol}^2 \left(\left(A^T\right)_{I',\{1,\dots,n\}} S^{-T}\right).$$

Substituting these values and $\det B = (\det S)^{n-1} \det A$ into (4.21) completes the proof. □

Lemma 4.9. *Let $t, k \in \mathbb{Z}$ such that $\deg f_2 \leq t \leq d$ and $0 \leq k < t$, and S be a real nonsingular $(d+k) \times (d+k)$ matrix. Then*

$$\begin{aligned}\operatorname{vol}\left(\left(\partial^k C_t\right) S\right) = |\det S| \left|\operatorname{lc}(f_1)^{t-\deg f_2} \operatorname{Res}(f_1, f_2)\right|^{t-k-1} \\ \cdot \operatorname{vol}\left(\big(\underbrace{x^{k-1}f_1, \dots, f_1}_{k \text{ terms}}, \underbrace{x^{d-t+k-1}f_2, \dots, f_2}_{d-t+k \text{ terms}}\big)_{d+k-1} S^{-T}\right).\end{aligned} \tag{4.22}$$

Proof. Let $t, k \in \mathbb{Z}$ such that $\deg f_2 \leq t \leq d$ and $0 \leq k < t$, and S be a real nonsingular $(d+k) \times (d+k)$ matrix. Let $\mathbb{A} = \mathbb{R}[u_0, \dots, u_d, v_0, \dots, v_d]$ where $u_0, \dots, u_d$ and $v_0, \dots, v_d$ are algebraically independent indeterminates over $\mathbb{R}$. Define $g_1 = \sum_{i=0}^{d} u_i x^i$, $g_2 = \sum_{i=0}^{d} v_i x^i$ and $p_1, \dots, p_d \in \mathbb{A}[x]$ by (4.11). Then $\operatorname{Bez}(g_1, g_2) = (p_1, \dots, p_d)$. Define a $k \times (d+k)$ matrix G and a $(d+k) \times (d+k)$ matrix B as follows:

$$G = \left(x^{k-1}g_1, x^{k-2}g_1, \dots, g_1\right)_{d+k-1} \quad \text{and} \quad B = \left(\begin{array}{c|c} G^T & \begin{matrix} 0_{k\times d} \\ \operatorname{Bez}(g_1, g_2) \end{matrix} \end{array}\right),$$

where $0_{m\times n}$ denotes the $m \times n$ matrix of zeros for all integers $m, n \geq 0$. The upper $k \times k$ submatrix of G^T is lower triangular, with each entry on its diagonal equal to u_d. Thus, (3.2) implies that $\det B = u_d^k (-1)^{d(d+1)/2} \operatorname{Res}(g_1, g_2) \in \mathbb{A}$, which is nonzero since $u_0, \dots, u_d$ and $v_0, \dots, v_d$ are algebraically independent over $\mathbb{R}$.

If $k \geq 1$, then (4.2) implies that

$$G \cdot \left(\operatorname{M}_{d,i+j-1}(g_1, g_2)\right)_{j=1,\dots,d+k}^{T} = \operatorname{S}_d(g_1, g_2)_{\{i,\dots,i+k-1\},\{1,\dots,2d-1\}} \cdot \boldsymbol{c}_d(g_1, g_2)^T = \mathbf{0}_k$$

for $i = 1, \dots, d-k$. Consequently, Lemma 4.6 implies that

$$(-1)^{d(d-1)/2} u_d^k \left(\operatorname{M}_{d,i+j-1}(g_1, g_2)\right)_{\substack{i=1,\dots,d-k \\ j=1,\dots,d+k}} \cdot B = \begin{pmatrix} 0_{(d-k)\times 2k} & \det B \cdot \operatorname{Id}_{d-k} \end{pmatrix}.$$

As B is nonsingular, it follows that

$$(\operatorname{adj} B)_{\{2k+1,\dots,d+k\},\{1,\dots,d+k\}} = (-1)^{d(d-1)/2} u_d^k \left(\operatorname{M}_{d,i+j-1}(g_1, g_2)\right)_{\substack{i=1,\dots,d-k \\ j=1,\dots,d+k}}.$$

Therefore, on the one hand,

$$\begin{aligned}\operatorname{vol}^2 \left((\operatorname{adj} B)_{\{d-t+2k+1,\dots,d+k\},\{1,\dots,d+k\}} S\right) \\ = u_d^{2k(t-k)} \operatorname{vol}^2 \left(\left(\operatorname{M}_{d,d-t+i+j-1}(g_1, g_2)\right)_{\substack{i=1,\dots,t-k \\ j=1,\dots,d+k}} S\right).\end{aligned} \tag{4.23}$$

On the other hand, as $\mathrm{Bez}(g_1, g_2)$ is symmetric (which is deduced from Lemma 4.6 by noting that $\operatorname{adj}\mathrm{Bez}(g_1, g_2)$ is symmetric), Lemma 4.8 implies that

$$\begin{aligned}(4.24)\quad \operatorname{vol}^2 \left((\operatorname{adj} B)_{\{d-t+2k+1,\dots,d+k\},\{1,\dots,d+k\}}\, S \right) \\ = \left(u_d^k \operatorname{Res}(g_1, g_2)\right)^{2(t-k-1)} (\det S)^2 \\ \cdot \operatorname{vol}^2 \left(\left(x^{k-1} g_1, \dots, g_1, p_1, \dots, p_{d-t+k}\right)_{d+k-1} S^{-T} \right).\end{aligned}$$

Write $f_1 = \sum_{i=0}^{d} a_{1,i} x^i$ and $f_2 = \sum_{i=0}^{d} a_{2,i} x^i$ such that the coefficients $a_{i,j}$ are integers. Define the evaluation homomorphism $\varphi : \mathbb{A} \to \mathbb{R}$ by $u_i \mapsto a_{1,i}$ and $v_i \mapsto a_{2,i}$ for $i = 0, \dots, d$. Then $\varphi(\operatorname{Res}(g_1, g_2)) = a_{1,d}^{d-\deg f_2} \operatorname{Res}(f_1, f_2)$. Extend φ entry-wise to matrices and let $\tilde{\varphi} : \mathbb{A}[x] \to \mathbb{R}[x]$ be the natural extension of φ. Then

$$\begin{aligned}\varphi \left(\left(\mathrm{M}_{d,d-t+i+j-1}(g_1, g_2)\right)_{\substack{i=1,\dots,t-k \\ j=1,\dots,d+k}} \right) &= \left(\mathrm{M}_{d,d-t+i+j-1}(f_1, f_2)\right)_{\substack{i=1,\dots,t-k \\ j=1,\dots,d+k}} \\ &= (-a_{1,d})^{d-t} \cdot \partial^k C_t,\end{aligned}$$

where the final equality follows from the recurrence relation (4.4). Therefore, computing the φ-images of (4.23) and (4.24) and equating shows that

$$\begin{aligned}(4.25)\quad \operatorname{vol}^2 \left(\left(\partial^k C_t\right) S \right) = (\det S)^2 \left(a_{1,d}^{t-\deg f_2} \operatorname{Res}(f_1, f_2) \right)^{2(t-k-1)} a_{1,d}^{-2(d-t+k)} \\ \cdot \operatorname{vol}^2 \left(\left(x^{k-1} f_1, \dots, f_1, \tilde{\varphi}(p_1), \dots, \tilde{\varphi}(p_{d-t+k})\right)_{d+k-1} S^{-T} \right).\end{aligned}$$

From the definition of $p_1, \dots, p_d$, it follows that if $\deg f_2 < d$, then

$$\tilde{\varphi}(p_i) = -f_2 \cdot \sum_{j=0}^{i-1} a_{1,d-i+1+j} x^j \quad \text{for } i = 1, \dots, d - \deg f_2.$$

Furthermore, if $d - t + k > d - \deg f_2$, then

$$\tilde{\varphi}(p_{d-\deg f_2+i}) = \left(\sum_{j=0}^{i-1} a_{2,\deg f_2 - i+1+j} x^j \right) f_1 - \left(\sum_{j=0}^{d-\deg f_2+i-1} a_{1,\deg f_2 - i+1+j} x^j \right) f_2$$

for $i = 1, \dots, \deg f_2 - t + k$. As $\deg f_2 - t + k - 1 \le k - 1$, it follows that

$$\begin{aligned}&\operatorname{vol}^2 \left(\left(x^{k-1} f_1, \dots, f_1, \tilde{\varphi}(p_1), \dots, \tilde{\varphi}(p_{d-t+k})\right)_{d+k-1} S^{-T} \right) \\ &\quad = a_{1,d}^{2(d-t+k)} \cdot \operatorname{vol}^2 \left(\left(x^{k-1} f_1, \dots, f_1, x^{d-t+k-1} f_2, \dots, f_2\right)_{d+k-1} S^{-T} \right).\end{aligned}$$

Substituting this equation into (4.25) and taking roots yields (4.22). □

Lemma 4.10. *Let $t, k \in \mathbb{Z}$ such that $\deg f_2 \le t \le d$ and $0 \le k < t$. Then*

$$\begin{aligned}(4.26)\quad \|\boldsymbol{c}_t\|_{2,s^{-1}}^{t-k} \ge s^{-\frac{t(d-t+k)+dk}{2}} \left| \operatorname{lc}(f_1)^{t-\deg f_2} \operatorname{Res}(f_1, f_2) \right|^{t-k-1} \\ \cdot \operatorname{vol} \left(x^{k-1} f_1(sx), \dots, f_1(sx), x^{d-t+k-1} f_2(sx), \dots, f_2(sx) \right)_{d+k-1}\end{aligned}$$

for all $s > 0$.

Proof. Let $t, k \in \mathbb{Z}$ such that $\deg f_2 \le t \le \deg f_1$ and $0 \le k < t$. For a real number $s > 0$, define

$$S_1 = s^{-\frac{d+t-2}{2}} \operatorname{diag}\left(1, s, \dots, s^{t-k-1}\right), \quad S_2 = \operatorname{diag}\left(1, s, \dots, s^{d+k-1}\right)$$

and

$$S_3 = \operatorname{diag}\big(\underbrace{s^{-d}, \ldots, s^{-(d+k-1)}}_{k \text{ terms}}, s^{-t}, \ldots, s^{-(d+k-1)}\big).$$

Then

$$\begin{aligned}&\left(x^{k-1} f_1, \ldots, f_1, x^{d-t+k-1} f_2, \ldots, f_2\right)_{d+k-1} S_2^{-T} \\ &\qquad = S_3 \left(x^{k-1} f_1(sx), \ldots, f_1(sx), x^{d-t+k-1} f_2(sx), \ldots, f_2(sx)\right)_{d+k-1}.\end{aligned}$$

Thus, Lemma 4.9 with $S = S_2$ implies that

$$\begin{aligned}(4.27)\quad &\operatorname{vol}\left(S_1 \left(\partial^k C_t\right) S_2\right) \\ &\qquad = |\det S_1|\,|\det S_2|\,|\det S_3|\,\left|\operatorname{lc}(f_1)^{t-\deg f_2} \operatorname{Res}(f_1, f_2)\right|^{t-k-1} \\ &\qquad\quad \cdot \operatorname{vol}\left(x^{k-1} f_1(sx), \ldots, f_1(sx), x^{d-t+k-1} f_2(sx), \ldots, f_2(sx)\right)_{d+k-1}.\end{aligned}$$

Recall that $\boldsymbol{c}_t = (c_{t,d+t-2}, \ldots, c_{t,0})$ and $C_t = (c_{t,d+t-i-j})_{i=1,\ldots,t;j=1,\ldots,d}$. Thus,

$$S_1 \left(\partial^k C_t\right) S_2 = \left(c_{t,d+t-2-(i+j-2)} s^{(i+j-2)-\frac{d+t-2}{2}}\right)_{i=1,\ldots,t-1;j=1,\ldots,d+1}.$$

Therefore, the row vectors of $S_1 \left(\partial^k C_t\right) S_2$ each have Euclidean length bounded by $\|\boldsymbol{c}_t\|_{2,s^{-1}}$. Consequently, Hadamard's determinant theorem implies that

$$(4.28)\qquad \operatorname{vol}\left(S_1 \left(\partial^k C_t\right) S_2\right) \leq \|\boldsymbol{c}_t\|_{2,s^{-1}}^{t-1}.$$

Calculating the determinants of S_1, S_2 and S_3 yields

$$(4.29)\qquad |\det S_1|\,|\det S_2|\,|\det S_3| = s^{-\frac{t(d-t+k)+dk}{2}}.$$

Combining (4.27), (4.28) and (4.29) gives (4.26), which completes the proof since s was chosen arbitrarily. □

To end the section, two corollaries to Lemma 4.10 are given. The first corollary establishes the lower bound on $\|\boldsymbol{c}_t\|_{2,s^{-1}}$ stated in Property (5) of Theorem 4.1, completing the proof of the theorem. The second corollary is utilised in the next section as part of the analysis of Montgomery's method.

Corollary 4.11. *The inequality*

$$\|\boldsymbol{c}_t\|_{2,s^{-1}} \geq s^{\frac{t-d}{2}} \left|s^{\deg f_2} \operatorname{lc}(f_2)\right|^{\frac{d}{t}-1} \left|\operatorname{lc}(f_1)^{t-\deg f_2} \operatorname{Res}(f_1, f_2)\right|^{1-\frac{1}{t}}$$

holds for $t = \deg f_2, \ldots, d$ *and all* $s > 0$.

Proof. For $t \in \mathbb{Z}$ such that $\deg f_2 \leq t \leq d$ and $s > 0$, the $(d-t) \times (d-t)$ submatrix of $(x^{d-t-1} f_2(sx), \ldots, f_2(sx))_{d-1}$ formed by columns $t - \deg f_2 + 1, \ldots, d - \deg f_2$ is upper triangular with $s^{\deg f_2} \operatorname{lc}(f_2)$ in each diagonal entry. By applying the Binet–Cauchy formula, it follows that $\operatorname{vol}^2(x^{d-t-1} f_2(sx), \ldots, f_2(sx))_{d-1}$ is equal to $(s^{\deg f_2} \operatorname{lc}(f_2))^{2(d-t)}$ plus some sum of squares. Thus, for all $s > 0$,

$$\operatorname{vol}\left(x^{d-t-1} f_2(sx), \ldots, f_2(sx)\right)_{d-1} \geq \left|s^{\deg f_2} \operatorname{lc}(f_2)\right|^{d-t} \quad \text{for } t = \deg f_2, \ldots, d.$$

Substituting these inequalities into (4.26) for $t = \deg f_2, \ldots, d$ and $k = 0$ completes the proof. □

Corollary 4.12. *The inequality*

$$\|\boldsymbol{c}_d\|_{2,s^{-1}}^{d-1} \geq s^{\frac{\deg f_2 - d}{2}} \left|\mathrm{lc}(f_1)^{d-\deg f_2} \operatorname{Res}(f_1, f_2)\right|^{d-2} |\sin\theta_s(f_1, f_2)| \, \|f_1\|_{2,s} \, \|f_2\|_{2,s}$$

holds for all $s > 0$.

Proof. For all $s > 0$, the Euclidean length of the first row vector of the matrix $(f_1(sx), f_2(sx))$ is $s^{d/2} \|f_1\|_{2,s}$, the Euclidean length of the second row vector is $s^{\deg f_2/2} \|f_2\|_{2,s}$, and the angle between the two row vectors is $\theta_s(f_1, f_2)$. Therefore,

$$\operatorname{vol}(f_1(sx), f_2(sx)) = s^{\frac{d+\deg f_2}{2}} \|f_1\|_{2,s} \, \|f_2\|_{2,s} \, |\sin\theta_s(f_1, f_2)| \quad \text{for all } s > 0.$$

Substituting this equation into (4.26) for $t = d$ and $k = 1$ completes the proof. □

5. Analysis of Montgomery's method

Montgomery's method is analysed in this section, providing criteria for the selection of geometric progressions that yield polynomials with optimal coefficient size and optimal resultant. In particular, the goal of this section is to prove the following theorem, which may be viewed as a converse to Theorem 4.1:

Theorem 5.1. *Let* $d \geq 2$ *and* $\boldsymbol{c} = [c_{2d-2}, \ldots, c_0] \in \mathbb{Z}^{2d-1}$ *be a geometric progression with ratio* r *modulo* N *such that* $C = C(c_{2d-2}, \ldots, c_0)$ *is nonsingular and* $\gcd(c_0, \ldots, c_{d-2}, N) = 1$. *Then, for* $f_1, f_2 \in \mathbb{Z}[x]$ *such that* $2 \leq \deg f_2 \leq \deg f_1$ *and* $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ *is a basis for the integer kernel of* ∂C, *the following properties hold:*

(1) $\deg f_1 = d$;
(2) r *is a root of* f_1 *and* f_2 *modulo* N;
(3) f_1 *and* f_2 *are coprime, with*

$$|\operatorname{Res}(f_1, f_2)| = \frac{|\det C|^{d-1}}{\Delta(\partial C)^{\deg f_2} \Delta(\widehat{\partial C})^{d-\deg f_2}} \quad \text{and} \quad \Delta(\mathrm{S}_d(f_1, f_2)) = \Delta(\boldsymbol{c}) \frac{|\det C|^{d-2}}{\Delta(\partial C)^{d-1}};$$

(4) *the inequalities*

$$\left\| \frac{\boldsymbol{c}}{\Delta(\boldsymbol{c})} \right\|_{2,s^{-1}}^{\frac{1}{d-1}} \leq s^{\frac{\deg f_2 - d}{2}} |\sin\theta_s(f_1, f_2)| \, \|f_1\|_{2,s} \, \|f_2\|_{2,s} \leq \frac{1}{N^{d-2}} \left\| \frac{\boldsymbol{c}}{\Delta(\boldsymbol{c})} \right\|_{2,s^{-1}}^{d-1}$$

hold for all $s > 0$.

Property (1) and Property (2) of Theorem 5.1 are proved in Section 2. The two remaining properties of the theorem are proved in the next section.

Recall from Section 2 that in Montgomery's method, the polynomials f_1 and f_2 are chosen such that $\{(f_1(sx))_d^T, (f_2(sx))_d^T\}$, where $d = \max\{\deg f_1, \deg f_2\}$, is a Lagrange-reduced basis for some $s > 0$. It follows that $|\sin\theta_s(f_1, f_2)| \geq \sqrt{3}/2$ for the chosen value of s (see [26, p. 41]). Combining the inequalities from Property (4) shows that any length $2d - 1$ geometric progression $\boldsymbol{c}$ that satisfies the condition of the theorem has norm satisfying $\|\boldsymbol{c}/\Delta(\boldsymbol{c})\|_{2,s^{-1}} \geq N^{1-1/d}$ for all $s > 0$. Thus, Montgomery's method is unforgiving of a poor choice of geometric progression. In particular, the method generates to two degree d polynomials with optimal coefficient size only if a geometric progression of almost minimal size is used.

Property (3) of Theorem 5.1 may aid the selection of parameters for specific geometric progression constructions by allowing parameters to be tuned so that polynomials with resultant equal to a small multiple of N are obtained. Before

completing the proof of Theorem 5.1 in the next section, Property (3) is used to compute the resultant given by two existing geometric progression constructions:

Example 5.2. For $d = 2$, several authors [22, 31, 27, 18, 8] propose using length $2d - 1 = 3$ geometric progressions of the form

$$[c_2, c_1, c_0] = \left[\frac{am^2 - kN}{p}, am, ap\right],$$

where a, k, p and m are nonzero integers such that $\gcd(m, p) = 1$ and $\gcd(a, N) = 1$. Letting $C = C(c_2, c_1, c_0)$, it follows that

$$\det C = -akN \quad \text{and} \quad \Delta(\partial C) = \Delta\left([c_2, c_1, c_0]\right) = \gcd(a, c_2).$$

Let $\tilde{a} = a/\gcd(a, c_2)$ and $\tilde{k} = k/\gcd(a, c_2)$, with the latter being an integer since $am^2 - kN = pc_2$ and $\gcd(a, N) = 1$. Property (3) of Theorem 5.1 implies that if f_1 and f_2 are quadratic polynomials whose coefficient vectors form a basis for the integer kernel of ∂C, then $\operatorname{Res}(f_1, f_2) = \pm\tilde{a}\tilde{k}N$ and $\Delta(\mathrm{S}_2(f_1, f_2)) = 1$.

Example 5.3. For $d = 3$, Koo, Jo and Kwon [18] and the author [8] propose using length $2d - 1 = 5$ geometric progressions of the form

$$[c_4, c_3, c_2, c_1, c_0] = \left[\frac{m(am^3 - kN)}{p^2}, \frac{am^3 - kN}{p}, am^2, amp, ap^2\right],$$

where a, k, p and m are nonzero integers such that $\gcd(m, p) = 1$ and $\gcd(a, N) = 1$. Letting $C = C(c_4, \ldots, c_0)$, it follows that $\det C = -a(kN)^2$,

$$\Delta(\partial C) = \gcd\left(\frac{am^3 - kN}{p^2}, am, ap\right) \cdot |k|N = \gcd\left(a, c_3/p\right) \cdot |k|N$$

and

$$\Delta([c_4, \ldots, c_0]) = \gcd\left(m\frac{am^3 - kN}{p^2}, p\frac{am^3 - kN}{p^2}, a\right) = \gcd(a, c_3/p).$$

Let $\tilde{a} = a/\gcd(a, c_3/p)$ and $\tilde{k} = k/\gcd(a, c_3/p)$. Property (3) of Theorem 5.1 implies that if f_1 and f_2 are cubic polynomials whose coefficient vectors form a basis for the integer kernel of ∂C, then $\operatorname{Res}(f_1, f_2) = \pm\tilde{a}^2\tilde{k}N$ and $\Delta(\mathrm{S}_3(f_1, f_2)) = |\tilde{a}|$.

5.1. **Proof of Theorem 5.1.** Theorem 4.1 shows that each pair of nonlinear number field sieve polynomials with maximum degree $d \geq 2$ appears in the kernel of the matrix $\partial C(c_{2d-2}, \ldots, c_0)$ for some geometric progression $[c_{2d-2}, \ldots, c_0]$. The following lemma shows that such a geometric progression is unique up to scalar multiple, thus allowing results from Section 4.1 to be used in the proof of Theorem 5.1:

Lemma 5.4. *Let f_1 and f_2 be coprime integer polynomials such that $2 \leq \deg f_2 \leq \deg f_1$, and $d = \deg f_1$. Then the vectors $(f_1)_d^T$ and $(f_2)_d^T$ are in the kernel of $\partial C(c_{2d-2}, \ldots, c_0)$ for some vector $\boldsymbol{c} = (c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ if and only if $\boldsymbol{c} = \pm(\Delta(\boldsymbol{c})/\Delta(\boldsymbol{c}_d(f_1, f_2)) \cdot \boldsymbol{c}_d(f_1, f_2)$.*

Proof. Let f_1 and f_2 be coprime integer polynomials such that $2 \leq \deg f_2 \leq \deg f_1$, $d = \deg f_1$ and $\boldsymbol{c}_d = \boldsymbol{c}_d(f_1, f_2)$. Let $\boldsymbol{c} = (c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ and $C = C(c_{2d-2}, \ldots, c_0)$. Then the ith entry of $\partial C \cdot (f_1)_d^T$ is equal to the ith entry of $\mathrm{S}_d(f_1, f_2) \cdot \boldsymbol{c}^T$ for $i = 1, \ldots, d-1$. Similarly, the ith entry of $\partial C \cdot (f_2)_d^T$ is equal to the $(d + i - 1)$th entry of $\mathrm{S}_d(f_1, f_2) \cdot \boldsymbol{c}^T$ for $i = 1, \ldots, d-1$. Thus, $(f_1)_d^T$ and $(f_2)_d^T$ are in the kernel of ∂C if and only if $\boldsymbol{c}^T$ is in the kernel of $\mathrm{S}_d(f_1, f_2)$.

As f_1 and f_2 are coprime, $\mathrm{Syl}(f_1, f_2)$ is nonsingular. The matrix $\mathrm{S}_{\deg f_2}(f_1, f_2)$ is the submatrix of $\mathrm{Syl}(f_1, f_2)$ obtained by first deleting its first and $(\deg f_2 + 1)$th rows, giving a matrix whose first column contains zeros, then deleting the first column of the resulting matrix. Therefore, $\mathrm{S}_{\deg f_2}(f_1, f_2)$ has full rank. Consequently, the recurrence relation (4.3) implies that $\mathrm{S}_d(f_1, f_2)$ has full rank. Lemma 4.2 and (4.2) imply that $\boldsymbol{c}_d^T$ is nonzero and belongs to the kernel of $\mathrm{S}_d(f_1, f_2)$. Thus, $\{\boldsymbol{c}_d^T/\Delta(\boldsymbol{c}_d)\}$ is a basis of the integer kernel of $\mathrm{S}_d(f_1, f_2)$. Hence, $\boldsymbol{c}^T$ is in the integer kernel of $\mathrm{S}_d(f_1, f_2)$ if and only if $\boldsymbol{c} = \pm(\Delta(\boldsymbol{c})/\Delta(\boldsymbol{c}_d)) \cdot \boldsymbol{c}_d$. □

Let $n \geq 2$ and A be an $n \times n$ matrix. For $i_1, i_2, j_1, j_2 \in \{1, \ldots, n\}$ such that $i_1 < i_2$ and $j_1 < j_2$, define $I_k = \{1, \ldots, n\} \setminus \{i_k\}$ and $J_k = \{1, \ldots, n\} \setminus \{j_k\}$ for $k = 1, 2$. Then

$$\det A \cdot \det A_{I_1 \cap I_2, J_1 \cap J_2} = \det \begin{pmatrix} \det A_{I_1, J_1} & \det A_{I_1, J_2} \\ \det A_{I_2, J_1} & \det A_{I_2, J_2} \end{pmatrix}, \tag{5.1}$$

which is known as Sylvester's identity (see [30, Theorem 4.1] for a proof).

The following lemma is obtained from Lemma 3.2 by computing the determinant of the matrix ψ for the case where H and the polynomials f_1 and f_2 are integral:

Lemma 5.5. *Let $d \geq 2$ and $(c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular. If f_1 and f_2 are integer polynomials such that $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis of the integer kernel of ∂C, then*

$$\operatorname{adj} C = \pm\Delta(\partial C) \cdot \mathrm{Bez}(f_1, f_2). \tag{5.2}$$

Proof. Let $d \geq 2$ and $(c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular. Define $(d+1) \times (d+1)$ matrices

$$B_{i,j} = \begin{pmatrix} & & & & 1 & & & \\ & & & & & & & \\ & & & & & 1 & & \\ c_{2d-2} & c_{2d-3} & \cdots & c_{2d-i-1} & \cdots & c_{2d-j-1} & \cdots & c_{d-2} \\ c_{2d-3} & c_{2d-4} & \cdots & c_{2d-i-2} & \cdots & c_{2d-j-2} & \cdots & c_{d-3} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ c_d & c_{d-1} & \cdots & c_{d-i+1} & \cdots & c_{d-j+1} & \cdots & c_0 \end{pmatrix},$$

where the omitted entries are zeros, for $1 \leq i, j \leq d+1$. Define $b_{i,j} = \det B_{i,j}$ for $1 \leq i, j \leq d+1$. Then $\{(-1)^{i+j+1} b_{i,j} \mid 1 \leq i < j \leq d\}$ is the set of $(d-1) \times (d-1)$ minors of ∂C. The matrix ∂C has full rank since C is nonsingular. Therefore, there exist indices k and ℓ such that $k < \ell$ and $b_{k,\ell} \neq 0$. The first and second column vectors of $\operatorname{adj} B_{k,\ell}$ are $(b_{1,\ell}, \ldots, b_{d+1,\ell})^T$ and $(b_{k,1}, \ldots, b_{k,d+1})^T$ respectively. Thus, these two vectors are nonzero and in the kernel of ∂C.

Let

$$B = \begin{pmatrix} b_{k,1} & \cdots & b_{k,d+1} \\ b_{1,\ell} & \cdots & b_{d+1,\ell} \end{pmatrix}.$$

Let j_1 and j_2 be indices such that $1 \leq j_1 < j_2 \leq d$. Define $I_n = \{1, \ldots, d+1\} \setminus \{n\}$ and $J_n = \{1, \ldots, d+1\} \setminus \{j_n\}$ for $n = 1, 2$. Then

$$\det(B_{k,\ell})_{I_1, J_n} = \begin{cases} (-1)^{j_n+1} b_{j_n,\ell} & \text{if } j_n \leq \ell \\ (-1)^{j_n} b_{\ell,j_n} & \text{if } j_n > \ell \end{cases} = (-1)^{j_n+1} b_{j_n,\ell} \quad \text{for } n = 1, 2.$$

Similarly, $\det(B_{k,\ell})_{I_2, J_n} = (-1)^{j_n} b_{k,j_n}$ for $n = 1, 2$. Finally,

$$\det(B_{k,\ell})_{I_1 \cap I_2, J_1 \cap J_2} = (-1)^{j_1+j_2+1} b_{j_1,j_2}.$$

Therefore, identity (5.1) implies that

$$\det \begin{pmatrix} b_{k,j_1} & b_{k,j_2} \\ b_{j_1,\ell} & b_{j_2,\ell} \end{pmatrix} = -b_{k,\ell} b_{j_1,j_2} \quad \text{for } 1 \leq j_1 < j_2 \leq d+1. \tag{5.3}$$

Consequently, $\Delta(B) = \pm\Delta(\partial C) \cdot b_{k,\ell}$, which is nonzero since ∂C has full rank and $b_{k,\ell} \neq 0$. Therefore, if V is a $(d+1) \times (d+1)$ unimodular matrix such that BV is in Hermite normal form, then

$$BV = \begin{pmatrix} 0_{2\times(d-1)} & H \end{pmatrix}$$

for some 2×2 integer matrix H. Performing elementary column operations on B does not change $\Delta(B)$. Thus, $\det H = \pm\Delta(B)$ is nonzero. Hence,

$$H^{-1}B = \begin{pmatrix} 0_{2\times(d-1)} & \mathrm{Id}_2 \end{pmatrix} V^{-1} \tag{5.4}$$

has integer entries and $(\partial C)(H^{-1}B)^T = 0_{2\times 2}$.

Suppose that f_1 and f_2 are integer polynomials such that $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis of the integer kernel of ∂C. Then there exists a 2×2 nonsingular integer matrix U such that

$$U\,(f_1, f_2) = H^{-1}B. \tag{5.5}$$

Therefore, Lemma 3.2 implies that

$$C^{-1} = -\frac{1}{\det \psi} \operatorname{Bez}(f_1, f_2), \tag{5.6}$$

where

$$\psi = \begin{pmatrix} c_{d-1} & \dots & c_0 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} (f_1, f_2)^T = \begin{pmatrix} c_{d-1} & \dots & c_0 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} B^T (HU)^{-T}.$$

Equation (5.4) implies that $\Delta(H^{-1}B) = 1$ since V is unimodular. Furthermore, $\Delta((f_1, f_2)) = 1$ since $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis of the integer kernel of ∂C: if $\Delta((f_1, f_2))$ were greater than one, then $(f_1)_d^T$ and $(f_2)_d^T$ would only generate a proper subgroup of the integer kernel [8, Section 3.1]. Thus, (5.5) implies that U is unimodular. Therefore,

$$\det \psi = \pm\frac{1}{\det H} \sum_{j=1}^{d} c_{d-j} \det \begin{pmatrix} b_{k,j} & b_{k,d+1} \\ b_{j,\ell} & b_{d+1,\ell} \end{pmatrix} = \pm\frac{1}{\Delta(\partial C)} \sum_{j=1}^{d} c_{d-j} b_{j,d+1}. \tag{5.7}$$

Expanding the determinant of C by minors along its last row shows that

$$\sum_{i=1}^{d} c_{d-i} b_{i,d+1} = \det C. \tag{5.8}$$

Hence, (5.6), (5.7) and (5.8) imply that $\operatorname{adj} C = \pm\Delta(\partial C) \cdot \operatorname{Bez}(f_1, f_2)$. □

The first assertion of Property (3) of Theorem 5.1, that f_1 and f_2 are coprime, follows from Lemma 5.2 since C is nonsingular by assumption. The next step in the proof of the theorem is to prove the formulae for $\operatorname{Res}(f_1, f_2)$ and $\Delta(\mathrm{S}_d(f_1, f_2))$.

Lemma 5.6. *Let $d \geq 2$ and $\boldsymbol{c} = (c_{2d-2}, \dots, c_0) \in \mathbb{Z}^{2d-1}$ such that the matrix $C = C(c_{2d-2}, \dots, c_0)$ is nonsingular. If $f_1, f_2 \in \mathbb{Z}[x]$ such that $2 \leq \deg f_2 \leq \deg f_1$ and $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis for the integer kernel of ∂C, then*

$$\operatorname{lc}(f_1)^{d-\deg f_2} \operatorname{Res}(f_1, f_2) = \pm\frac{(\det C)^{d-1}}{\Delta(\partial C)^d} \quad \textit{and} \quad \Delta\left(\mathrm{S}_d(f_1, f_2)\right) = \Delta(\boldsymbol{c}) \frac{|\det C|^{d-2}}{\Delta(\partial C)^{d-1}}.$$

Proof. Let $d \geq 2$ and $\boldsymbol{c} = (c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular. Suppose that $f_1, f_2 \in \mathbb{Z}[x]$ such that $2 \leq \deg f_2 \leq \deg f_1$ and $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis for the integer kernel of ∂C. Then (5.2) holds and computing the determinant of both sides of the equation yields

$$(\det C)^{d-1} = \pm\Delta(\partial C)^d \operatorname{lc}(f_1)^{d-\deg f_2} \operatorname{Res}(f_1, f_2).$$

As C is nonsingular, $\det C$ and $\Delta(\partial C)$ are nonzero. Thus, $\operatorname{Res}(f_1, f_2)$ is nonzero and Lemma 5.4 implies that $\boldsymbol{c} = \pm\left(\Delta(\boldsymbol{c})/\Delta(\boldsymbol{c}_d(f_1, f_2))\right) \cdot \boldsymbol{c}_d(f_1, f_2)$. Therefore, on the one hand, Corollary 4.7 implies that

$$\operatorname{adj}\operatorname{Bez}(f_1, f_2) = \pm\frac{\Delta(\boldsymbol{c}_d(f_1, f_2))}{\Delta(\boldsymbol{c})} C = \pm\frac{\Delta\left(\operatorname{S}_d(f_1, f_2)\right)}{\Delta(\boldsymbol{c})} C$$

On the other hand, computing the adjoint of both side of (5.2) yields

$$(\det C)^{d-2} C = \pm\Delta(\partial C)^{d-1} \operatorname{adj}\operatorname{Bez}(f_1, f_2).$$

As C has at least one nonzero entry, it follows that

$$\pm(\det C)^{d-2} = \Delta(\partial C)^{d-1} \frac{\Delta\left(\operatorname{S}_d(f_1, f_2)\right)}{\Delta(\boldsymbol{c})},$$

where the right hand side is positive. □

The following lemma and its subsequent corollary complete the proof of Property (3) of Theorem 5.1 by showing that $\operatorname{lc}(f_1) = \pm\Delta(\widehat{\partial C})/\Delta(\partial C)$ when the degree of f_2 is strictly less than d:

Lemma 5.7. *Let $d \geq 2$, $(c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular, and integers $b_{i,j}$, for $1 \leq i, j \leq d+1$, be defined as in the proof of Lemma 5.5. Then there exist integers $x_2, \ldots, x_{d+1}$ such that $\sum_{k=2}^{d+1} x_k b_{k,1} = \Delta(\widehat{\partial C})$ and, for any such integers, the set*

$$\left\{ \frac{1}{\Delta(\partial C)} \sum_{k=2}^{d+1} x_k \left(b_{k,1}, b_{k,2}, \ldots, b_{k,d+1}\right)^T, \frac{1}{\Delta(\widehat{\partial C})} \left(b_{1,1}, b_{2,1}, \ldots, b_{d+1,1}\right)^T \right\}$$

is a basis of the integer kernel of ∂C.

Proof. Let $(c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular Then $\Delta(\widehat{\partial C})$ is nonzero since C is nonsingular. Define $b_{i,j}$ as in the proof of Lemma 5.5 for $1 \leq i, j \leq d+1$. For distinct indices i and j, $b_{i,j}$ is up to sign equal to the determinant of the $(d-1) \times (d-1)$ submatrix of ∂C obtained by deleting columns i and j. Thus, $\gcd(b_{2,1}, b_{3,1}, \ldots, b_{d+1,1}) = \Delta(\widehat{\partial C})$ and there exist integers $x_2, \ldots, x_{d+1}$ such that $\sum_{k=2}^{d+1} x_k b_{k,1} = \Delta(\widehat{\partial C})$. For such integers, define

$$\boldsymbol{b}_1 = \frac{1}{\Delta(\partial C)} \sum_{k=2}^{d+1} x_k \left(b_{k,1}, \ldots, b_{k,d+1}\right)^T \quad \text{and} \quad \boldsymbol{b}_2 = \frac{1}{\Delta(\widehat{\partial C})} \left(b_{1,1}, \ldots, b_{d+1,1}\right)^T.$$

Write $\boldsymbol{b}_i = (\beta_{i,1}, \ldots, \beta_{i,d+1})^T$ for $i = 1, 2$, and let $B = (\beta_{i,j})_{i=1,2;j=1,\ldots,d+1}$. As $b_{i,i} = 0$ for $i = 1, \ldots, d+1$, it follows that $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ have integer entries. Moreover, it follows from the proof of Lemma 5.5 that $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ belong to the integer kernel of ∂C. Therefore, $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ is a basis for the integer kernel of ∂C if and only if $\Delta(B) = 1$: $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are linearly independent if and only if $\Delta(B) \neq 0$; and if $\Delta(B) \neq 0$, then $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ is a basis of an index $\Delta(B)$ subgroup of the integer kernel [8, Section 3.1].

For $1 \leq i < j \leq d+1$,

$$\det \begin{pmatrix} \beta_{1,i} & \beta_{1,j} \\ \beta_{2,i} & \beta_{2,j} \end{pmatrix} = \frac{1}{\Delta(\partial C)\Delta(\widehat{\partial C})} \sum_{k=2}^{d+1} x_k \det \begin{pmatrix} b_{k,i} & b_{k,j} \\ b_{i,1} & b_{j,1} \end{pmatrix}.$$

Therefore, (5.3) implies that

$$\det \begin{pmatrix} \beta_{1,i} & \beta_{1,j} \\ \beta_{2,i} & \beta_{2,j} \end{pmatrix} = \frac{1}{\Delta(\partial C)\Delta(\widehat{\partial C})} \sum_{k=2}^{d+1} -x_k b_{k,1} b_{i,j} = -\frac{b_{i,j}}{\Delta(\partial C)}$$

for $1 \leq i < j \leq d+1$. The greatest common divisor of the $b_{i,j}$, for $1 \leq i < j \leq d+1$, is equal to $\Delta(\partial C)$. Hence, $\Delta(B) = 1$. □

Corollary 5.8. *Let $d \geq 2$ and $(c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that the matrix $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular. If $f_1, f_2 \in \mathbb{Z}[x]$ such that $\deg f_2 \leq \deg f_1$ and $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis for the integer kernel of ∂C, then $\Delta(\widehat{\partial C})/\Delta(\partial C)$ divides $\operatorname{lc}(f_1)$. Furthermore, if $\deg f_2 < \deg f_1$, then $\operatorname{lc}(f_1) = \pm\Delta(\widehat{\partial C})/\Delta(\partial C)$.*

Proof. Let $d \geq 2$ and $(c_{2d-2}, \ldots, c_0) \in \mathbb{Z}^{2d-1}$ such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular. Define $\boldsymbol{b}_1 = (\beta_{1,1}, \ldots, \beta_{1,d+1})^T$ and $\boldsymbol{b}_2 = (\beta_{2,1}, \ldots, \beta_{2,d+1})^T$ as in the proof of Lemma 5.7. Then $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ is a basis for the integer kernel of ∂C,

$$\beta_{1,1} = \frac{1}{\Delta(\partial C)} \sum_{k=2}^{d+1} x_k b_{k,1} = \frac{\Delta(\widehat{\partial C})}{\Delta(\partial C)} \neq 0 \quad \text{and} \quad \beta_{2,1} = \frac{b_{1,1}}{\Delta(\widehat{\partial C})} = 0.$$

Suppose that $f_1, f_2 \in \mathbb{Z}[x]$ such that $\deg f_2 \leq \deg f_1$ and $\{(f_1)_d^T, (f_2)_d^T\}$ is a basis of the integer kernel of ∂C. Then there exists a unimodular matrix $U = (u_{i,j})_{i=1,2;j=1,2}$ such that

$$U \cdot (\beta_{i,j})_{i=1,2;j=1,\ldots,d+1} = (f_1, f_2)_d.$$

As C is nonsingular, $\deg f_1 = d$. Therefore, $\operatorname{lc}(f_1) = u_{1,1}\beta_{1,1}$ since $\beta_{2,1} = 0$. If $\deg f_2 < \deg f_1$, then $u_{2,1} = 0$ since $\beta_{1,1} \neq 0$ and $\beta_{2,1} = 0$. Thus, if $\deg f_2 < \deg f_1$, then $u_{1,1} = \pm 1$ since U is unimodular. □

Property (4) of Theorem 5.1 is now proved, completing the proof of the theorem:

Lemma 5.9. *Let $d \geq 2$ and $\boldsymbol{c} = [c_{2d-2}, \ldots, c_0]$ be a geometric progression modulo N such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular and $\gcd(\Delta(\boldsymbol{c}), N) = 1$. If $f_1, f_2 \in \mathbb{Z}[x]$ such that $2 \leq \deg f_2 \leq \deg f_1$ and $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis of the integer kernel of ∂C, then*

$$\|\boldsymbol{c}/\Delta(\boldsymbol{c})\|_{2,s^{-1}} \leq \frac{|\det C|^{d-2}}{\Delta(\partial C)^{d-1}} \|\boldsymbol{c}\|_{2,s^{-1}} \leq \left(s^{\frac{\deg f_2 - d}{2}} |\sin\theta_s(f_1, f_2)| \, \|f_1\|_{2,s} \, \|f_2\|_{2,s}\right)^{d-1}$$

and

$$s^{\frac{\deg f_2 - d}{2}} |\sin\theta_s(f_1, f_2)| \, \|f_1\|_{2,s} \, \|f_2\|_{2,s} \leq \frac{1}{\Delta(\partial C)} \|\boldsymbol{c}\|_{2,s^{-1}}^{d-1} \leq \frac{1}{N^{d-2}} \|\boldsymbol{c}/\Delta(\boldsymbol{c})\|_{2,s^{-1}}^{d-1}$$

for all $s > 0$.

Proof. Let $d \geq 2$ and $\boldsymbol{c} = [c_{2d-2}, \ldots, c_0]$ be a geometric progression modulo N such that $C = C(c_{2d-2}, \ldots, c_0)$ is nonsingular and $\gcd(\Delta(\boldsymbol{c}), N) = 1$. Suppose that $f_1, f_2 \in \mathbb{Z}[x]$ such that $2 \leq \deg f_2 \leq \deg f_1$ and $\left\{(f_1)_d^T, (f_2)_d^T\right\}$ is a basis for the integer kernel of ∂C. Let $\boldsymbol{c}_d = \boldsymbol{c}_d(f_1, f_2)$. Then Lemma 5.6 implies that $\operatorname{Res}(f_1, f_2)$

and $\Delta(\mathrm{S}_d(f_1, f_2))$ are nonzero. Thus, f_1 and f_2 are coprime and Lemma 5.4 implies that $\boldsymbol{c}_d = \pm(\Delta(\mathrm{S}_d(f_1, f_2))/\Delta(\boldsymbol{c})) \cdot \boldsymbol{c}$. Therefore, Lemma 5.6 implies that

$$\|\boldsymbol{c}_d\|_{2,s^{-1}} = \frac{\Delta(\mathrm{S}_d(f_1, f_2))}{\Delta(\boldsymbol{c})} \|\boldsymbol{c}\|_{2,s^{-1}} = \frac{|\det C|^{d-2}}{\Delta(\partial C)^{d-1}} \|\boldsymbol{c}\|_{2,s^{-1}} \quad \text{for all } s > 0. \tag{5.9}$$

As $\Delta(\mathrm{S}_d(f_1, f_2))$ is a nonzero integer, the inequality $\|\boldsymbol{c}_d\|_{2,s^{-1}} \geq \|\boldsymbol{c}/\Delta(\boldsymbol{c})\|_{2,s^{-1}}$ holds for all $s > 0$. Combining this inequality with (5.9) and the upper bound on $\|\boldsymbol{c}_d\|_{2,s^{-1}}$ provided by Lemma 4.3 yields the first set of inequalities stated in the corollary for all $s > 0$.

Corollary 4.12, Lemma 5.6 and (5.9) imply that

$$s^{\frac{\deg f_2 - d}{2}} |\sin\theta_s| \, \|f_1\|_{2,s} \, \|f_2\|_{2,s} \leq \left| \frac{\Delta(\partial C)^d}{(\det C)^{d-1}} \right|^{d-2} \|\boldsymbol{c}_d\|_{2,s^{-1}}^{d-1} = \frac{\|\boldsymbol{c}\|_{2,s^{-1}}^{d-1}}{\Delta(\partial C)}, \tag{5.10}$$

where $\theta_s = \theta_s(f_1, f_2)$, for all $s > 0$. Let r be the ratio of $\boldsymbol{c}$ modulo N. Then subtracting r times row $i+1$ of ∂C from row i for $i = 1, \ldots, d-2$ produces a matrix whose first $d-2$ rows contain multiples of N. As $\Delta(\boldsymbol{c})$ divides each entry of ∂C and $\gcd(\Delta(\boldsymbol{c}), N) = 1$, it follows that $\Delta(\boldsymbol{c})^{d-1} N^{d-2}$ divides $\Delta(\partial C)$. Thus, $\Delta(\boldsymbol{c})^{d-1} N^{d-2} \leq \Delta(\partial C)$ since $\Delta(\partial C) \neq 0$. Combining this inequality with (5.10) completes the proof of the second set of inequalities stated in the corollary. □

In the proof of Lemma 5.9, the assumption that $\boldsymbol{c}$ is a geometric progression modulo N such that $\gcd(\Delta(\boldsymbol{c}), N) = 1$ is only used to prove the last inequality of the lemma. Setting $\boldsymbol{c} = (0, \ldots, 0, 1, 0, \ldots, 0)$ where the 1 appears in the dth coordinate shows that the remaining inequalities cannot be improved by a constant factor without using the assumption. If $\deg f_2 = d$, then (5.9) and Corollary 4.11 imply that

$$\frac{|\det C|^{d-2}}{\Delta(\partial C)^{d-1}} \|\boldsymbol{c}\|_{2,s^{-1}} \geq |\mathrm{Res}(f_1, f_2)|^{1-1/d} \quad \text{for all } s > 0.$$

If the inequality is strict, then Lemma 5.9 improves upon the lower bound on $|\sin\theta_s(f_1, f_2)| \, \|f_1\|_{2,s} \, \|f_2\|_{2,s}$ provided by Lemma 3.1.

Acknowledgements

Part of this work was performed while the author was employed by The University of Queensland, Brisbane, Australia. The author is grateful to Cyril Bouvier for many helpful comments and discussions.

References

1. A. C. Aitken, *Determinants and matrices*, third ed., Oliver and Boyd, Edinburgh, 1944.
2. Shi Bai, Cyril Bouvier, Alexander Kruppa, and Paul Zimmermann, *Better polynomials for GNFS*, Preprint, 12 pages, 2014, `https://hal.inria.fr/hal-01089507`.
3. Shi Bai, Richard P. Brent, and Emmanuel Thomé, *Root optimization of polynomials in the number field sieve*, ArXiv e-Print archive, `arXiv:1212.1958 [math.NT]`, 2012, `http://arxiv.org/abs/1212.1958`.
4. Shi Bai, Emmanuel Thomé, and Paul Zimmermann, *Factorisation of RSA-704 with CADO-NFS*, Cryptology ePrint Archive, Report 2012/369, 2012, `http://eprint.iacr.org/2012/369.pdf`.
5. Richard A. Brualdi and Hans Schneider, *Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, and Cayley*, Linear Algebra Appl. **52/53** (1983), 769–791.

6. E. R. Canfield, Paul Erdős, and Carl Pomerance, *On a problem of Oppenheim concerning "factorisatio numerorum"*, J. Number Theory **17** (1983), no. 1, 1–28.
7. Henri Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics, vol. 138, Springer-Verlag, Berlin, 1993.
8. Nicholas Coxon, *On nonlinear polynomial selection for the number field sieve*, ArXiv e-Print archive, `arXiv:1109.6398 [math.NT]`, 2011, `http://arxiv.org/abs/1109.6398`.
9. Richard Crandall and Carl Pomerance, *Prime numbers: A computational perspective*, second ed., Springer, New York, 2005.
10. M'hammed El Kahoui, *An elementary approach to subresultants theory*, J. Symbolic Comput. **35** (2003), no. 3, 281–292.
11. E Fischer, *Über den hadamardschen determinantensatz*, Arch. Math. Phys. **13** (1908), no. 3, 32–40.
12. Domingo Gomez, Jaime Gutierrez, Álvar Ibeas, and David Sevilla, *Common factors of resultants modulo p*, Bull. Aust. Math. Soc. **79** (2009), no. 2, 299–302.
13. Jacques Hadamard, *Résolution d'une question relative aux déterminants*, Bull. des Sci. Math. **17** (1893), 240–246.
14. Georg Heinig and Karla Rost, *Introduction to Bezoutians*, Numerical methods for structured matrices and applications (D.A. Bini, V. Mehrmann, V. Olshevsky, E. Tyrtsyhnikov, and M. van Barel, eds.), Oper. Theory Adv. Appl., vol. 199, Birkhäuser Verlag, Basel, 2010, pp. 25–118.
15. Thorsten Kleinjung, *On polynomial selection for the general number field sieve*, Math. Comp. **75** (2006), no. 256, 2037–2047.
16. ———, *Polynomial selection*, Slides presented at the CADO workshop on integer factorization, Nancy, France, 2008, `http://cado.gforge.inria.fr/workshop/slides/kleinjung.pdf`.
17. Thorsten Kleinjung, Kazumaro Aoki, Jens Franke, Arjen K. Lenstra, Emmanuel Thomé, Joppe W. Bos, Pierrick Gaudry, Alexander Kruppa, Peter L. Montgomery, Dag Arne Osvik, Herman te Riele, Andrey Timofeev, and Paul Zimmermann, *Factorization of a 768-bit RSA modulus*, Advances in cryptology–CRYPTO 2010 (Tal Rabin, ed.), Lecture Notes in Comput. Sci., vol. 6223, Springer, Berlin, 2010, pp. 333–350.
18. Namhun Koo, Gooc Hwa Jo, and Soonhak Kwon, *On nonlinear polynomial selection and geometric progression (mod N) for number field sieve*, Cryptology ePrint Archive, Report 2011/292, 2011, `http://eprint.iacr.org/2011/292.pdf`.
19. F. I. Lander, *The Bezoutian and the inversion of Hankel and Toeplitz matrices (in Russian)*, Mat. Issled. **9** (1974), no. 2 (32), 69–87, 249–250.
20. A. K. Lenstra and H. W. Lenstra, Jr. (eds.), *The development of the number field sieve*, Lecture Notes in Mathematics, vol. 1554, Springer-Verlag, Berlin, 1993.
21. Bhubaneswar Mishra, *Algorithmic algebra*, Springer-Verlag New York, Inc., New York, 1993.
22. Peter L. Montgomery, *Small geometric progressions modulo n*, Unpublished note of 2 pages, 1993, revised 1995 and 2005.
23. ———, *Searching for higher-degree polynomials for the general number field sieve*, PowerPoint presentation, 34 pages, 2006, `http://www.ipam.ucla.edu/publications/scws1/scws1_6223.ppt`.
24. Brian A. Murphy, *Modelling the yield of number field sieve polynomials*, Algorithmic number theory (Joe P. Buhler, ed.), Lecture Notes in Comput. Sci., vol. 1423, Springer, Berlin, 1998, pp. 137–150.
25. ———, *Polynomial selection for the number field sieve integer factorisation algorithm*, Ph.D. thesis, Australian National University, 1999.
26. Phong Q. Nguyen and Brigitte Vallée (eds.), *The LLL algorithm: Survey and applications*, Information Security and Cryptography, Springer-Verlag, Berlin, 2010.
27. Thomas Prest and Paul Zimmermann, *Non-linear polynomial selection for the number field sieve*, J. Symbolic Comput. **47** (2012), no. 4, 401–409.
28. R. L. Rivest, A. Shamir, and L. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Comm. ACM **21** (1978), no. 2, 120–126.
29. Tom Sederberg, Ron Goldman, and Hang Du, *Implicitizing rational curves by the method of moving algebraic curves*, J. Symbolic Comput. **23** (1997), no. 2-3, 153–175.
30. Eugene Tyrtyshnikov, *Hankel minors and Pade approximations*, Numerical methods for structured matrices and applications (D.A. Bini, V. Mehrmann, V. Olshevsky, E. Tyrtsyhnikov,

and M. van Barel, eds.), Oper. Theory Adv. Appl., vol. 199, Birkhäuser Verlag, Basel, 2010, pp. 431–439.
31. Ronnie S. Williams, Jr., *Cubic polynomials in the number field sieve*, Master's thesis, Texas Tech University, 2010.

INRIA / CNRS / Universite de Lorrainé, Campus Scientifique, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France
E-mail address: nicholas.coxon@inria.fr