ELSEVIER Journal of Pure and Applied Algebra 117&118 (1997) 229–251

JOURNAL OF PURE AND APPLIED ALGEBRA

Certified approximate univariate GCDs

Ioannis Z. Emiris[a,*], André Galligo[a,b], Henri Lombardi[c]

[a] *Projet S.A.F.I.R., I.N.R.I.A., B.P. 93, Sophia-Antipolis 06902, France*

[b] *Laboratoire de Mathématiques, Université de Nice – Sophia-Antipolis, Parc Valrose, Nice 06108, Cedex 2, France*

[c] *Laboratoire de Mathématiques, Université de Franche – Comté, Besançon 25030, France*

Abstract

We study the approximate GCD of two univariate polynomials given with limited accuracy or, equivalently, the exact GCD of the perturbed polynomials within some prescribed tolerance. A perturbed polynomial is regarded as a family of polynomials in a classification space, which leads to an accurate analysis of the computation. Considering only the Sylvester matrix singular values, as is frequently suggested in the literature, does not suffice to solve the problem completely, even when the extended euclidean algorithm is also used. We provide a counter-example that illustrates this claim and indicates the problem's hardness. SVD computations on subresultant matrices lead to upper bounds on the degree of the approximate GCD. Further use of the subresultant matrices singular values yields an approximate syzygy of the given polynomials, which is used to establish a gap theorem on certain singular values that certifies the maximum-degree approximate GCD. This approach leads directly to an algorithm for computing the approximate GCD polynomial. Lastly, we suggest the use of weighted norms in order to sharpen the theorem's conditions in a more intrinsic context. © 1997 Elsevier Science B.V.

1991 Math. Subj. Class.: 15A18, 65Y20, 68Q40

1. Introduction

The question of computing the approximate greatest common divisor (GCD) of a polynomial pair is being studied with renewed interest, as illustrated by the variety of

* Corresponding author. E-mail: emiris@sophia.inria.fr.

PII S0022-4049(97)00013-3

different approaches to the problem within the last couple of years [6, 8, 16–18, 23, 26]; Section 2 presents a comprehensive account of previous work. In the same area lies the problem of computing approximate solutions to systems of polynomials whose coefficients are only imperfectly known.

Such questions relate to both algebraic and numerical computation and belong to an area sometimes called *seminumerical computation*. The grand project of this area is to cross-fertilize the two fields and use the advantages of each to facilitate computation in the other. Here, we exploit the mathematical veracity of algebra to provide a solid foundation for performing numerical computation, while we exploit the speed of the latter.

In addition to the richness of mathematical issues involved, the answers to problems on imperfectly known polynomials have important practical ramifications. Whenever laboratory measurements are involved, data may be given by floating point coefficients to limited accuracy or only a certain number of significant digits may be obtainable efficiently. To mention only a sample of applications, there is a multitude of graphics and modeling, robotics, vision and control theory problems where noise corrupts the input parameters [10, 17, 20, 25].

Our first contribution is a counterexample to a direct approach relying only on the Sylvester matrix singular values and on the extended euclidean algorithm [6]. This discussion completes, in a sense, the counterexample in [8] that showed that Euclid's algorithm only gives a lower bound to the maximum degree of the approximate GCD. We conclude that Euclid's algorithm is unable to find the maximum-degree GCD polynomial within some guaranteed error, contrary to claims in certain papers such as [16]. This illustrates the inherent difficulty of the problem.

The main contribution of this paper is a gap theorem on the singular values of the subresultant matrices that *guarantees* the degree for the approximate GCD and, moreover, that this degree is maximum within the given tolerance. The current gap theorem is much tighter than the one obtained in [8]; that article relied on a geometric approach based on the polynomial roots via Ostrowski's theorem. Here, a direct algebraic approach is adopted that yields a gap with linear dependence on the singular value that is almost zero, whereas the old result had a polynomial dependence. The present approach leads to a polynomial-time algorithm in the degrees of the input and output polynomials, based on the singular value decomposition (SVD) of subresultant matrices.

Our approach generalizes the usual notion of backward error, since we solve exactly a slightly perturbed problem. A perturbed polynomial is regarded as a family of polynomials in a classification space, which leads to an accurate analysis of the computation. Different solutions are compatible with different degrees of uncertainty. Trying to maximize the degree of the GCD is the natural approach in the presence of noise.

Definition 1. Fix integers n, m and a metric $|\cdot|$ on the spaces of univariate polynomials $\mathscr{P}_n, \mathscr{P}_m \subset \mathbb{C}[x]$ of degree bounded by n and m respectively. Given $f \in \mathscr{P}_n$, $g \in \mathscr{P}_m$ and

$\varepsilon > 0$, the *degree* of the ε-GCD is defined to be the *maximum* integer r such that there exist $\hat{f} \in \mathscr{P}_n$, $\hat{g} \in \mathscr{P}_m$, with $|f - \hat{f}|, |g - \hat{g}| \leq \varepsilon$ and $\deg(\gcd(\hat{f}, \hat{g})) = r$.

The polynomial $\gcd(\hat{f}, \hat{g})$ is the ε-GCD of f, g. In this paper, we are concerned with computing the degree and the actual ε-GCD polynomial. Additionally, we may further consider bounding the error in the computed GCD, i.e., if we are given $\theta > 0$, we would like to find a polynomial $h \in \mathscr{P}_r$ such that $|h - \gcd(\hat{f}, \hat{g})| \leq \theta$.

Corollary 12 to the main theorem imposes some mild assumptions in order to simplify the conditions under which the ε-gcd degree is guaranteed to be equal to r. If τ_k is the smallest singular value of the kth subresultant mapping and $n \geq m$, then

$$\tau_{r-1} \leq 2^{-5n-3} \tau_r^{n+2} \varepsilon \quad \Rightarrow \quad \deg \varepsilon\text{-gcd} = r.$$

Main Theorem 6 gives more precise, albeit more involved bounds.

We are interested in floating point computations, but we ignore roundoff error, assuming that the algorithms are executed in sufficiently high precision to make the latter too small compared to the allowed tolerance. For now the norm used is the standard L_2-norm, also known as euclidean norm. The present approach lends itself to a direct generalization to weighted norms, as indicated in Section 8.

The paper is structured as follows. The next section describes existing work in the area. Section 3 defines the norms of interest and provides a list of bounds on the norm of polynomial products, as well as relations among the different norms. Section 4 introduces singular values of the Sylvester matrix and the subresultant chain and mentions some known bounds on the degree of the ε-GCD. A counterexample to the method of [6], using only the Sylvester matrix singular values, is described in Section 5 in order to illustrate the problem's difficulty. The main gap theorem is derived in Section 6, together with the conditions under which it certifies the ε-GCD degree. It leads to an algorithmic method for computing the approximate GCD polynomial in Section 7. Section 8 proposes weighted norms and we conclude with open questions.

2. Previous work

Among the euclidean algorithms that compute exact GCDs, it is known that the subresultant version is the most efficient, since it strikes a balance between coefficient growth and computational effort. Subresultant chains were essentially introduced in [27] and used for computing GCDs in [5, 4]. The same objects had been studied in a very general setting in [14] and rediscovered in the latter terms by [13].

A significant portion of the literature is devoted to methods derived from Euclid's algorithm and its extensions. Schönhage [24] proposes ways to compute the *quasi-GCD*, under the particular assumption that the coefficients of the given polynomials can be given to arbitrary accuracy by some oracle. The algorithm is not simple as it uses a special change of variable in order to control coefficient size and optimize complexity on a *pointer machine*.

An ε-GCD is sought by Hribernig and Stetter [16], who use the classical euclidean algorithm in order to identify the clusters of polynomial roots. An improved approach has been recently proposed in [26]. Noda and Sasaki came to study approximate GCDs via the need to define approximate square-free decompositions and introduced a scaled euclidean algorithm [22]. The extended euclidean algorithm and its variants offer no guarantee that an ε-GCD has been found. It only returns a common divisor within the prescribed tolerance, called an ε-divisor in [8], where a counterexample was given to illustrate the limitations of this method. Approximate GCDs have been studied in relation to various applications, including the computation of proper parameterizations and rational curve degree reduction [25].

Numerical GCDs have been studied in the control theory literature, where numerical computation, even with rational input, is bound to produce some error in the result. Karkanias and Mitrouli [17] use standard backward error analysis techniques to show that the numerical GCD obtained by an SVD computation will be sufficiently close to the exact one. However, this approach can only return an upper bound on the degree of the ε-GCD.

The SVD of Sylvester's matrix has long been known within the numerical computation community to be rather stable. Corless et al. [6] emphasized the merits of this approach in the setting of seminumerical computation and computer algebra. Their problem is slightly different, since the a priori bound ε is not guaranteed to bound the perturbation. However, nor does the a posteriori bound $\hat{\varepsilon}$ correspond necessarily to the perturbation that *maximizes* the approximate GCD degree. This is illustrated in Section 5 by a counterexample.

Corless et al. extend this work to polynomial systems, taking advantage of resultant formulations for approximating the common roots. They offer heuristics and examples based on Lazard's resultant formulation [19]. This subject deserves a more rigorous study in light of the recent interest in resultant methods, especially since other approaches, such as the Newton matrix [7], may also treat over-constrained systems.

A recent approach consists in regarding the problem as an optimization question, where we try to minimize the distance of the given polynomials to the perturbed pair. Karmarkar and Lakshman [18] prove that the complexity of this optimization problem is polynomial in the degrees of the given polynomials and the bit size of their coefficients and simply exponential in the degree of the approximate GCD. They apply their techniques to perturbing a polynomial so that it has multiple roots, a problem studied in [15].

The univariate GCD identifies the common roots of the given polynomials. The inverse viewpoint is also interesting, as illustrated in [23] where approximations to the roots of the given polynomials are computed and then matched in order to arrive at the approximate GCD. Pan studies the combinatorial problem that must be solved and shows that the complexity is polynomial in the input degrees as well as the degree of the GCD.

Emiris et al. [8] formalized the discussion and demonstrated that the variants of Euclid's algorithm only supply a lower bound on the degree. They proved sufficient

conditions for obtaining upper bounds on the degree which, coupled with the euclidean algorithm, lead to heuristics for computing the degree accurately. Moreover, they provided a gap theorem on the singular values of the subresultant matrices which guarantees the degree, thus offering the first complete certification condition. In the notation of the present paper, their main result [8, Theorem 6.8] required that

$$\tau_{r-1} \leq \varepsilon \quad \text{and} \quad \tau_r > \varepsilon\sqrt{d-r+1} > d^3(\gamma+1)^{3d}\tau_{r-1}^{1/d},$$

where $d = n+m-r+1$, $\gamma = 2|f,g|/(\tau_r - 4|f,g|\tau_{r-1})$ and $|f,g|^2 = |f|^2+|g|^2$ in order to guarantee an ε-GCD degree of r. Yet, this gap was too loose to be effective, essentially because of its dependence on Ostrowski's theorem. The present paper continues this work in the sense that it sharpens this gap to obtain a linear dependence on τ_{r-1}. The approach is direct and uses solely some properties of Euclid's algorithm and of the singular values of certain subresultant matrices. It leads to an efficient algorithm of polynomial complexity in the degrees of the input and the GCD polynomials.

We concentrate on monomial bases for convenience. Different bases, such as the Bernstein polynomials [9], may improve stability but are beyond the scope of this paper.

3. Matrix and polynomial norms

This section contains a series of properties useful in our presentation. We equip the space of all linear transformations from $\mathbb{C}^{n+m-2r}$ to $\mathbb{C}^{n+m-r}$ with the following *operator* norm.

$$\|A\| = \sup_x \frac{\|Ax\|}{\|x\|} \quad \text{for any mapping } A \text{ and vector } x \in \mathbb{C}^{n+m-2r},$$

where $\|x\|$ represents the L_2-norm of x. If we denote by $a_{i,j}$ the entries of the matrix of A it can be shown that

$$\|A\|^2 \leq \sum_{i,j} |a_{i,j}|^2,$$

where $|c|$ is the module of the complex number c.

We define polynomial norms

$$|P|_l = (|p_0|^l + \cdots |p_d|^l)^{1/l}, \quad l \geq 1$$

and

$$|P|_\infty = \max\{|p_0|,\ldots,|p_d|\} \quad \text{where } P = \sum_{i=0}^{d} p_i x^{d-i} \in \mathbb{C}[x].$$

We shall denote $|P|_2$ simply as $|P|$. The following relations are known:

$$|P|_\infty \leq |P| \leq |P|_1 \leq (d+1)|P|_\infty,$$

$$\frac{|P|_1}{\sqrt{d+1}} \le |P| \le \sqrt{d+1}|P|_\infty.$$

For any $P_1, P_2 \in \mathbb{C}[x]$, let d stand for the sum of their degrees. [3, Theorem 1.1 and Remark p. 231] shows that

$$|P_1P_2|_l \le 2^{d(1-1/l)}|P_1|_l|P_2|_l, \quad 1 \le l \le \infty, \qquad |P_1||P_2| \le 2^d|P_1P_2|. \tag{1}$$

For monic P_1, P_2, [11] proved $|P_1|_\infty|P_2|_\infty \le \sqrt{d+1}\ 2^d|P_1P_2|_\infty$.

We let the norm of a polynomial *pair* be the square root of the sum of the two squared norms, and denote it $|\cdot,\cdot|_l$. In particular, for two polynomials P_1 and P_2, $|P_1, P_2| = \sqrt{|P_1|^2 + |P_2|^2}$.

Weighted norms assign different weights to different coefficients in order to take into account the importance of middle coefficients in polynomial multiplication.

The weighted L_l-norm, for $l \ge 1$, is defined as follows:

$$\langle P\rangle_l = \left(|p_0|^l \Big/ \binom{d}{0}^{l-1} + \cdots + |p_i|^l \Big/ \binom{d}{i}^{l-1} + \cdots + |p_d|^l \Big/ \binom{d}{d}^{l-1} \right)^{1/l}.$$

We denote by $\langle P\rangle$ the L_2 weighted norm $\langle P\rangle_2$.

The first merit of this norm is that it remains invariant, for *bivariate homogeneous* polynomials, under unitary changes of variables. This is the reason it has been used in invariant theory [28]. Let $\bar{a}$ denote the complex conjugate of a, then

$$\begin{pmatrix} y \\ y_0 \end{pmatrix} = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \begin{pmatrix} x \\ x_0 \end{pmatrix}, \quad |a|^2 + |b|^2 = 1 \Rightarrow \langle P(x, x_0)\rangle = \langle P(y, y_0)\rangle. \tag{2}$$

The second advantage is that the norm of the product is rather tightly bound above and below by the product of the norms [3, Theorem 1.2]. Let d_i be the degree of P_i, $i = 1, 2$, then the following bounds are best possible:

$$\binom{d_1+d_2}{d_1}^{-1/2} \langle P_1\rangle\langle P_2\rangle \le \langle P_1P_2\rangle \le \langle P_1\rangle\langle P_2\rangle. \tag{3}$$

Some simple properties needed below are:

$$\langle P_1 + P_2\rangle \le \langle P_1\rangle + \langle P_2\rangle, \qquad \langle cP\rangle = |c|\langle P\rangle,$$

for any complex constant c.

The standard and weighted norms are related as follows [3], where $\deg P = d$ and $\deg P_i = d_i$:

$$\begin{aligned} &(d!)^{1/l-1}|P|_l \le \langle P\rangle_l \le |P|_l, \quad l \ge 1, \\ &\binom{d}{\lfloor d/2\rfloor}^{-1/2} |P| \le \langle P\rangle \le |P|, \\ &\langle P_1\rangle\langle P_2\rangle \le \sqrt{d_1+d_2+1}\ 2^{(d_1+d_2)/2}|P_1P_2|_l, \quad l = 1,\ 2,\ \infty. \end{aligned} \tag{4}$$

Theorem 2. *For any* $P_1, P_2 \in \mathbb{C}[x]$ *with respective degrees* d_1, d_2,

$$\frac{|P_1||P_2|}{2^{d_1+d_2}} \leq |P_1 P_2| \leq \sqrt{2}|P_1||P_2|.$$

A slightly tighter lower bound is

$$|P_1||P_2|\binom{d_1+d_2}{d_1}^{-1/2}\binom{d_1}{\lfloor d_1/2\rfloor}^{-1/2}\binom{d_2}{\lfloor d_2/2\rfloor}^{-1/2} \leq |P_1 P_2|.$$

Proof. For the first statement, the left-hand side bound is simply bound (1). To prove the right-hand side bound, assume without loss of generality, that $d_1 \geq d_2$ and denote the coefficients of P_1 and P_2, respectively, by $(a_0, \ldots, a_{d_1})$ and $(b_0, \ldots, b_{d_2})$. Then, partitioning the coefficients of the product and applying the Cauchy–Schwartz inequality to each of the three sums we can upper bound the squared norm $|P_1 P_2|^2$ by

$$\left(\sum_{k=0}^{d_2}\sum_{i=0}^{k}|a_i b_{k-i}|^2 + \sum_{k=2d_2+1}^{d_1+d_2}\sum_{i=k-d_2}^{\min\{k,d_1\}}|a_i b_{k-i}|^2\right) + \sum_{k=d_2+1}^{2d_2}\sum_{i=k-d_2}^{\min\{k,d_1\}}|a_i b_{k-i}|^2$$

$$\leq 2\sum_{i=0}^{d_1}\sum_{j=0}^{d_2}|a_i b_j|^2 = 2|P_1|^2|P_2|^2.$$

The last statement of the theorem follows from bound (3) and the relation between euclidean and weighted norms in (4). Asymptotically, the squared divisor of the product of the norms is estimated as follows, by applying Stirling's approximation:

$$\binom{d_1+d_2}{d_1}\binom{d_1}{\lfloor d_1/2\rfloor}\binom{d_2}{\lfloor d_2/2\rfloor} = \Theta\left(\frac{2^{d_1+d_2+2}(d_1+d_2)^{d_1+d_2+1/2}}{\pi^{3/2}d_1^{d_1+1}d_2^{d_2+1}}\right),$$

which is upper bounded by $2^{2(d_1+d_2)}$. □

For simplicity, we usually apply the first lower bound. For completeness, we give bounds on the coefficients of polynomial divisors.

Proposition 3 (Mignotte [21] and Beauzamy [2]). *Let* $P_1, P_2 \in \mathbb{C}[x]$, $\deg P_i = d_i$, *and let* P_2 *divide* P_1. *Write* P_2 *as* $p_0 x^{d_2} + \cdots + p_{d_2}$. *Then*

$$|p_i| \leq \binom{d_2}{i}|P_1|, \qquad 0 \leq i \leq d_2.$$

If, moreover, $P_1, P_2 \in \mathbb{Z}[x]$ *and* P_1 *has a nonzero constant term, then*

$$|p_i| \leq \sqrt{\frac{1}{2}\binom{d_2}{i}\binom{d_1}{d_2}}\langle P_1\rangle,\ 0 \leq i \leq d_2 \quad \textit{and} \quad |P_2|_\infty \leq \sqrt{\frac{\sqrt{27}\,3^{d_1}}{4\pi d_1}}\langle P_1\rangle.$$

4. Subresultant matrix singular values

This section introduces subresultant matrices and their singular values and states some bounds on the degree of the approximate GCD derived in [8]. We apply the powerful tool of singular value decomposition but we only require the minimal singular value of a subresultant matrix.

The minimal singular value of a matrix is the reciprocal of the square root of the operator norm of the inverse of the corresponding Gram matrix. This property can be used for computing the minimal value without performing an SVD. This possibility deserves further study.

Formally, any linear map Φ between $\mathbb{C}^p$ and $\mathbb{C}^q$ equipped with their usual hermitian norms can be written, after suitable orthogonal changes of coordinates, as a matrix whose only nonzero entries are real, nonnegative and on the diagonal. The ordered diagonal elements

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(p,q)}$$

are called the *singular values* of the map Φ and can be computed, together with the coordinate changes, by an SVD. The stability of computing the overall SVD is quantified in [1]. The singular values of Φ describe how the map Φ deforms the objects from the source space to the target space. σ_1 is the norm of the map Φ. Let us denote by $S_1(F)$ the unit sphere of any subspace F of $\mathbb{C}^p$. Then,

$$\sigma_{r+1} = \inf_{\substack{F \subset \mathbb{C}^p: \\ \operatorname{codim} F = r}} \left\{ \sup_{x \in S_1(F)} \|\Phi(x)\| \right\}, \qquad \sigma_r = \sup_{\substack{F \subset \mathbb{C}^q: \\ \dim F = r}} \left\{ \inf_{x \in S_1(F)} \|\Phi(x)\| \right\}.$$

The rank of Φ is larger than or equal to r if and only if $\sigma_r > 0$. In the real case, if $p \leq q$, the first p elements of the new base of $\mathbb{R}^q$ are given by the principal axes of the ellipsoid image of the unit ball. The knowledge of the singular values allows a discussion on the numerical rank of Φ. It can be shown [12, Corollary 2.3.3] that if we perturb Φ by a linear map $\Delta\Phi$ of norm $\|\Delta\Phi\|$, such that $\sigma_r > \|\Delta\Phi\| > \sigma_{r+1}$, then the rank of $\Phi + \Delta\Phi$ may get down to r but cannot reach $r-1$. On the other hand, it may go up to $\min(p,q)$ but this does not depend on the norm of $\Delta\Phi$.

The problem at hand is somewhat related to the widely studied problem of sensitivity of the eigenvalues and rank of an arbitrary numerical matrix. Recall that the singular values of matrix Φ correspond to the square roots of the eigenvalues of matrix $\Phi^H\Phi$. The numerical sensitivity of the eigenvalues is studied in [12, Section 7.2] and the references thereof; a powerful tool is Gershgorin's circle theorem. Our treatment does not rely on these general results for it exploits the rich structure of subresultant matrices by looking directly at their singular values and the implications for the corresponding polynomials.

Let polynomials f, g have, respectively, degree n, m, where $f = f_0x^n + \cdots + f_n$. To every polynomial pair and $0 \leq r \leq m \leq n$ we associate the subresultant mapping

$\mathrm{Sy}_r(f,g)$ given by

$$\mathrm{Sy}_r(f,g) : (u,v) \mapsto uf + vg, \qquad \deg u \leq m-r-1, \quad \deg v \leq n-r-1.$$

Its matrix in the usual monomial basis has $m-r$ columns corresponding to f, $n-r$ columns corresponding to g and a total of $m+n-r$ rows. Below zero entries are left empty:

$$\mathrm{Sy}_r(f,g) = \begin{bmatrix} f_0 & & & g_0 & & \\ f_1 & f_0 & & g_1 & g_0 & \\ \vdots & \ddots & & \vdots & \ddots & \\ & \vdots & f_0 & & \vdots & g_0 \\ f_n & & & g_m & & \\ & f_n & \vdots & & g_m & \vdots \\ & & \ddots & & & \ddots \\ & & f_n & & & g_m \end{bmatrix}.$$

If $r = 0$, the Sylvester mapping and the respective matrix is denoted $\mathrm{Sy}_0(f,g)$. The following properties are well known:

- $\mathrm{Sy}_0(f,g)$ is of rank $m+n-r$ if and only if $\deg(\gcd(f,g)) = r$.
- $\mathrm{Sy}_r(f,g)$ has full rank i.e. $m+n-2r$ if and only if $\deg(\gcd(f,g)) \leq r$.

If $\sigma_1 \geq \cdots \geq \sigma_{m+n}$ are the singular values of the Sylvester matrix $\mathrm{Sy}_0(f,g)$ and $\gamma_{m-1} \geq \cdots \geq \gamma_0 \geq 0$ are the m last singular values of $\mathrm{Sy}_0(f,g)$, then $\gamma_k = \sigma_{m+n-k}$.

Let us denote by τ_r or $\tau_r(f,g)$ the last singular value of $\mathrm{Sy}_r(f,g)$. Then, $\tau_0 = \gamma_0 = \sigma_{m+n}$ whereas $\tau_m > 0$ because $\deg(\gcd(f,g)) \leq m$ under hypothesis $m \leq n$. Moreover, we extend a known result [8, Lemma 5.1] by simply applying the definition of τ_l.

Lemma 4. *With the previous notation, we have*

$$\tau_m \geq \tau_{m-1} \geq \cdots \geq \tau_1 \geq \tau_0 = \gamma_0.$$

Proof. By definition, τ_r and τ_{r+1} are the minimum norms of the image of any element of the unit ball under the respective subresultant mapping. For an appropriate polynomial pair (u,v) we have

$$\tau_{r+1} = \frac{|uf+vg|}{|u,v|} = \frac{|xuf+xvg|}{|xu,xv|} \geq \tau_r.$$

This concludes the proof for $r = 0,\ldots,m-1$. □

The following proposition bounds the degree from above, solely on the basis of the singular values. An analogous result is used in [6].

Proposition 5 (Emiris et al. [8, Proposition 5.4]). *With the previous notation,*

$$\varepsilon \leq \frac{\tau_r(f,g)}{\sqrt{m+n-2r}} \quad \textit{and} \quad \varepsilon \leq \frac{\gamma_r(f,g)}{\sqrt{m+n}} \textit{ each implies } \deg \varepsilon\text{-gcd}(f,g) \leq r.$$

Applying the stronger Theorem 2 derived in this paper, we can improve the first condition to $\varepsilon \leq \tau_r/\sqrt{2}$.

5. On the problem's hardness

Here we describe a counterexample to the method proposed in [6], demonstrating the insufficiency of Euclid's algorithm and the need for certified bounds on the degree of the approximate GCD. The example of [8, Section 3] showed that the euclidean algorithm alone gives only a lower bound on the degree.

Algorithm 2.4 in [6] claims to compute an $\hat{\varepsilon}$-GCD, where $\hat{\varepsilon}$ is a posteriori bound, but the given proof for Claim 2 is incomplete. Indeed, with our notation, nothing forbids that we have

$$\gamma_r(f,g) \leq \varepsilon < \gamma_{r+1}(f,g)/\sqrt{m+n} \leq \hat{\varepsilon},$$

whereas there exist perturbations $\leq \hat{\varepsilon}$ but superior to ε yielding a GCD degree of degree $r+2$, where $\hat{\varepsilon}$ is a posteriori bound defined as $\max\{|\Delta f|, |\Delta g|\}$. Thus, computing a common divisor of $(f+\Delta f, g+\Delta g)$ with $\max\{|\Delta f, \Delta g|\} \leq \hat{\varepsilon}$ does not imply that it is an $\hat{\varepsilon}$-GCD of (f,g), since the degree is *not* necessarily maximized. The abstract setting is illustrated in Fig. 1. The algorithm would return $r+1$ whereas the degree of $\hat{\varepsilon}$-GCD may be $r+2$. To be rigorous, one would have to restart the whole process with $\hat{\varepsilon}$. The discontinuity that lies at the heart of the matter, and which does not depend on the euclidean procedure, is formalized in [8, Section 4.3].

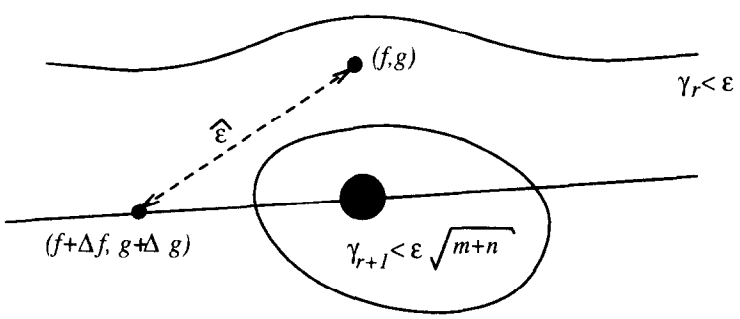

Fig. 1. The fat point represents the pairs with a GCD of degree $\geq r+2$, whereas the straight line includes the pairs corresponding to degree $\geq r+1$. The top curve is the boundary of neighborhood $\gamma_r \leq \varepsilon$ and the closed curve is neighborhood $\gamma_{r+1} \leq \varepsilon\sqrt{m+n}$. Here, the degree of $\hat{\varepsilon}$-GCD may be superior to $r+1$, suggested by SVD with the given ε.

A concrete instance that satisfies the conditions above is the following. Let

$$f = (x-1)(x-2) = x^2 - 3x + 2,$$

$$g = (x-1.08)(x-1.82) = x^2 - 2.9x + 1.9656,$$

$$\varepsilon = 0.016.$$

The singular values of the Sylvester matrix are

$$\sigma_4 = \gamma_0 = 0.01563304540 < \sigma_3 = \gamma_1 = 0.03335507319 < \sigma_2$$

$$= \gamma_2 = 3.110021071 < \sigma_1 = \gamma_3 = 6.698774418.$$

The exact GCD is 1. Algorithm 2.4 of Corless et al. sets the approximate GCD degree to $4-k=1$, since

$$\sigma_{k+1} = 0.01563304540 \leq \varepsilon < \frac{\sigma_k}{\sqrt{2+2}} = 0.01667753660.$$

Running Euclid's algorithm and stopping at a linear polynomial, one obtains approximate GCD,

$$g - f = 0.1x - 0.0344 = 0.1(x - 0.344).$$

Minimizing the perturbations which allow this to be a common divisor we obtain

$$\Delta f = \Delta g = 0.1135283784x^2 + 0.3300243558x + 0.9593731274$$

$$\Rightarrow \min_{\Delta f, \Delta g} \max\{|\Delta f|, |\Delta g|\} = 1.020882738 = \hat{\varepsilon}.$$

But we can find perturbations of norm smaller than $\hat{\varepsilon}$ that lead to a common quadratic divisor $(x-0.96)(x-2.04)$ with

$$\Delta f = -0.0416, \qquad \Delta g = -0.1x - 0.0072 \Rightarrow \max\{|\Delta f|, |\Delta g|\} = 0.0416.$$

This shows that Euclid's algorithm is not a safe method for the approach used in [6], for it does not return an $\hat{\varepsilon}$-GCD as claimed. Stronger upper as well as lower bounds and gap theorems are, therefore, needed, and they constitute precisely the focus of this paper.

6. A certification theorem

This section contains the main result. The proof traces an algorithm, detailed in the next section, for computing the ε-GCD.

Let f and g denote two polynomials of degree n and m such that $|f| = |g| = 1$. For readability of the formulae we assume $n \geq m$. We denote by $\tau_0 \leq \cdots \leq \tau_{m-1} \leq \tau_m$ the increasing sequence of the minimal singular values of the subresultant mappings

$\mathrm{Sy}_r(f, g)$ of f and g. A tolerance ε on the norm of an admissible perturbation is fixed; refer to Definition 1.

The main theorem is the following; its assumptions are used throughout the section. Note that there is no intrinsic reason to exclude r from being zero except that then τ_{r-1} would be undefined, so we restrict attention to $r \geq 1$.

Theorem 6. *Suppose that* $\tau_{r-1} \leq \varepsilon \leq \tau_r/\sqrt{2}$ *for* $\varepsilon \leq 1$, $1 \leq r \leq m \leq n$ *and*

$$\tau_r > \tau_{r-1} 2^{2n+m-2r}. \tag{5}$$

If, moreover,

$$\left(1 + \frac{2 + \tau_r^2}{\tau_r - \tau_{r-1}}\right)^{n+1} \left(1 + \frac{2^{2n+m-2r}}{\tau_r - \tau_{r-1} 2^{2n+m-2r}}\right) \tau_{r-1} \leq \varepsilon,$$

then the degree of the ε-$\mathrm{GCD}(f, g)$ *is equal to* r.

The proof takes the rest of the section. First, we prove a proposition for bounding the norms of the results of polynomial division. Let $lc(f)$ express the leading coefficient of polynomial f.

Proposition 7. *Let* A *and* B *be polynomials of degree* α *and* β, *respectively, such that* $\alpha \geq \beta$. *Then, polynomial division of* A *by* B *is written*:

$$A = BQ + R, \qquad \deg(Q) = \alpha - \beta, \quad \deg(R) \leq \beta - 1.$$

Let leading coefficient $b_0 = lc(B)$, *let polynomial* $C = B - b_0 x^\beta$, *and let bound* $M \geq 1 + |C/b_0|$, *then*

$$|Q| \leq \frac{|A|}{|b_0|} \frac{M^{\alpha-\beta+1} - 1}{M - 1} \quad \text{and} \quad |R| \leq |A| M^{\alpha-\beta+1}.$$

Proof. We write the division algorithm as a sequence of $\alpha - \beta + 1$ subtractions. More precisely, we consider the sequence of partial remainders A_i, for $\alpha > i \geq \beta - 1$, such that $A_{\beta-1} = R$ and $\deg A_i = i$. If we set $A_\alpha = A$, then

$$A_{i+1} = -\frac{lc(A_{i+1})}{b_0} B x^{i+1} + A_i, \qquad \alpha > i \geq \beta - 1,$$

therefore,

$$A_i = A_{i+1} - lc(A_{i+1}) x^{i+1} + lc(A_{i+1}) x^{i+1-\beta} C/b_0, \qquad \alpha > i \geq \beta - 1.$$

Then,

$$|A_i| \leq |A_{i+1}|(1 + |C|/|b_0|) \leq |A_\alpha|(1 + |C|/|b_0|)^{\alpha-i}, \qquad \alpha > i \geq \beta - 1,$$

which implies

$$|lc(A_i)| \leq |A_\alpha| M^{\alpha-i}, \qquad \alpha > i \geq \beta - 1.$$

Since $b_0 Q = \sum_{i=\beta-1}^{\alpha-1} lc(A_{i+1}) x^{i+1}$ and $R = A_{\beta-1}$, the claim follows. □

We now derive an approximate syzygy by Sy_{r-1} and show that the syzygy polynomials have not-too-small leading coefficients and, moreover, are relatively prime. Hence we find the degree of the approximate least common multiple (LCM) of f, g which implies the degree of the approximate GCD.

We denote by u and v polynomials of degree (less than or equal to) $m-r$ and $n-r$ respectively, and $|u,v| = \sqrt{|u|^2 + |v|^2} = 1$. Polynomials u, v are defined by the subresultant mapping Sy_{r-1} for which the singular value τ_{r-1} is smaller than ε. Polynomial pair u, v can be regarded as an *approximate syzygy* of the input pair f, g. More precisely we set

$$uf - vg = T, \qquad |T| = \tau_{r-1} = \min_{u,v} \frac{|uf - vg|}{|u,v|}.$$

We also define U and V by $u = u_0 x^{m-r} + U$, $v = v_0 x^{n-r} + V$. Therefore, $|U,V| \leq 1$. We shall show that the two leading coefficients u_0 and v_0 of u and v cannot be too small and hence do not vanish. More precisely, we have the following bounds.

Lemma 8. *With the above notation,*

$$\max\{|u_0|, |v_0|\} \geq \frac{\tau_r - \tau_{r-1}}{2 + \tau_r^2} \quad \textit{and} \quad \min\left\{\frac{|U|}{|u_0|}, \frac{|V|}{|v_0|}\right\} \leq \frac{2 + \tau_r^2}{\tau_r - \tau_{r-1}}.$$

Proof. We have $|U|^2 = |u|^2 - |u_0|^2$, $|V|^2 = |v|^2 - |v_0|^2$, $uf - vg = (Uf - Vg) + (u_0 x^{m-r} f - v_0 x^{n-r} g)$ and $|u_0 f_0 - v_0 g_0| = |T_0| \leq \tau_{r-1}$. Without loss of generality, we suppose $|u_0| \geq |v_0|$ so it suffices, for the first claim, to show that $|u_0|$ satisfies the bound.

Then $|U,V|^2 = 1 - |u_0|^2 - |v_0|^2 \geq 1 - 2|u_0|^2$. Moreover, since U and V have degree one less than u and v, the definition of τ_r (see Section 4) implies

$$\tau_r \leq \frac{|Uf + Vg|}{|U,V|} \Rightarrow |Uf - Vg| \geq \tau_r |U,V|.$$

We distinguish two cases on whether $\tau_{r-1} \geq \tau_r |U,V|$ or not. In the first case, $\tau_{r-1} \geq \tau_r |U,V| \geq \sqrt{1 - 2|u_0|^2}\, \tau_r$ which means

$$|u_0|^2 \geq \frac{\tau_r^2 - \tau_{r-1}^2}{2\tau_r^2} \Rightarrow |u_0| \geq \frac{\tau_r - \tau_{r-1}}{\sqrt{2}\tau_r} \geq \frac{\tau_r - \tau_{r-1}}{2 + \tau_r^2}.$$

So $|u_0|$ satisfies the first inequality in this case.

In the second case, $\tau_{r-1} < \tau_r |U,V|$, therefore $|uf - vg| < |Uf - Vg|$, hence we may write $|u_0 x^{m-r} f - v_0 x^{n-r} g| \geq |Uf - Vg| - \tau_{r-1}$ and we get

$$\max\{|u_0|, |v_0|\} \geq \tfrac{1}{2}(|Uf - Vg| - \tau_{r-1}) \Rightarrow 2\max\{|u_0|, |v_0|\} \geq |U,V|\tau_r - \tau_{r-1},$$

where the last step applies the triangular inequality and hypothesis $\tau_{r-1} < \tau_r |U, V|$. As $|u_0| \geq |v_0|$ then we have

$$2|u_0| \geq \sqrt{1 - 2|u_0|^2}\ \tau_r - \tau_{r-1},$$

where the right-hand side is non-negative by hypothesis. The second degree inequality implies

$$|u_0|^2(4 + 2\tau_r^2) + 4|u_0|\tau_{r-1} + \tau_{r-1}^2 - \tau_r^2 \geq 0$$

$$\Rightarrow |u_0| \geq \frac{-2\tau_{r-1} + \tau_r\sqrt{4 - 2\tau_{r-1}^2 + 2\tau_r^2}}{4 + 2\tau_r^2}.$$

Now,

$$4 - 2\tau_{r-1}^2 + 2\tau_r^2 \geq 4 \Rightarrow |u_0| \geq \frac{\tau_r - \tau_{r-1}}{2 + \tau_r^2}$$

and the first claim is established. To establish the second inequality assume, without loss of generality, that $|u_0| \geq |v_0|$ and plug the corresponding bound in

$$\frac{|U|}{|u_0|} \leq \frac{1}{|u_0|} \leq \frac{2 + \tau_r^2}{\tau_r - \tau_{r-1}}.$$

Therefore, one of $|U|/|u_0|$ and $|V|/|v_0|$ must be upper bounded by the latter function of τ_r, τ_{r-1}. □

Therefore, the degrees of u, v are, respectively, $m - r$ and $n - r$. Let us denote by B the following function of τ_r, τ_{r-1}:

$$B = 1 + \frac{2 + \tau_r^2}{\tau_r - \tau_{r-1}}.$$

The next step consists in a division that defines Q and R. Whether the divisor is u or v depends on whose leading coefficient is larger. Without loss of generality, we assume in the sequel that $|U|/|u_0| \leq |V|/|v_0|$. If not, an analogous approach leads to similar results and algorithms by starting with v instead of u.

Lemma 9. *For $|U|/|u_0| \leq |V|/|v_0|$, there exist two polynomials Q and R such that, with the previous notation,*

$$uf - vg = T = Qu + R,$$

where $\deg(Q) \leq n$, $\deg(R) \leq m - r - 1$ *and*

$$|Q| \leq B^{n+1}\tau_{r-1}, \qquad |R| \leq B^{n+1}\tau_{r-1}.$$

Proof. We use the two previous lemmas. With $|u_0| \geq |v_0|$ we set A, M, α and β of Proposition 7 equal, respectively, to $T, B, n + m - r$ and $m - r$. □

In order to complete the proof of the main theorem, we apply the following procedure. The goal is to establish an expression of the form $R = us - vt$, where s and t are two polynomials of degree less than or equal to $n-r-1$ and $m-r-1$ respectively. Then, we shall define the deformations $\hat{f}$ and $\hat{g}$ of f and g such that

$$\hat{f} = f - Q - s, \qquad \hat{g} = g - t, \qquad \deg \hat{f} = n, \qquad \deg \hat{g} = m.$$

The exact relation $u\hat{f} - v\hat{g} = 0$ is then satisfied, thus $\hat{f}, \hat{g}$ admit an exact GCD of degree r. These are the target perturbed polynomials. Then, we estimate the norms of the perturbations $|\hat{f} - f|$, $|\hat{g} - g|$.

The fact that u, v are relatively prime is established in the following lemma. This is the crucial step in the proof of the theorem, since it implies that there exists an approximate LCM of f, g close to uf and to vg, of degree $n + m - r$.

Lemma 10. *If $\tau_r > \tau_{r-1} 2^{2n+m-2r}$, then there exists a constant $a \in \mathbb{C}$ and two unique polynomials u_1 and v_1 of degree (less than or equal to) $n-r-1$ and $m-r-1$ such that $|u_1, v_1| = \sqrt{|u_1|^2 + |v_1|^2} = 1$ and*

$$uu_1 - vv_1 = a, \qquad a \in \mathbb{C} \text{ such that } \frac{|a|}{\sqrt{2}} \geq \frac{\tau_r}{2^{2n+m-2r}} - \tau_{r-1}. \tag{6}$$

Proof. If u and v are relatively prime, the existence of polynomials u_1, v_1 and of constant a follows from Bézout's identity $uu_1 - vv_1 = a$. This also implies uniqueness and $a \neq 0$. Otherwise such a relation still exists with $a = 0$ but it is not unique. Without loss of generality, $|v| \geq |u|$, hence $|v| \geq 1/\sqrt{2}$. Multiplying Bézout's relation by f, we get $fuu_1 - fvv_1 = af$. Substituting uf by $vg + T$ we obtain $af = v(u_1 g - v_1 f) + Tu_1$.

By definition of τ_r, we have $|u_1 g - v_1 f| \geq \tau_r$. Recalling that $|f| = 1$ and $|u_1| \leq 1$, Theorem 2 implies $|Tu_1| \leq \tau_{r-1}\sqrt{2}$ and

$$|v(u_1 g - v_1 f)| \geq \frac{|v|\tau_r}{2^{2n+m-2r-1}} \Rightarrow |a| \geq \frac{\tau_r}{2^{2n+m-2r-1/2}} - \tau_{r-1}\sqrt{2},$$

where the last step relies on the hypothesis on the gap between τ_r, τ_{r-1}. We get a similar inequality if we suppose $|u| \geq |v|$ and multiply by g instead of f, namely

$$|a| \geq \frac{\tau_r}{2^{n+2m-2r-1/2}} - \tau_{r-1}\sqrt{2}.$$

By recalling that $n \geq m$ we arrive at the lower bound on $|a|$. □

Multiplying by R the Bézout relation in the previous lemma gives

$$uRu_1 - vRv_1 = aR.$$

Performing polynomial division of Ru_1/a by v, the remainder is s; in an analogous way we calculate t. This defines unique polynomials s and t of degree less than or equal to $n-r-1$, $m-r-1$. Furthermore, polynomial $R = su - vt + uv(\cdot)$ is of degree $\leq m-r-1$ hence there is no uv-term on the right-hand side of the expression,

hence $R = us - vt$. The algorithm in the next section computes directly s, t from the latter Bézout relationship without computing u_1, v_1 and a. These three quantities were introduced solely for the purposes of exposition.

Lemma 11. *Suppose* $\tau_r > \tau_{r-1} 2^{2n+m-2r}$. *Let* $s = (Ru_1/a) \bmod v$ *and* $t = (Rv_1/a) \bmod u$ *be polynomials of degree less than or equal to* $n - r - 1$ *and* $m - r - 1$, *respectively. They satisfy* $R = us - vt$, *where* R *is defined in Lemma 9. Moreover,*

$$\text{both } |s| \text{ and } |t| \text{ are upper bounded by } B^{n+1}\tau_{r-1} \Big/ \left(\frac{\tau_r}{2^{2n+m-2r}} - \tau_{r-1}\right).$$

Proof. The equivalence of the two definitions of s, t follows from the previous discussion. To bound the norms, the proof resembles to that of the previous lemma. Substitute $uf = T + vg$ into $Rf = usf - vtf$ to get $Rf = s(T + vg) - vtf$. Thus,

$$|Rf| = |v(sg - tf) + Ts| \geq |(|v(sg - tf)| - |Ts|)| .$$

This implies, given that $|f| = 1$ and assuming $|v(sg - tf)| \geq |Ts|$,

$$|R|\sqrt{2} \geq |v(sg - tf)| - |Ts| \geq \frac{|v||sg - tf|}{2^{2n+m-2r-1}} - \tau_{r-1}|s,t|\sqrt{2},$$

by repeated application of both bounds of Theorem 2. Assuming that $|v| \geq |u|$ we can lower bound the former by $1/\sqrt{2}$. By definition of τ_r, $\tau_r \leq |sg - tf|/|s,t|$. Therefore,

$$|R| \geq |s,t| \left(\frac{\tau_r}{2^{2n+m-2r}} - \tau_{r-1}\right) \Rightarrow |s,t| \leq |R| \Big/ \left(\frac{\tau_r}{2^{2n+m-2r}} - \tau_{r-1}\right).$$

The expression in the first parenthesis provides a lower bound on $|v(sg - tf)| - |Ts|$ which is positive by the lemma's hypothesis, hence the previous assumption that $|v(sg - tf)| \geq |Ts|$ is valid. If, on the other hand, $|u| \geq |v|$, then an analogous argument gives the same bound except that the exponent of 2 is $n + 2m - 2r$. Keeping the largest exponent and bounding $|R|$ by Lemma 9 yields the result. □

Summarizing we have an approximate LCM equal to $u(f - Q - s) = v(g - t)$ and

$$\varepsilon\text{-gcd} = \frac{g - t}{u} = \frac{f - Q - s}{v}.$$

Furthermore, the perturbations are bounded as follows:

$$|\hat{f} - f| \leq |Q| + |s| \leq B^{n+1} \left(1 + \frac{2^{2n+m-2r}}{\tau_r - \tau_{r-1}2^{2n+m-2r}}\right) \tau_{r-1},$$

$$|\hat{g} - g| = |t| \leq B^{n+1} \frac{2^{2n+m-2r}}{\tau_r - \tau_{r-1}2^{2n+m-2r}} \tau_{r-1}.$$

The following quantity provides a bound on both perturbations as a function of $m, n, r, \tau_{r-1}, \tau_r$:

$$E = B^{n+1}\left(1 + \frac{2^{2n+m-2r+1/2}}{\tau_r - \tau_{r-1}2^{2n+m-2r+1/2}}\right)\tau_{r-1}.$$

This discussion proves the main certification theorem.

Theorem 6. *Suppose that* $\tau_{r-1} \le \varepsilon \le \tau_r/\sqrt{2}$ *for* $\varepsilon \le 1$, $1 \le r \le m \le n$, *and that* $\tau_r > \tau_{r-1}2^{2n+m-2r}$. *Then, with the above notation,*

$$E \le \varepsilon \Rightarrow \deg \varepsilon\text{-gcd}(f, g) = r.$$

The main limitation of this result is the exponential factor in the required gap, which reduces the practical significance of the theorem. Notice, however, that this is the only available guarantee and that in real-world situations the norms should not attain these worst-case bounds.

A simpler version of the gap is now given, thus clarifying its dependence on the various quantities. Typically, ε is so small that τ_r can be upper bounded by a constant somewhat smaller than unity.

Corollary 12. *Suppose that* $\tau_{r-1} \le \varepsilon \le \tau_r/\sqrt{2}$, $\tau_r \le 23/25$ *and* $1 \le r \le m \le n$. *Then*

$$\tau_{r-1} \le 2^{-5n-3}\tau_r^{n+2}\varepsilon \Rightarrow \deg \varepsilon\text{-gcd}(f, g) = r.$$

Proof. We demonstrate that the conditions of the main theorem are satisfied. The present gap implies

$$\tau_{r-1} < \tau_r 2^{-4n-1} \tag{7}$$

and the gap of the main theorem follows. Moreover, $\tau_{r-1} < \tau_r/32$, since $n \ge 1$. This is used for upper bounding B.

$$B < 1 + \frac{2 + \tau_r^2}{31\tau_r/32} < \frac{4}{\tau_r},$$

where the last inequality used the bound on τ_r. Another consequence of bound (7) is

$$\tau_{r-1}2^{2n+m-2r} < \tau_r 2^{-n-1} \Rightarrow \tau_r - \tau_{r-1}2^{2n+m-2r} > \frac{3\tau_r}{4}.$$

This is used in upper bounding E by the new bound on ε.

$$E < \left(\frac{4}{\tau_r}\right)^{n+1}\left(1 + \frac{2^{3n}}{3\tau_r/4}\right)\tau_{r-1} \le \frac{2^{2n+2}}{3\tau_r^{n+2}}\,\frac{12}{11}2^{3n+2}\,\tau_{r-1} < \frac{2^{5n+3}}{\tau_r^{n+2}}\tau_{r-1}. \qquad \square$$

Here the linear dependence of the gap on τ_{r-1} is made obvious, in contrast to the gap in [8] where the dependence was polynomial. Yet, there still remains the polynomial dependence on τ_r and the exponential dependence on n, both of which may provide some room for improvement by the use of weighted norms.

7. The algorithms

The main theorem and its proof lead directly to an algorithm. An important step is the use of an SVD to define an approximate syzygy $uf + vg$ of minimum norm for given polynomials f, g. Any subresultant matrix Sy_k can be written as a product $A\Sigma B$ where A, B are square orthogonal matrices and Σ is diagonal with the dimensions of S_k. The last row of B contains the coefficients of polynomial pair (u, v) with degrees bounded by $n-k-1$, $m-k-1$ and unit pair norm. The $(n+m-2k)$th column of A expresses the syzygy polynomial of degree bounded by $n+m-1$ divided by the minimum singular value σ_{m+n-2k}. In other words, polynomial $uf + vg$ has norm σ_{m+n-2k} which is the minimum norm of any polynomial in the image of Sy_k; see Section 3.

Algorithm 1. Given polynomials $f \in \mathscr{P}_n$, $g \in \mathscr{P}_m$ and tolerance $\varepsilon > 0$, the following procedure computes the ε-GCD of f and g under the hypotheses of Theorem 6 and the perturbed polynomials for which this is an exact GCD.

(i) If $\tau_0 > \varepsilon$ then output 1 and terminate. Initialize r to 1. While the minimum singular value τ_r of subresultant matrix $\mathrm{Sy}_r(f, g)$ is smaller than $\varepsilon\sqrt{2}$, increment r. In practice, we may not have to compute the singular values of all subresultant matrices, since a rather accurate indication of r is given simply by observing the singular values of the Sylvester matrix.

(ii) Compute constant E defined in Theorem 6. If $\tau_r \leq \tau_{r-1} 2^{2n+m-2r}$ or $E < \varepsilon$ then the hypotheses of Theorem 6 are not satisfied and we are not able to guarantee the success of this procedure a priori. The algorithm continues and tests the norm of the computed perturbations at the end.

(iii) Apply an SVD to write $\mathrm{Sy}_{r-1} = A\Sigma B$, where A, B are orthogonal and Σ diagonal. Read off pair (u, v) from the last row of B by inverting the coefficients of v. Obtain polynomial T by multiplying the polynomial in the $(n+m-2r+2)$th column of A by τ_{r-1}, thus satisfying $uf - vg = T$ and $|T| = \tau_{r-1}$.

(iv) The rest of the algorithm relies on the condition $|U|/|u_0| \leq B - 1$; if this does not hold, then exchange u, v and f, g. Apply standard polynomial division of T by u to compute Q as the quotient and R as the remainder, as in Lemma 9, i.e., $T = uQ + R$.

(v) Compute s and t such that $R = us - vt$ by solving for X, whose image under the Sylvester matrix transformation $\mathrm{Sy}_0(u, v)$ is vector Y expressing polynomial R. Polynomial pair (s, t) is expressed by the solution X of $\mathrm{Sy}_0(u, v)X = Y$.

(vi) Compute polynomials $\hat{f} = f - Q - s$ and $\hat{g} = g - t$, which have an exact GCD of degree r. If any of $|\hat{f} - f|$ and $|\hat{g} - g|$ are larger than ε, then the algorithm terminates unsuccessfully. Otherwise, polynomials $\hat{f}$, $\hat{g}$ lie within the prescribed tolerance and are, therefore, output. The approximate LCM is $u(f - Q - s)$ or $v(g - t)$ and the algorithm returns

$$\varepsilon\text{-gcd} = \frac{g - t}{u}$$

which is equal to $(f - Q - s)/v$ within the error of numerical calculation.

A note on numerical stability. In step (iii) the result depends solely on the stability of computing singular values and not on the stability of computing the singular vectors. In other words, we obtain polynomials u, v, T with $uf - vg = T$ and $|T|$ very close to τ_{r-1} even if (u, v) is not exactly the best approximate syzygy.

An alternative computation, once the degree is determined, is to perform some optimization process to compute the coefficients of the approximate GCD which minimize the perturbation achieving the computed degree. If a gap is not found for any r, then the best we can obtain is a lower and upper bound on the degree; this is the case of an unstable situation. A lower bound is obtainable by Euclid's algorithm, as explained below. Upper bounds are given by Proposition 5 under certain hypotheses on the singular values of the subresultant matrices.

In order to find the approximate GCD polynomial, either optimization or Euclid's algorithm can be used to yield a sequence of GCDs, each associated with a tolerance. In [8, Section 3] the latter approach was studied in depth and different variants were described based on the plain, pseudo-division and subresultant polynomial remainder sequences. Here we sketch briefly the first approach for the sake of completeness. The following is essentially the extended euclidean algorithm.

Algorithm 2. Given polynomials $f \in \mathscr{P}_n, g \in \mathscr{P}_m$ and tolerance $\varepsilon > 0$, the following procedure computes a polynomial that divides exactly $\hat{f} \in \mathscr{P}_n$, $\hat{g} \in \mathscr{P}_m$, such that $|f - \hat{f}|$, $|g - \hat{g}| \leq \varepsilon$.

(i) Initialize $F_1 = f$, $F_2 = g$ and $j = 2$.

(ii) Division step: Compute Q_j and F_{j+1} such that $F_{j-1} = Q_j F_j + F_{j+1}$. Also compute $R_j^{(i)}$, $i = 1, 2$, such that $R_j^{(i)} = Q_j R_{j-1}^{(i)} + R_{j-2}^{(i)}$, $i = 1, 2$.

(iii) Termination test: If the maximum of $|R_{j-1}^{(1)} F_{j+1}|$ and $|R_{j-1}^{(2)} F_{j+1}|$ is bounded by ε, then return F_j. Otherwise increment j and go to step (ii).

Proposition 13 (Emiris et al. [8, Proposition 3.3]). *Suppose that after executing the above algorithm on polynomials F_1, F_2, we obtain polynomials F_r and $R_{r-2}^{(i)}$ for $i = 1, 2$, such that $|R_{r-2}^{(i)} F_r| \leq \varepsilon$, $i = 1, 2$. Then*

$$\deg(\varepsilon - \gcd(F_1, F_2)) \geq \deg F_{r-1}.$$

Let $d_i = \deg F_r + \deg R_{r-2}^{(i)}$, $i = 1, 2$, which means that for $r > 2$, $d_1 = \deg F_r + \deg F_{r-2} - \deg F_2$ and $d_2 = \deg F_r + \deg F_{r-2} - \deg F_3$. Then,

$$|F_r| \leq \varepsilon \cdot \min_{i=1,2} \left\{ \frac{2^{d_i}(d_i + 1)^{3/2}}{|R_{r-2}^{(i)}|} \right\}.$$

The above algorithm, just as any other variant of Euclid's algorithm, will only produce a lower bound on the degree of the ε-GCD. Note also that the test condition does not change monotonically with the candidate degree, so the user should choose the best GCD candidate after inspecting the output for all possible degrees.

The input may be ill-conditioned with respect to our main theorem in the sense that its hypothesis may be unsatisfied, yet the upper and lower bound on the approximate degree can coincide. In this case the optimal degree is computed and a valid GCD polynomial is the one found by the extended euclidean Algorithm 2.

The arithmetic complexity of the algorithms is polynomial in the degrees of the given polynomial and the degree of the output ε-gcd.

MAPLE code for all operations is available from the first author.

8. Weighted norms

As we have seen in Section 3, weighted norms possess several advantages. This section indicates how these advantages can be exploited and provides the machinery for sharpening the certification theorem.

First, weighted norms are almost multiplicative with relatively small multiplicative constants. Second, they are invariant by unitary changes of coordinates. The use of the first property is straightforward because it suffices to replace the powers of 2 or the factorials by a smaller binomial coefficient in the degrees.

The use of the second property is more elaborate, since we have to perform a unitary change of coordinates in order to increase the minimum quotient value between $|U|/|u_0|$ and $|V|/|v_0|$. Eventually, in the exposition of the previous Section, one should adapt the given estimates of $|U|/|u_0|, |V|/|v_0|$ by replacing e.g. the equality $|U|^2 = |u|^2 - |u_0|^2$ by the weighted one. In fact, the leading coefficient u_0 (respectively v_0) is the value "at infinity" of u (respectively v). More precisely, we choose the unitary change of variables so that it maximizes the maximum of the leading coefficients in the image, under this transformation, of u and v.

The new (geometric) setting is more intrinsic since it corresponds to the pair of sets of roots of f and g on the projective complex line. In particular, we wish to consider the metric space of one-dimensional subspaces of a hermitian plane, endowed with the following natural metric: the distance between the complex vector lines generated by vectors a and b equals the scalar product $a^{\mathrm{T}}b$ divided by the product of the vector (hermitian) norms. This metric space is isometric to a euclidean 2-dimensional sphere. Unitary transformations of the initial hermitian plane give all direct isometries of the euclidean sphere. In a numerical setting, it is necessary to put a metric on the Riemann sphere and the way we have just indicated is the most natural one.

Fig. 2 shows how the distances between roots change when they are mapped from affine to projective space. It gives a rough indication that projective space allows us to take advantage of the change of variables.

To use weighted norms in practice, we write the subresultant mappings of Section 4. Recall that $\deg f = n, \deg g = m$,

$$\mathrm{Sy}_r(f,g) : (u,v) \mapsto uf + vg, \qquad \deg u \leq m - r - 1, \quad \deg v \leq n - r - 1.$$

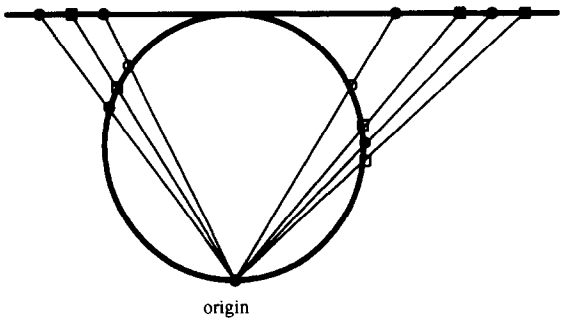

Fig. 2. The straight line represents affine space and the circle represents projective space. The correspondences between the roots in the two spaces is shown; squares and little circles label the roots of the two polynomials.

In a *weighted monomial basis* the new matrix $\mathrm{Sw}_r(f,g)$ is obtained from $\mathrm{Sy}_r(f,g)$ by

multiplying column j corresponding to f, for $1 \le j \le m-r$ by $\sqrt{\binom{m-r-1}{j-1}}$,

multiplying column $m-r+j$ corresponding to g, for $1 \le j \le n-r$ by $\sqrt{\binom{n-r-1}{j-1}}$,

dividing row i, $1 \le i \le m+n-r$ by $\sqrt{\binom{m+n-r-1}{i-1}}$.

The new sequence of minimum singular values τ_i comes from these adapted matrices and has analogous properties, namely, the properties of Section 4 become:

- $\mathrm{Sw}_0(f,g)$ is of rank $m+n-r$ if and only if $\deg(\gcd(f,g))=r$.
- $\mathrm{Sw}_r(f,g)$ has full rank i.e. $m+n-2r$ if and only if $\deg(\gcd(f,g)) \le r$.

The intrinsic significance of the new singular values relies on the fact that they remain invariant subject to unitary change of variables as defined in (2).

Routines on MAPLE implementing the basic weighted-norms routines, including the subresultant matrix in the weighted basis, are available upon request from the first author.

9. Conclusion

Extending to systems of polynomials is an important open question. We may consider the generalization of the Sylvester matrix, i.e., the multivariate resultant as given by classical elimination theory or, if only affine roots are interesting, the more recent sparse resultant. In any case, different matrix formulations exist which should be compared from a seminumerical point of view. Two examples include Newton matrices [7], which generalize Macaulay's matrix, and Bézout–Dixon matrices.

Another issue concerns the case of real polynomials f, g and whether their ε-GCD has real coefficients or not. Interestingly, algebraic procedures such as Euclid's algorithm and SVD yield answers in the real space, whereas an optimization problem may lead to complex coefficients; see the formal geometric setting discussed in [8] in terms of complex varieties. A complex common solution implies the existence of another complex solution, namely its conjugate, which corresponds to the same minimum distance.

It was supposed that roundoff error is very small. A more careful study would attempt to incorporate computational accuracy, especially when this is significant with respect to ε. This would assert the cost attached to increasing the precision so that the error is negligible for the specific computation.

Acknowledgements

We thank the referees for suggesting several improvements in the presentation. All authors partially supported by European ESPRIT project FRISCO (Reactive LTR 21.024).

References

[1] Z. Bai, J. Demmel and A. McKenney, On the conditioning of the nonsymmetric eigenproblem: theory and software, Tech. Report CS-89-86, Univ. of Tennessee, October 1989.
[2] B. Beauzamy, Products of polynomials and a priori estimates for coefficients in polynomial decompositions: a sharp result, J. Symbol. Comput. 13 (1992) 463–472.
[3] B. Beauzamy, E. Bombieri, P. Enflo and H.L. Montgomery, Products of polynomials in many variables, J. Number Theory 36 (1990) 219–245.
[4] W.S. Brown, On Euclid's algorithm and the computation of polynomial greatest common divisors, J. ACM 18 (1971) 476–504.
[5] G.E. Collins, Subresultants and reduced polynomial remainder sequences, J. ACM 14 (1967) 128–142.
[6] R.M. Corless, P.M. Gianni, B.M. Trager and S.M. Watt, The singular value decomposition for polynomial systems, in: Proc. ACM Internat. Symp. on Symbolic and Algebraic Computation (1995) 195–207.
[7] I.Z. Emiris and J.F. Canny, Efficient incremental algorithms for the sparse resultant and the mixed volume, J. Symbol. Comput. 20 (1995) 117–149.
[8] I.Z. Emiris, A. Galligo and H. Lombardi, Numerical univariate polynomial GCD, in: J. Renegar, M. Shub and S. Smale, eds., The Mathematics of Numerical Analysis, Lectures in Applied Mathematics, Vol. 32 (AMS, Providence, RI, 1996) 323–343.
[9] R.T. Farouki and V.T. Rajan, On the numerical condition of polynomials in Bernstein form, Comput.-Aided Des. 4 (1987) 191–216.
[10] O.D. Faugeras, Three-Dimensional Computer Vision: a Geometric Viewpoint (MIT Press, Cambridge, MA, 1993).
[11] A.O. Gelfond, Transcendental and Algebraic Numbers (Dover, New York, 1960).
[12] G.H. Golub and C.F. van Loan, Matrix Computations (The Johns Hopkins University Press, Baltimore, ML, 1989).
[13] L. Gonzalez-Vega, H. Lombardi, T. Recio and M.-F. Roy, Determinants and real roots of univariate polynomials, in: Proc. 25 Years of Quantifier Elimination and Cylindrical Algebraic Decomposition, Linz, October 1993, Texts and Monographs in Symbolic Computation (Springer, Berlin, 1995).
[14] W. Habicht, Zur inhomogenen Eliminationstheorie, Commut. Math. Helvetici 21 (1948) 79–98.

[15] D.G. Hough, Explaining and Ameliorating the Ill Condition of Zeros of Polynomials, Ph.D. Thesis, Comp. Science Div., Univ. of California, Berkeley, 1977.
[16] V. Hribernig and H.J. Stetter, Detection and validation of clusters of polynomial zeros, J. Symbolic Computation, 1996. Special Issue on Validation and Symb. Computing, to appear.
[17] N. Karkanias and M. Mitrouli, A matrix pencil based numerical method for the computation of the GCD of polynomials, IEEE Trans. Automat. Control 39 (1994) 977–981.
[18] N. Karmarkar and Y.N. Lakshman, Approximate polynomial greatest common divisors and nearest singular polynomials, in: Proc. ACM Internat Symp. on Symbolic and Algebraic Computation (1996).
[19] D. Lazard, Résolution des systèmes d'équations algébriques, Theoret. Comput. Sci. 15 (1981) 77–110.
[20] J.-P. Merlet, Les Robots Parallèles, Traités de Nouvelles Technologiques (Paris, Hermès, 1990).
[21] M. Mignotte, An inequality about factors of polynomials, Math. Comput. 28 (1974) 1153–1157.
[22] M.-T. Noda and T. Sasaki, Approximate GCD and its application to ill-conditioned algebraic equations, J. Comput. Appl. Math. 38 (1991) 335–351.
[23] V.Y. Pan, Numerical computation of a polynomial GCD and extensions, manuscript, Technical Report 2969, INRIA, Sophia-Antipolis, France, 1996.
[24] A. Schönhage, Quasi-GCD computations, J. Complexity 1 (1985) 118–137.
[25] T.W. Sederberg and G.-Z. Chang, Best linear common divisors for approximate degree reduction, Comput.-Aided Des. 25 (1993) 163–168.
[26] H.J. Stetter, Analysis of zero clusters in multivariate polynomial systems, in: Proc. ACM Intern. Symp. on Symbolic and Algebraic Computation (1996).
[27] J.J. Sylvester, On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraic common measure, Philosophical Trans. 143 (1853) 407–548.
[28] E.K. Wakeford, Apolarity and canonical forms, Proc. London Math. Soc. 18 (1918–1919) 403–410.