When are two numerical polynomials relatively prime?

Bernhard Beckermann
Laboratoire d'Analyse Numérique et d'Optimisation,
Université des Sciences et Technologies de Lille,
59655 Villeneuve d'Ascq Cedex, France
e–mail: bbecker@ano.univ-lille1.fr
and
George Labahn
Department of Computing Science
University of Waterloo, Waterloo, Ontario, Canada
e–mail: glabahn@daisy.uwaterloo.ca

Jan 7, 1998

Abstract

Let a and b be two polynomials having numerical coefficients. We consider the question: when are a and b relatively prime? Since the coefficients of a and b are approximant, the question is the same as: when are two polynomials relatively prime, even after small perturbations of the coefficients?

In this paper we provide a numeric parameter for determining that two polynomials are prime, even under small perturbations of the coefficients. Our methods rely on an inversion formula for Sylvester matrices to establish an effective criterion for relative primeness. The inversion formula can also be used to approximate the condition number of a Sylvester matrix.

1 Introduction

Let $\mathbb{C}[z]$ be the space of polynomials over the complex numbers and let $a, b \in \mathbb{C}[z]$ be polynomials

$$a(z) = a_0 + a_1 z + \ldots + a_m z^m, \quad b(z) = b_0 + b_1 z + \ldots + b_n z^n, \quad a_m, b_n \neq 0$$

of degree m and n, respectively. The GCD of a and b is given by

$$gcd(a,b)(z) = \prod_{\gamma \in A \cap B} (z - \gamma), \quad \text{where} \quad a(z) = a_m \cdot \prod_{\alpha \in A} (z - \alpha), \quad b(z) = b_n \cdot \prod_{\beta \in B} (z - \beta).$$

This is well defined by the Fundamental Theorem of Algebra. We are interested in the question: when are two polynomials, a, b relatively prime, that is, when do a and b have no common roots?

In the case of exact arithmetic determining if two polynomials are relatively prime is well known. This is not the case in the presence of finite precision arithmetic. In this case a computer will not necessarily decide correctly whether two given polynomials with rational coefficients are coprime. For instance, after transforming the coefficients of the polynomials

$$a(z) = (z - \frac{1}{3})(z - \frac{5}{3}) = z^2 - 2z + \frac{5}{9}, \quad b(z) = z - \frac{1}{3}$$

into (decimal) floating point numbers, the resulting polynomials are coprime. Also, the polynomials

$$a(z) = 50z - 7, \quad b(z) = z - \frac{1}{7}$$

are not coprime within a precision of two (decimal) digits.

A more reliable computer answer may be expected for the problem of deciding whether two polynomials remain coprime even after perturbation of coefficients by quantities bounded in norm by some ϵ. This is the type of problem that is of interest in applications such as robotics and control theory [11, 16] where the input data is only known up to some fixed accuracy or where noise is present in the input parameters. In this paper we provide a parameter to determine coprimeness of two numeric polynomials. This parameter is based on quantities which are efficiently obtainable. Indeed in [2] we present an algorithm for computing this parameter that is both numerically stable and at the same time is typically an order of magnitude faster than alternate methods. Because of this efficiency, computing this parameter as an initial test for coprimeness may always be done before starting the more expensive computation of an ϵ–GCD [5, 6, 10, 14, 15].

In fact, we are very much interested in determining some non–trivial numerical ϵ–GCD if the answer to the above question is no. This problem has been treated by several authors each with a different notion of greatest common divisor. These include methods that are based on optimization techniques [5, 10] which are probably numerically stable but quite expensive and others which are more or less based on classical Euclidean concepts [6, 14] but for which one is unable to guarantee numerical stability [2]. Finally we mention the quasi-gcd of Schönhage [15] where the use of an oracle makes it difficult to judge the practical use.

It is well known that the Sylvester matrix of two polynomials plays a vital role in determining the greatest common divisor of two polynomials. The magnitude of the inverse of the Sylvester matrix is important in determining the distance to the closest polynomials having a common root. In our case, we use a new inversion formula for the Sylvester matrix to obtain an estimate of the magnitude of the inverse in terms of only the magnitude of the first and last columns of the inverse. We show that our estimate is better for determining the distance to the closest polynomials having a common root than that provided by the magnitude of the inverse of the Sylvester matrix.

The remainder of the paper organized is as follows. In the next section we place our problem in a linear algebraic setting making use of Sylvester's matrix. Section 3 gives a new inverse formula for Sylvester's matrix while our new "coprime" measure follows in Section 4. Section 5 gives a refinement of our primeness measure. The final sections include some examples and give a conclusion.

Notation: For the remainder of this paper we make use of the following notation: we denote the 1–Hölder vector norm on $\mathbb{C}^n$ as well as the subordinate matrix norm by $||\cdot||$. For $c \in \mathbb{C}[z]$, $c(z) = c_0 + .. + c_n z^n$ we set $\vec{c} = (c_0, .., c_n)^T$ as the vector of coefficients. Our norm on $\mathbb{C}[z]$ is given by

$$||c|| := ||\vec{c}|| = \sum_j |c_j|,$$

and on $\mathbb{C}[z]^{r\times s}$, the space of $r \times s$ matrices with polynomial entries, by

$$||(c_{j,k})|| := ||(||c_{j,k}||)|| = \max_k \sum_j ||c_{j,k}||.$$

Note that for all $c, d \in \mathbb{C}[z]$ we have $||c \cdot d|| \leq ||c|| \cdot ||d||$, and this inequality also holds for polynomial matrices of appropriate size.

With this notation we can restate our problem as follows.

Definition 1.1 *For $a, b \in \mathbb{C}[z]$ let*

$$\epsilon(a, b) := \inf\{||(a - a^*, b - b^*)|| : \ (a^*, b^*) \text{ have a common root, } \deg a^* \leq m, \deg b^* \leq n\},$$

that is, any polynomials a^*, b^* satisfying $||(a-a^*, b-b^*)|| \leq \epsilon < \epsilon(a,b)$ and the above degree restrictions are coprime. We will then refer to a, b as being ϵ–prime.

We are interested in computing approximately sharp "simple" lower bounds for $\epsilon(a,b)$.

2 Inversion of Sylvester's Matrix

It is well-known that the greatest common divisor problem can be placed in a linear algebra setting. This has the advantage that it allows one to make use of concepts from numerical linear algebra (such as condition number) to give information on the numerical gcd problem.

Let $S(a,b)$ denote the Sylvester Matrix for (a,b), that is,

$$S(a,b) = \underbrace{\begin{bmatrix} a_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_m & & \ddots & a_0 \\ 0 & a_m & & a_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_m \end{bmatrix}}_{n} \underbrace{\begin{bmatrix} b_0 & 0 & \cdots & 0 \\ b_1 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ b_n & & \ddots & b_0 \\ 0 & b_n & & b_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_n \end{bmatrix}}_{m} \in \mathbb{C}^{(m+n)\times(m+n)}.$$

Sylvester's criterion from 1853 states that two polynomials are relatively prime if and only if $S(a,b)$ is non–singular (see, e.g., [8]). Numerically it is known that

Lemma 2.1 *For any two polynomials a and b we have*

$$\epsilon(a,b) \geq \frac{1}{||S(a,b)^{-1}||}.$$

Proof: According our choice of matrix norms we have

$$||S(a,b)|| = \max\{\sum_{j=0}^{m} |a_j|, \sum_{j=0}^{m} |b_j|\} = ||(a,b)||.$$

Consequently, using a Theorem of Gastinel [9, Theorem 6.5, p.123] we obtain

$$\begin{aligned} \epsilon(a,b) &= \inf\{||S(a,b) - S(a^*,b^*)|| : S(a^*,b^*) \text{ singular}\} \\ &\geq \min\{||S(a,b) - B|| : \ B \text{ singular}\} = \frac{1}{||S(a,b)^{-1}||}. \end{aligned}$$

□

Remark 2.2 *In the case of the Euclidean norm, we have*

$$j = 1, ..., m+n: \quad \sigma_j = \min\{||S(a,b) - B||_2 : \ defect(B) \geq j\},$$

with $\sigma_1 \leq \sigma_2 \leq ... \leq \sigma_{m+n}$ being the singular values of $S(a,b)$. This allows one to define the ϵ–defect of the Sylvester matrix, which has been chosen by Corless, Gianni, Trager and Watt [5] as the degree of some ϵ–GCD.

Remark 2.3 *From the proof of Lemma 2.1 we see that the quantity* $||(a,b)||/\epsilon(a,b)$ *may be considered as a structured* 1*–condition number of* $S(a,b)$ *in the class of Sylvester matrices (i.e., we consider only perturbations of* $S(a,b)$ *being themselves Sylvester matrices). More generally, the distance to the set of polynomials with GCD having a certain degree (see [5, 6, 10]) may be understood as a structured singular value (with respect to the* 1*–Hölder norm) of a Sylvester matrix*

$$\begin{aligned}\epsilon_j(a,b) &:= \min\{||S(a,b) - S(a^*,b^*)|| : \ \textit{defect}(S(a^*,b^*)) \geq j\} \\ &= \min\{||(a,b) - (a^*,b^*)|| : \textit{degree of GCD of } (a^*,b^*) \textit{ is at least } j \ \}.\end{aligned}$$

Lemma 2.1 states that if we perturb the coefficients of our polynomials by any ϵ less than the reciprocal of the norm of the inverse of the Sylvester matrix then we still have relatively prime polynomials. In fact, a test for coprimeness based on the size of the norm of the inverse of the Sylvester matrix is already included as a special case in the SVD GCD algorithm proposed by Corless, Gianni, Trager and Watt [5, p.198], and in [6, Algorithm 1] of Emiris, Galligo and Lombardi. However, in our case we do not want to estimate the reciprocal of the norm of the inverse by the singular value decomposition of the Sylvester matrix. This decomposition is expensive and does not take advantage of the special structure of a Sylvester matrix. Our goal is to find an easily computable bound that lies between $\epsilon(a,b)$ and the reciprocal of the norm of the inverse. This gives a criterion for numerical coprimeness that is both more precise and also less expensive to compute than previous methods.

Note that $\mathbb{C}[z]$ is a principal ideal domain, so that we have $< a,b >=< gcd(a,b) >$ for any two polynomials a, b (where $< .. >$ denotes the ideal generated by the specific elements). Thus, determining if a and b are relatively prime is the same as solving

$$\exists\, u, v \in \mathbb{C}[z], \deg u < m, \deg v < n : \quad a \cdot v + b \cdot u = 1. \tag{1}$$

Equation (1) is the same as

$$S(a,b) \cdot \begin{bmatrix} \vec{v} \\ \vec{u} \end{bmatrix} = \big(1, 0, \ldots, 0\big)^T \tag{2}$$

so that two polynomials are relatively prime if and only if one can determine the first column of the inverse of their corresponding Sylvester matrix. That this is equivalent to Sylvester's criterion is obvious from the next lemma which gives the inverse of a Sylvester matrix entirely in terms of the first column of its inverse.

Lemma 2.4 *Let* $f(z) = f_{-1}z^{-1} + ... + f_{1-m-n}z^{1-m-n} = \frac{u(z)}{a(z)} + \mathcal{O}(z^{-m-n})_{z\to\infty}$. *Then* $S(a,b)$ *is invertible with inverse given by*

$$\begin{bmatrix} v_0 & 0 & \cdots & \cdots & \cdots & 0 & b_0 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & & \vdots & \vdots & \ddots & \ddots & & & \vdots \\ v_{n-1} & \cdots & v_0 & 0 & \cdots & 0 & b_{n-1} & \cdots & b_0 & 0 & \cdots & 0 \\ u_0 & 0 & \cdots & \cdots & \cdots & 0 & -a_0 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & & \vdots & \vdots & \ddots & \ddots & & & \vdots \\ u_{m-1} & \cdots & u_0 & 0 & \cdots & 0 & -a_{m-1} & \cdots & -a_0 & 0 & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ 0 & f_{-1} & \cdots & f_{1-m-n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & f_{-1} \end{bmatrix} \tag{3}$$

Proof: Note that equation (1) gives $\frac{u(z)}{a(z)} + \frac{v(z)}{b(z)} = \frac{1}{a(z)\cdot b(z)} = \mathcal{O}(z^{-m-n})_{z\to\infty}$ and so

$f(z) = -\frac{v(z)}{b(z)} + \mathcal{O}(z^{-m-n})_{z\to\infty}$. Thus we have

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ 0 & f_{-1} & \cdots & f_{1-m-n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & f_{-1} \end{bmatrix} \cdot S(a,b) = \begin{bmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & \ddots & \ddots & \vdots & b_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ a_m & & \ddots & a_0 & b_n & & \ddots & b_0 \\ 0 & \ddots & & a_1 & 0 & \ddots & & b_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_m & 0 & \cdots & 0 & b_n \\ u_0 & 0 & \cdots & 0 & -v_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ u_{m-1} & & \ddots & 0 & -v_{n-1} & & \ddots & 0 \\ 0 & \ddots & & u_0 & 0 & \ddots & & -v_0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{m-1} & 0 & \cdots & 0 & -v_{n-1} \end{bmatrix}.$$

The inverse formula follows directly by multiplying the right side of the previous equation with the matrix on the left of equation (3). □

Remark 2.5 *We note that for our Sylvester inversion formula it is not important that b has precise degree n. In fact in the case $m = \deg a \geq n \geq \deg b$ all formulas remain valid.*

Remark 2.6 *Similar inversion formula can also be derived for matrices that express information about the existence of common roots, in particular for the Bézout matrix of two polynomials. If we assume, without loss of generality, that $m = \deg a \geq degb$, and choose $m = n$, then the Sylvester matrix $S(a,b)$ has size $2m \times 2m$, and we may partition it into four square blocks as follows*

$$S(a,b) = \begin{bmatrix} L(a) & L(b) \\ U(a) & U(b) \end{bmatrix}.$$

In this case, the matrix $B(a,b) := U(a) \cdot L(b) - U(b) \cdot L(a)$ coincides up to some reordering of columns and rows with the Bézout of a and b as considered by Fiedler [7, Chapter 7, p.164ff]. By making some block manipulations and using a similar argument as in Lemma 2.4 we obtain

$$B(a,b)^{-1} = \begin{bmatrix} f_{-m} & f_{-m-1} & \cdots & f_{1-2m} \\ f_{1-m} & f_{-m} & \cdots & f_{2-2m} \\ \vdots & \vdots & & \vdots \\ f_{-1} & f_{-2} & \cdots & f_{-m} \end{bmatrix}.$$

In other words, the inverse of $B(a,b)$ is a (Toeplitz) block found in the factorization of the inverse of $S(a,b)$.

3 Coprime Parameters

For our purposes we use our inversion formula to obtain information on the magnitude of the inverse of a Sylvester matrix. In this section we give an upper bound for the norm of the inverse of a Sylvester

matrix. This gives us a (initial) numerical parameter that can be used to determine if two polynomials are coprime.

Theorem 3.1 *Let u, v be polynomials of degrees at most $m-1$ and $n-1$ solving equation (1). Then*

$$\left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right| \leq ||S(a,b)^{-1}|| \leq \left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right| + 2 \cdot ||f|| \cdot ||(a,b)||. \tag{4}$$

Proof: Since (v, u) defines the first column of the inverse of $S(a, b)$ the inequality on the left of (4) follows directly from the definition of our polynomial and matrix norms. The bound on the right follows from our inverse formula. □

Theorem 3.1 gives a bound for the norm of the inverse of the Sylvester matrix in terms of the cofactors u, v, and the easily computable first coefficients f_j of the power series u/a. However it still remains to determine how good (or bad) such a bound will be. In particular, we need to determine the size of the coefficients f_j.

As a first step we note that Sylvester's matrix has a certain interesting duality property. Namely, let

$$\underline{a}(z) = z^m \cdot a(1/z), \quad \underline{b}(z) = z^n \cdot b(1/z), \tag{5}$$

that is,

$$\underline{a}(z) = a_m + a_{m-1}z + \ldots + a_0 z^m, \quad \underline{b}(z) = b_n + b_{n-1}z + \ldots + b_0 z^n.$$

The Sylvester matrices $S(a, b)$ and $S(\underline{a}, \underline{b})$ are the same up to reordering of rows and columns. In particular, their inverses have the same matrix norms. As such it is of interest to look at solutions to the diophantine equations

$$\underline{a}(z) \cdot \tilde{v}(z) + \underline{b}(z) \cdot \tilde{u}(z) = 1 \tag{6}$$

with $\tilde{u}, \tilde{v}$ being polynomials of degrees $m-1$ and $n-1$, respectively. Letting $\underline{u}(z) = z^{m-1} \cdot \tilde{u}(1/z)$ and $\underline{v}(z) = z^{n-1} \cdot \tilde{v}(1/z)$, equation (6) is the same as

$$a(z) \cdot \underline{v}(z) + b(z) \cdot \underline{u}(z) = z^{m+n-1}, \quad \text{that is, } S(a,b) \cdot \begin{bmatrix} \vec{v} \\ \vec{\underline{u}} \end{bmatrix} = \Big(0, \ldots, 0, 1\Big)^T. \tag{7}$$

The polynomials $\underline{v}, \underline{u}$ define a Padé approximant [3] of type $(m-1, n-1)$ for the power series $-b(z)/a(z)$.

Let

$$\kappa := \left|\left|\begin{bmatrix} v & \underline{v} \\ u & \underline{u} \end{bmatrix}\right|\right| = \max\{\left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right|, \left|\left|\begin{bmatrix} \underline{v} \\ \underline{u} \end{bmatrix}\right|\right|\}.$$

We may combine the results of Theorem 3.1 and (7), and obtain at the same time an upper bound for $||f||$.

Corollary 3.2 *With u, v and $\underline{u}, \underline{v}$ solutions of (1) and (7) we have*

$$\kappa \leq ||S(a,b)^{-1}|| \leq \kappa + 2 \cdot ||f|| \cdot ||(a,b)||, \tag{8}$$

where $||f|| = ||\underline{v} \cdot u - \underline{u} \cdot v||$. Furthermore, $||f|| \leq \kappa^2$.

Proof: The two inequalities in (8) are clear from Theorem 3.1. To determine $||f||$ we have that

$$\begin{aligned}
&f(z) - z^{1-m-n} \cdot [\underline{v}(z) \cdot u(z) - \underline{u}(z) \cdot v(z)] \\
&= \frac{u(z) - a(z) \cdot z^{1-m-n} \cdot [\underline{v}(z) \cdot u(z) - \underline{u}(z) \cdot v(z)]}{a(z)} + \mathcal{O}(z^{-m-n})_{z\to\infty} \\
&= \frac{z^{1-m-n} \cdot [b(z) \cdot \underline{u}(z) \cdot u(z) + a(z) \cdot \underline{u}(z) \cdot v(z)]}{a(z)} + \mathcal{O}(z^{-m-n})_{z\to\infty} \\
&= z^{1-m-n} \cdot \frac{\underline{u}(z)}{a(z)} + \mathcal{O}(z^{-m-n})_{z\to\infty} = \mathcal{O}(z^{-m-n})_{z\to\infty},
\end{aligned}$$

and so $f(z) = z^{1-m-n} \cdot [\underline{v}(z) \cdot u(z) - \underline{u}(z) \cdot v(z)]$. □

Numerical experiences seem to indicate [3, 4] that, for correctly scaled a and b, the quantity $||S(a,b)^{-1}||$ is proportional to κ and not of size κ^2. A slight generalization of Corollary 3.2 gives us more information about a class of polynomials (a,b) where this property is true.

First notice that for any Laurent polynomial $g(z) = g_{-1}z^{-1} + g_{-2}z^{-2} + ...$ we have

$$(\underline{v}(z) + g(z) \cdot b(z)) \cdot u(z) - (\underline{u}(z) - g(z) \cdot a(z)) \cdot v(z) = z^{m+n-1} \cdot f(z) + g(z) = z^{m+n-1} \cdot f(z) + \mathcal{O}(z^{-1}).$$

In other words, the polynomial part of the left hand side equals f, and $||f|| \leq ||(u,v)^T|| \cdot ||(\underline{v} + g \cdot b, \underline{v} + g \cdot a)||$, where we may choose a g to improve the upper bound for $||f||$ given in Corollary 3.2. More generally, let g_a, g_b be polynomials verifying

$$\deg g_a < m+n, \quad \deg g_b < m+n, \quad z^n a(z) g_a(z) + z^m b(z) g_b(z) = z^{2(m+n)-1} + \mathcal{O}(z^{m+n-1})_{z\to\infty}. \quad (9)$$

Then we have that $(g_a, -g_b) \cdot (u,v)^T = z^{2(m+n)-1} \cdot f + \mathcal{O}(z^{m+n-1})_{z\to\infty}$, again allowing for an estimate of $||f||$. Using Theorem 3.1 we obtain

Corollary 3.3 *Denote by $\rho_{m+n}(a,b)$ the minimum of the set of all products $||(a,b)|| \cdot ||(g_a, g_b)^T||$ where the pair (g_a, g_b) verifies (9). Then with u, v solutions of (1) we have*

$$\left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right| \leq ||S(a,b)^{-1}|| \leq \Big(1 + 2 \cdot \rho_{m+n}(a,b)\Big) \cdot \left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right|.$$

Note that $\rho_{m+n}(a,b)/||(a,b)||$ may be estimated above for instance by $||(\underline{u}, \underline{v})^T||$ in terms of the cofactors of the diophantine equation (7), or by $\rho_{m+n}(a,0)/||a||$ (resp. $\rho_{m+n}(0,b)/||b||$), the norm of the polynomial obtained by the first $m+n$ coefficients of the power series at zero of $\underline{a}(z)^{-1}$ (and of $\underline{b}(z)^{-1}$, respectively). Therefore, the quantity $\rho_{m+n}(a,b)$ may be close to one even if the Sylvester matrix $S(a,b)$ is ill–conditioned (see for instance the numerical results of [2]).

4 Closest Common Roots

In the previous section we obtained an upper bound (c.f. Corollary 3.2) for the norm of the inverse of the Sylvester matrix. Assuming, for the time being, that the computation of both (v,u) and $(\underline{v}, \underline{u})$ can be done in an efficient way (cf. [2]), we will have an effective method of determining when two polynomials are relatively prime. The only drawback to the above method is that our parameter (in this case $1/(\kappa + 2||f|| \cdot ||(a,b)||)$ which is a lower bound for $1/||S(a,b)^{-1}||$ and hence for $\epsilon(a,b)$), may

be too small since it could potentially be on the order of $1/\kappa^2$. In order to obtain a more precise bound we require a more detailed study for determining $\epsilon(a,b)$. The following statement is probably well known, however for the sake of completeness we provide a proof.

Theorem 4.1 *We have*

$$\epsilon(a,b) = \inf_{z\in\overline{\mathbb{C}}} \left|\left| \left(\frac{a(z)}{||(1,z^m)||}, \frac{b(z)}{||(1,z^n)||} \right) \right|\right| \tag{10}$$

where $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$. *The infimum on the right side is attained for a* $z^* =: \operatorname{ccr}(a,b) \in \overline{\mathbb{C}}$ *(called the "closest common root").*

Proof: Let $h(a,b,z) = \left|\left| \left(\frac{a(z)}{||(1,z^m)||}, \frac{b(z)}{||(1,z^n)||} \right) \right|\right|$. To see that $\epsilon(a,b) \geq h(a,b,z)$ for some $z \in \overline{\mathbb{C}}$, let a^* and b^* have the common root z. From Hölder's inequality we get

$$|a(z)| = |a(z) - a^*(z)| \leq ||\vec{a} - \vec{a}^*||_1 \cdot ||(1,z,...,z^m)^T||_\infty = ||a - a^*|| \cdot \max\{1, |z|^m\}$$

with a similar inequality for b. Therefore

$$||(a-a^*, b-b^*)|| = \max\{||a-a^*||, ||b-b^*||\} \geq h(a,b,z).$$

Taking the infimum on both sides leads to the first half of our assertion.

Note that the function $z \to h(a,b,z)$ is continuous over $\mathbb{C}$, and therefore attains its minimum on the unit disk. Also, we have $h(a,b,z) = h(\underline{a}, \underline{b}, 1/z)$, leading to

$$\inf_{z\in\overline{C}} h(a,b,z) = \min\{\inf_{|z|\leq 1} h(a,b,z), \inf_{|z|\leq 1} h(\underline{a},\underline{b},z)\},$$

showing that the infimum in (10) is attained. To show equality in (10), suppose that z^* is the closest common root of a, b, and consider

$$a^*(z) = a(z) - a(z^*) \cdot \begin{cases} 1 & \text{if } |z^*| < 1, \\ (z/z^*)^m & \text{otherwise,} \end{cases}$$

along with a similar b^*. □

Remark 4.2 *We see from the proof of Theorem 4.1 that* $\operatorname{ccr}(a,b)$ *is in fact the common root of the polynomials* a^*, b^* *which under all pairs of not coprime polynomials have minimal distance to* a, b. *Here* $\operatorname{ccr}(a,b) = \infty$ *is equivalent to saying that* $\deg a^* \leq m-1$ *and* $\deg b^* \leq n-1$. *Also,*

$$\epsilon(a,b) \begin{cases} = \min\limits_{|z|\leq 1} ||(a(z), b(z))|| \leq \min\limits_{|z|\leq 1} ||(\underline{a}(z), \underline{b}(z))|| & \text{if } |\operatorname{ccr}(a,b)| \leq 1, \\ = \min\limits_{|z|\leq 1} ||(\underline{a}(z), \underline{b}(z))|| \leq \min\limits_{|z|\leq 1} ||(a(z), b(z))|| & \text{if } |\operatorname{ccr}(a,b)| \geq 1, \end{cases}$$

with $\underline{a}, \underline{b}$ *as in (5).*

Remark 4.3 *A statement similar to Theorem 4.1 can also be made for other Hölder vector norms, and one may in addition consider weighted norms (useful, for example, in cases where only some of the*

coefficients may have inaccuracies). For instance, let $\alpha, \beta \in \mathbb{C}[z]$ be of degree m, and n, respectively, with positive coefficients. Then (compare also [5, Remark 4])

$$\inf\Big\{\sum_{j=0}^{m}\frac{|a_j-a_j^*|^2}{\alpha_j}+\sum_{j=0}^{n}\frac{|b_j-b_j^*|^2}{\beta_j}\ :\ (a^*,b^*)\ \textit{have a common root},\ \deg a^*\le m, \deg b^*\le n\Big\}$$
$$=\inf_{z\in\overline{\mathbb{C}}}\frac{|a(z)|^2}{\alpha(|z|^2)}+\frac{|b(z)|^2}{\beta(|z|^2)}.$$

Corless et al. [5, Section 2.6] and Karmarkar et al. [10] proposed to apply standard optimization algorithms for calculating a numerical GCD, and in particular for determining such a 2–counterpart of $\epsilon(a,b)$. Of course, for the problem of coprimeness it is preferable to take the above expression on the right since the number of free parameters is reduced from $m+n+1$ to 1.

One easily shows that $\epsilon(a,b)=\epsilon(b,a)=\epsilon(\underline{a},\underline{b})$, and that $\epsilon(a,b)\le\min\{||a||,||b||\}$. Also, it seems to be clear that a,b may not be ϵ–prime if they have zeros that are too close. In fact, denoting by z_a a zero of a, and by z_b a zero of b, respectively, we may show the estimate

$$\epsilon(a,b)\le\max\{m\cdot||a||,n\cdot||b||\}\cdot\frac{|z_a-z_b|}{\max\{1,|z_a|\}\cdot\max\{1,|z_b|\}},$$

where the distance of zeros is measured in some "chordal" metric.

From Lemma 2.1 and Theorem 4.1 we have

$$\epsilon(a,b)\ge\frac{1}{||S(a,b)^{-1}||}=\min_{y\neq0}\frac{||y\cdot S(a,b)||}{||y||},\quad \epsilon(a,b)=\min_{z\in\overline{\mathbb{C}}}\frac{||y(z)\cdot S(a,b)||}{||y(z)||},$$

where $y(z)=(1,z,...,z^{m+n-1})$. At present our "coprimeness parameter" requires the potentially large overestimate for $||S(a,b)^{-1}||$ given by Theorem 3.1 and Corollary 3.2. Can we improve this, for example by the following

$$\epsilon(a,b)\overset{?}{\ge}\frac{1}{\kappa}=\min\Big\{\frac{1}{||S(a,b)^{-1}\cdot e_1||},\frac{1}{||S(a,b)^{-1}\cdot e_{m+n}||}\Big\}?$$

In other words, can the norm of the inverse be replaced by only the norm of the first and/or last column of the inverse?

Corollary 4.4 *There holds $\epsilon(a,b)\ge\frac{1}{\kappa}$, and, more precisely,*

$$\min_{|z|\le1}||(a(z),b(z))||\ge\frac{1}{||(v,u)^T||},\quad \min_{|z|\le1}||(\underline{a}(z),\underline{b}(z))||\ge\frac{1}{||(\underline{v},\underline{u})^T||}.$$

Proof: In view of Remark 4.2, the estimate for $\epsilon(a,b)$ is a consequence of the other two estimates. In order to prove the second one, notice that

$$\min_{|z|\le1}||(a(z),b(z))||\ge\min_{|z|\le1}\frac{\left\|(a(z),b(z))\cdot\begin{bmatrix}v(z)\\u(z)\end{bmatrix}\right\|}{\left\|\begin{bmatrix}v(z)\\u(z)\end{bmatrix}\right\|}\ge\frac{1}{\left\|\begin{bmatrix}v\\u\end{bmatrix}\right\|}.$$

Here we have used the fact that, for every polynomial matrix U, there holds

$$\max_{z\le1}||U(z)||\le||U||.$$

Finally, the third estimate follows by symmetry. □

Remark 4.5 *From Remark 4.2 and the proof of Corollary 4.4 we see that, provided* $|\mathrm{ccr}\,(a,b)| \leq 1$, *we have the estimate* $\epsilon(a,b) \geq 1/||(u^*,v^*)^T||$ *for any polynomials* u^*, v^* *satisfying* $a \cdot v^* + b \cdot u^* = 1$, *even if the degree constraints of (1) are not valid. Thus, the bounds of Corollary 4.4 may be improved by considering* $(u^*,v^*) = (u,v) + \alpha \cdot (a,-b)$, *where* $\alpha \in \mathbb{C}[z]$ *is chosen in order to minimize the norm of* $(u^*,v^*)^T$. *In this context it is interesting to mention that by the Corona Theorem [13, Appendix 3] we may find functions* $u^\#, v^\#$ *analytic and bounded in the unit disk (i.e., elements of the Hardy space* H^∞*) such that*

$$a \cdot v^\# + b \cdot u^\# = 1, \quad \max_{|z|\leq 1} \sqrt{|u^\#(z)|^2 + |v^\#(z)|^2} \leq \frac{1}{\epsilon} + \frac{7 \cdot \sqrt{\log 1/\epsilon} + 20 \cdot \log 1/\epsilon}{\epsilon^2}, \tag{11}$$

provided that $\epsilon \leq \sqrt{|a(z)|^2 + |b(z)|^2} \leq 1$ *for all* $|z| \leq 1$. *Consequently, in the case* $||(a,b)^T|| \leq 1$ *it seems to be possible to find polynomials* (u^*,v^*) *by a suitable limiting procedure with* $1/||(u^*,v^*)^T||$ *lying between* $\epsilon(a,b)$ *and roughly its square.*

Remark 4.2 and Corollary 4.4 tell us that it is sufficient to solve only a single diophantine equation in order to determine an effective bound for $\epsilon(a,b)$, provided that we know in advance that the closest common root lies in or outside the unit disk. In some cases, such a localization of the closest common root may be given.

Lemma 4.6 *Suppose the roots of* a *and* b *all lie in the unit disk. Then* $|\mathrm{ccr}\,(a,b)| \leq 1$.

Proof: From (5) and Theorem 4.1 we know that

$$\epsilon(a,b) = \min\Big\{\min_{|z|\leq 1} ||(a(z),b(z))||, \min_{|z|\leq 1} ||(\underline{a}(z),\underline{b}(z))||\Big\}.$$

Thus for the assertion of Lemma 4.6 it is sufficient to show that $|a(z)| \leq |\underline{a}(\overline{z})|$ (and analogously that $|b(z)| \leq |\underline{b}(\overline{z})|$) for all $|z| < 1$, where as usual $\overline{z}$ denotes the complex conjugate of z. If $x_1, .., x_m$ are the roots of a, then this follows from

$$\frac{|a(z)|}{|\underline{a}(\overline{z})|} = \frac{|a(z)|}{|z|^m \cdot |a(1/\overline{z})|} = \prod_{j=1}^{m} \frac{|z - x_j|}{|1 - \overline{z} \cdot x_j|}$$

being less or equal to one for any $|z| < 1$, since $|z - x|/|1 - \overline{x} \cdot z| \leq 1$ for all x, z lying in the unit disk. □

For example, suppose that the roots of the polynomial c lie in the unit disk. By the Gauß–Lucas Theorem [12, p.22], the zeros of the derivative of c lie in the convex hull of the set of its zeros, and hence also in the unit disk. Thus if a, b are any (higher order) derivatives of c, then $|ccr(a,b)| \leq 1$.

5 Examples

In order to illustrate and to compare the findings of the preceding sections, we consider the following three simple examples

Example 5.1 *Let $a(z) = z^n - 1$, and $b(z) = b_0 + .. + b_n z^n$ with $||b|| \leq 1$. Then*

$$S(a,b) = \begin{bmatrix} -I_n & L \\ I_n & U \end{bmatrix} \quad \text{and} \quad S(a,b)^{-1} = \begin{bmatrix} -(U+L)^{-1} \cdot U & (U+L)^{-1} \cdot L \\ (U+L)^{-1} & (U+L)^{-1} \end{bmatrix}.$$

Therefore

$$||S(a,b)||_1 = 2, \quad \frac{||S(a,b)^{-1}||_1}{||(L+U)^{-1}||_2} \in [\frac{1}{\sqrt{n}}, 2\sqrt{n}].$$

The matrix $L+U$ is circulant, with eigenvalues $b(\omega^j)$, $j = 0, 1, .., n-1$, where ω is a primitive n-th root of unity. Moreover, $L+U$ is normal, and therefore $||(L+U)^{-1}||_2 = \max\{1/|\lambda| : \lambda$ is an eigenvalue of $L+U\}$. From Lemma 2.1 and Theorem 4.1 we then have

$$B := \min_j |b(\omega^j)| \geq \epsilon(a,b) \geq \frac{1}{||S(a,b)^{-1}||_1} \geq \frac{B}{2\sqrt{n}}.$$

Note also that Corollary 3.3 applies in this context with $\rho_{2n}(a,b) \leq 4$, giving

$$\left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right| \leq ||S(a,b)^{-1}||_1 \leq 9 \cdot \left|\left|\begin{bmatrix} v \\ u \end{bmatrix}\right|\right|.$$

Example 5.2 *Let $a(z) = z^m$, and $b(z) = b_0 + .. + b_n z^n$ with $||b|| \leq 1$. In this case the cofactors can be obtained explicitly: u from $b(z) \cdot u(z) = 1 + \mathcal{O}(z^m)_{z \to 0}$ via the Taylor expansion of $1/b$ at zero and v as a corresponding remainder. Then*

$$S(a,b) = \begin{bmatrix} 0 & L \\ I_n & U \end{bmatrix}, \quad ||S(a,b)||_1 = 1, \quad S(a,b)^{-1} = \begin{bmatrix} -U \cdot L^{-1} & I_n \\ L^{-1} & 0 \end{bmatrix}, \quad \frac{||S(a,b)^{-1}||_1}{||u||} \in [1,2],$$

the latter observation being in accordance with Corollary 3.3 since $\rho_{m+n}(a,b) = 1$. For example, if $b(z) = (1-2z)/3$ then

$$\frac{1}{6} \cdot \frac{1}{2^m - 1} \leq \frac{1}{||S(a,b)^{-1}||_1} \leq \epsilon(a,b) \leq a(1/2) = 2^{-m},$$

a consequence of $u(z) = 3 \cdot (1 + 2z + 4z^2 + ... + (2z)^{m-1})$.

Example 5.3 *With the same setting as in Example 5.2, let $b(z) = (\frac{1-z}{2})^m$. Then $||b|| = 1$, and*

$$u(z) = 2^m \cdot (1-z)^{-m} + \mathcal{O}(z^m) = 2^m \cdot \sum_{j=0}^{m-1} \binom{m-1+j}{j} z^j$$

with

$$||u|| = 2^m \cdot \binom{2m-1}{m-1} \approx \frac{2^{3m-1}}{\sqrt{\pi \cdot m}}.$$

This gives

$$\sqrt{\pi m} \cdot 8^{-m} \approx \frac{1}{2||u||} \approx \frac{1}{||S(a,b)^{-1}||_1} \leq \epsilon(a,b) = b(1/3) = 3^{-m}.$$

Thus the criterion of Corollary 4.4 does not always yield sharp bounds, since for large m we have $1/\kappa \approx 1/||S(a,b)^{-1}||_1 \not\approx \epsilon(a,b)$.

From Example 5.3 we also see that a "small" $\epsilon(a,b)$ in general does not imply that a has a root which is "close" to one of the roots of b.

6 Conclusion

We have considered the problem of determining when two polynomials are numerically relatively prime. A parameter has been given that improves a previous existing measure for numerical primeness. A sharper measure can be given in the case where it is known that the two polynomials have all their roots in the unit disk. This parameter is based on quantities which are efficiently obtainable in a numerically stable way [2]. The efficiency and numerical correctness of such a computation makes a good initial test for coprimeness before starting the more expensive computation of an ϵ–GCD.

References

[1] B. Beckermann, The stable computation of formal orthogonal polynomials, *Numerical Algorithms* **11** (1996) 1-23.

[2] B. Beckermann and G. Labahn, A fast, numerically stable Euclidean-like algorithm for detecting relatively prime numerical polynomials. *Journal of Symbolic Computation* (this issue).

[3] S. Cabay and R. Meleshko, A weakly stable Algorithm for Padé Approximants and the Inversion of Hankel matrices, *SIAM J. Matrix Analysis and Applications* **14** (1993) 735-765.

[4] S. Cabay, A. R. Jones and G. Labahn, Experiments with a Weakly Stable Algorithm for Computing Padé-Hermite and Simultaneous Padé Approximants, *ACM Trans. of Mathematical Software (TOMS)* **23**(1) (1997) 91-110.

[5] R.M. Corless, P.M. Gianni, B.M. Trager & S.M. Watt, The Singular Value Decomposition for Polynomial Systems, Proceedings ISSAC '95, ACM Press (1995) 195-207.

[6] I. Emiris, A. Galligo and H. Lombardi, Certified approximate univariate GCDs, *J. Pure and Applied Algebra*, **117** (1997) 229-251.

[7] M. Fiedler, *Special matrices and their application in numerical mathematics*, Martinus Nijhoff Publishers, Dordrecht (1986).

[8] K.O. Geddes, S.R. Czapor and G. Labahn, *Algorithms for Computer Algebra* (Kluwer, Boston, MA, 1992)

[9] N.J. Higham, *Accuracy and Stability of Numerical Algorithms* (SIAM, Philadelphia, 1996).

[10] N. Karmarkar and Y.N. Laksmann, Approximate polynomial greatest common divisors and nearest singular polynomials, Proceedings ISSAC '96, ACM Press (1996) 35-43.

[11] T. Kailath, Linear Systems, Prentice–Hall (1980).

[12] M. Marden, Geometry of Polynomials, *Math. Surveys* **3** (Amer. Math. Soc. Providence, RI, 1966).

[13] N.K. Nikol'skii, *Treatise of the shift operator*, (Springer, Berlin, Heidelberg, 1986).

[14] M.-T. Noda & T. Sasaki, Approximate GCD and its applications to ill–conditioned algebraic equations, *J.CAM* **38** (1991) 335-351.

[15] A. Schönhage, Quasi–GCD Computations, *J. Complexity* **1**(1985) 118-137.

[16] T.W. Sederberg and G.Z. Chang, Best linear common divisors for approximate degree reduction, Computer-Aided Design 25 (1993) 163-168.