POLYNOMIALS AND LEMNISCATES OF INDEFINITENESS

MARKO HUHTANEN* AND OLAVI NEVANLINNA†

Abstract. For a large indefinite linear system, there exists the option to directly precondition for the normal equations. Matrix nearness problems are formulated to assess the attractiveness of this alternative. Polynomial preconditioning leads to polynomial approximation problems involving lemniscate-like sets, both in the plane and in $\mathbb{C}^{n\times n}$. A natural matrix analytic extension for lemniscates is introduced. Operator theoretically one is concerned with polynomial unitarity and associated factorizations for the inverse. For the speed of convergence and lemniscate asymptotics, the notion of quasilemniscate arises. In the L^2-norm algorithms for solving the problem are devised.

Key words. indefiniteness, preconditioning, normal equations, polynomial preconditioning, lemniscate, polynomially unitary operator, GMRES

AMS subject classifications. 41A10, 47A10, 65F08

1. Introduction.

Iterative methods of numerical linear algebra for solving large linear systems and eigenvalue problems are connected with classical approximation theory in many ways [24, 11, 6, 21, 17]. In this paper an attempt is made to do this for indefiniteness. A large linear system involving a matrix $A \in \mathbb{C}^{n\times n}$ can be regarded as being indefinite if one is close to solving the normal equations with iterative methods which try to avoid using them; see [2] for indefinite problems and [29, Chapter 8] on using the normal equations. This is a well-known manifestation of indefiniteness, at least for Hermitian problems. The so-called generalized saddle point problems provide a large pool of non-Hermitian indefinite problems; see [2, p.4] and references therein. In deciding whether one actually should consider using the normal equations, one is led to inspect the matrix nearness problem

$$\min_{M\in\mathcal{V},\, B\in\mathcal{U}_k} \|AM - B\|$$

for $k \ll n$.[1] Here the matrix subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ models a chosen preconditioning strategy while $\mathcal{U}_k$ denotes the rank-k neighborhood of unitary matrices. This paper is concerned with polynomial preconditioning and $k = 0$ corresponding to unitary matrices $\mathcal{U}$.

In polynomial preconditioning $\mathcal{V}$ consists of polynomials in A of degree $j-1$ at most, so that the problem becomes that of finding

$$l_j(A) = \min_{p\in\mathcal{P}_{j-1},\, U\in\mathcal{U}} \|Ap(A) - U\|,$$

where $\mathcal{P}_{j-1}$ denotes the set of polynomials of degree $j-1$ at most. To analyze this, denote by Λ the spectrum of A. To optimally form the normal equations, one invariably ends up studying the polynomial approximation problem

$$l_j(\Lambda) = \min_{p\in\mathcal{P}_{j-1}} \max_{\lambda\in\Lambda} \big|\, |\lambda p(\lambda)| - 1 \big|\,; \tag{1.1}$$

* Department of Mathematical Sciences, University of Oulu, 90570 Oulu 57, Finland, (Marko.Huhtanen@aalto.fi).

† Department of Mathematics and Systems Analysis, Aalto University, P.O.Box 11100, FIN-00076 Aalto, Finland, (Olavi.Nevanlinna@aalto.fi).

[1] In preconditioning for the normal equations there are two options. Either first form AA^* and thereafter precondition this positive definite matrix. Or, like in this paper, first precondition A and thereafter form the normal equations.

see [10, Section 5] for the Hermitian indefinite case. To contrast this with the polynomial approximation problem of the GMRES (generalized minimal residual) method, obviously

$$l_j(\Lambda) \leq \min_{p\in\mathcal{P}_{j-1}} \max_{\lambda\in\Lambda} |\,\lambda p(\lambda) - 1| \tag{1.2}$$

holds. Plane geometrically $l_j(\Lambda)$ can be interpreted as describing how well Λ can be "circled" under a polynomial map of degree j having a zero at the origin. Attaining zero corresponds to Λ being a subset of a lemniscate with a locus at the origin. Of course, study of lemniscates and associated extremal problems is part of classical polynomial analysis and potential theory [30, 9, 3]. See also [4].

Here it is shown that lemniscate-like sets appear naturally in matrix (operator) and spectral theory as well. The so-called polynomial unitarity leads to an extension of the notion of algebraic operator and related factorizations for the inverse; see Definition 2.3 for the notion of polynomially unitary. The concept of lemniscate turns out have an obvious matrix theoretic extension to $\mathbb{C}^{n\times n}$, reducing to the classical planar lemniscate in the scalar case $n = 1$. Then the associated "lemniscate asymptotics" is concerned with the decay of $l_j(\Lambda)$ as j grows. In particular, it is shown that there holds $l_j(\Lambda) \leq l_j(A)$. It is also shown that

$$l_j(A) = \min_{p\in\mathcal{P}_{j-1}} \max_{\|x\|=1} |\|Ap(A)x\| - 1|, \tag{1.3}$$

providing an alternative way of measuring the deviation of A from being polynomially unitary of degree j. The quantity $l_j(A)$ is unitarily invariant such that if A is normal, then $l_j(\Lambda) = l_j(A)$ holds.

Since the spectrum is rarely known exactly, or the dimension n is not known in advance, assume $\Lambda \subset \mathbb{C}$ is compact containing the spectrum of A. Bearing in mind that the GMRES method is well-suited for definite problems (in some sense), the inequality (1.2) yields a way to quantify indefiniteness. The difference between these two scalars can be arbitrarily wide while varying between zero and one. The GMRES method, when analyzed based on Λ, can be guaranteed to converge only if 0 is not in the polynomial convex hull of Λ. However, having 0 in the polynomial convex hull is not an obstruction as such for $l_j(\Lambda)$ to attain zero. That is, now the situation is far from being so clear cut and it is an intriguing problem to provide conditions on Λ under which $\lim_{j\to\infty} l_j(\Lambda)$ vanishes. There arises a link with the Riemann mapping theorem and its extensions; see Theorem 3.3. If the convergence is rapid enough, the notion of quasilemniscate arises. For the converse, if $\lim_{j\to\infty} l_j(\Lambda) > 0$, then Λ is called severely indefinite. Also for this sufficient conditions are given.

To numerically solve the problem (1.1) in the L^2-norm, an algorithm to satisfy necessary orthogonality conditions for the solution is devised. Resulting in a differential equation and a descend method, the associated flow moves points towards polynomials satisfying these orthogonality conditions. Being straightforward to implement, a numerical experiment is given to illustrate its performance.

The paper is organized as follows. Section 2 is concerned with matrix analysis of preconditioning for the normal equations. Polynomial unitarity is defined and associated with notions such as that of algebraic operator. In this connection, lemniscates replace the role of the exact location of the spectrum. Section 3 is concerned with lemniscates having a locus at the origin, their extensions and related (asymptotic) approximation problems. The notion of quasilemniscate is introduced. In Section 4 algorithms for computing polynomials in the L^2-norm are described.

2. Indefiniteness, polynomial unitarity and lemniscates. In what follows, indefiniteness is addressed in terms of whether preconditioning for the normal equations is an attractive alternative for iteratively solving a linear system.

2.1. Idefiniteness and the normal equations. Consider iteratively solving a large linear system

$$Ax = b \tag{2.1}$$

with a nonsingular $A \in \mathbb{C}^{n\times n}$ and $b \in \mathbb{C}^n$. If one is close to solving the normal equations with methods which try to avoid using them, the problem can be regarded as being indefinite. Although not a rigorous definition, this is a well-known manifestation of indefiniteness, at least for Hermitian problems. Dating from the early days of analysis, indefiniteness was originally designed for classifying quadratic forms [13, p.112]. Modern iterative methods, however, require more flexible notions. For more rigor and flexibility, denote by $\mathcal{U} \subset \mathbb{C}^{n\times n}$ the set of unitary matrices.

DEFINITION 2.1. *The set*

$$\mathcal{U}_k = \{U + F \in \mathbb{C}^{n\times n} \, : \, U \in \mathcal{U}, \; \mathrm{rank}(F) \leq k\}$$

is said to be the rank-k neighborhood of unitary matrices.[2]

Whether one actually should consider using the normal equations, there exist the alternatives to precondition either A^*A or A. Even though A^*A has the pleasant property of being positive definite, the former option leads to a considerable loss of information. For the latter option, to model a preconditioning strategy, denote by $\mathcal{V}$ a matrix subspace of $\mathbb{C}^{n\times n}$. Then consider the matrix nearness problem

$$\min_{M\in\mathcal{V},\, B\in\mathcal{U}_k} \|AM - B\| \tag{2.2}$$

for $k \ll n$. Motivated by direct methods, the case $k = 0$ was studied in [18] where it was shown that any solution in the operator norm yields the best conditioned element of the matrix subspace $A\mathcal{V}$. For the other extreme, the case $\dim \mathcal{V} = 1$ has been completely solved (together with numerically stable algorithms) in [16].

In view of choosing an iterative method for solving the linear system (2.1), the magnitude of (2.2) is of interest. Whenever moderate for $k \ll n$, an arguable alternative is to apply the CG method on the normal equations for the preconditioned linear system

$$AMy = b. \tag{2.3}$$

Namely, as an extreme, suppose (2.2) equals zero with a nonsingular matrix $AM = U + F$. Then we have

$$(AM)^*AM = I + U^*F + F^*(AM),$$

i.e., a small rank perturbation of the identity matrix. In this case the CG method consumes at most $2\,\mathrm{rank}(F)$ iterates for the exact solution.

Regarding preconditioning, there are many $n^2/2$ dimensional alternatives for $\mathcal{V}$ to attain zero in (2.2). Of course, a realistic assumption is to accept at most $O(n)$ free

[2]This can be regarded as providing a finite dimensional analogue of the essential untarity for Hilbert space operators [23, 5].

parameters. For a classical option, consider the case of $\mathcal{V}$ being the Krylov matrix subspace

$$\mathcal{K}_j(A;I) = \operatorname{span}\{I, A, \ldots, A^{j-1}\}$$

corresponding to polynomial preconditioning the original linear system. Then (2.2) transforms into

$$\min_{p\in\mathcal{P}_{j-1},\, B\in\mathcal{U}_k} \|Ap(A) - B\|. \tag{2.4}$$

For a realistic number of free parameters, in practice $j \ll n$ is assumed.

2.2. Polynomial unitarity and lemniscates. In what follows, we are concerned with

$$l_j(A) = \min_{p\in\mathcal{P}_{j-1},\, U\in\mathcal{U}} \|Ap(A) - U\| \tag{2.5}$$

which corresponds to choosing $k = 0$ in (2.4). Then we are measuring how close to $\mathcal{U}$ one can polynomial precondition the matrix A. Although not entirely satisfactory in assessing whether one should consider preconditioning for the normal equations, this problem is involved with fundamental properties of A. First of all, when j is allowed to grow, we have a measure of nonsingularity in the following sense.

PROPOSITION 2.2. *A matrix $A \in \mathbb{C}^{n\times n}$ is invertible if and only if*

$$\min_{p\in\mathcal{P}_{j-1},\, U\in\mathcal{U}} \|Ap(A) - U\| = 0$$

for some j.

Proof. If A is singular and $\sum_{k=1}^{j} \alpha_k A^k$ is unitary for some j, then $A\sum_{k=1}^{j} \alpha_k A^{k-1}$ is nonsingular. Thereby so is A, leading to a contradiction.

For the converse, suppose A is invertible and take the minimal polynomial $p(\lambda) = \sum_{k=0}^{j} c_k\lambda^k$, with $c_j = 1$, of A, i.e., $\sum_{k=0}^{j} c_k A^k = 0$. Now A is invertible if and only if $c_0 \neq 0$. Thereby $\frac{-1}{c_0}\sum_{k=1}^{j} c_k A^k = I$ is unitary. ◻

With obvious changes, the claim of this proposition holds when (2.5) is replaced with the so-called ideal GMRES approximation problem

$$\min_{p\in\mathcal{P}_{j-1}} \|Ap(A) - I\|, \tag{2.6}$$

i.e., A is invertible if and only if zero is attained for some j. The ideal GMRES approximation problem has received a lot of attention; see [11, 6] and references therein.[3] It aims at assessing how attractive iterative solving of (2.1) is by executing the GMRES method. Regarding our problem, (2.6) obviously yields an upper bound on (2.5) by the fact that our purpose now is to reach the set of unitary matrices and not just the identity matrix in particular. It is noteworthy that if (2.5) decays slowly while j grows, then the GMRES method can be expected to be a poor choice for solving linear systems involving A. A rapid decay of (2.5) means that using the CG method on the normal equations for $Ap(A)$ is an attractive choice.

Let us now consider (2.5) when zero is attained.

[3]The ideal GMRES approximation problem is closely connected with the Chebyshev approximation problem which is concerned with minimizing $||p(A)||$ over the monic polynomials p of degree j.

DEFINITION 2.3. *A matrix $A \in \mathbb{C}^{n\times n}$ is said to be polynomially unitary of degree j if $Ap(A)$ is unitary for a polynomial p of the least possible degree $j-1$.*

The more general notion of polynomial normality is defined in an analogous way [20, 15].

DEFINITION 2.4. *A matrix $A \in \mathbb{C}^{n\times n}$ is polynomially normal of degree j if there exists a (nonzero) polynomial p of the least possible degree $j-1$ such that $Ap(A)$ is normal.*

These are both unitarily invariant notions. They are related by the fact there exists an algorithm for computing a (nonzero) polynomial p of the least possible degree such that $Ap(A)$ is normal [15, Section 5.]. Then, to make $Ap(A)$ unitary, find a polynomial q with a vanishing constant term such that the spectrum of $q(Ap(A))$ is a subset of the unit circle. Of course, although straightforward, this construction cannot be expected to be optimal.

If A is polynomially unitary of degree j, then its inverse can be explicitly factored as

$$A^{-1} = p(A)U^*, \tag{2.7}$$

where $Ap(A) = U$ is unitary for some $p \in \mathcal{P}_{j-1}$. Being somehow able to explicitly represent the inverse is exceptional. In this sense polynomially unitary matrices of degree j generalize the set of invertible algebraic matrices of degree j corresponding to the special case of $U = I$.[4] (Bear in mind that the success of the GMRES method is based on an inexpensive construction of a low degree polynomial in A approximating the inverse of A, possibly locally in a Krylov subspace.) For the latter, the spectrum consists of at most j distinct points whereas for polynomially unitary operators we have the following relaxation.

THEOREM 2.5. *Let $A \in \mathbb{C}^{n\times n}$. Then for any eigenvalue λ of A and any polynomial p there holds*

$$||\lambda p(\lambda)| - 1| \le \min_{U\in\mathcal{U}} \|Ap(A) - U\| = \max_{\|x\|=1} |\|Ap(A)x\| - 1|. \tag{2.8}$$

Proof. The distance of a matrix $M \in \mathbb{C}^{n\times n}$ to the set of unitary matrices is given by $\max\{|\sigma_1 - 1|, |\sigma_n - 1|\}$, where σ_j denote the singular values of M; see, e.g., [14, p.454]. Let $M = Ap(A)$ with the singular value decomposition $M = Q_1\Sigma Q_2^*$, where Q_1 and Q_2 are unitary. Then $\max_{\|x\|=1} |\|Ap(A)x\| - 1| = \max_{\|x\|=1} |\|\Sigma x\| - 1|$. This equals $\max\{|\sigma_1 - 1|, |\sigma_n - 1|\}$, yielding the claim concerning the equality.

The inequality follows from the singular value inequalities for the eigenvalues, i.e., there holds $\sigma_n \le |\lambda p(\lambda)| \le \sigma_1$. □

Consequently, if (2.7) holds, we may only conclude that the spectrum of A is a subset of the lemniscate

$$\{\lambda \in \mathbb{C} : \ |\lambda p(\lambda)| = 1\},$$

i.e., a continuum.

The equality in (2.8) allows us to alternatively define $l_j(A)$ according to (1.3). It is attractive since, being exclusively based on the usage of norm, it extends to Banach spaces in a natural way.

[4] An invertible matrix (operator) A is algebraic of degree j if and only if $q(A) = 0$ for a monic polynomial $q(\lambda) = \lambda^j + c_{j-1}\lambda^{j-1} + \cdots + c_0$ with $c_0 \neq 0$. Whence $A^{-1} = p(A)$ with $p(\lambda) = \frac{-1}{c_0\lambda}(q(\lambda) - c_0)$.

COROLLARY 2.6. *Denote by Λ the spectrum of A. Then*

$$l_j(\Lambda) \leq l_j(A) \tag{2.9}$$

such that equality holds if A is normal.

Proof. For any $U \in \mathcal{U}$, there exists a polynomial p realizing $\min_{p\in\mathcal{P}_{j-1}} \|Ap(A) - U\|$. By compactness of $\mathcal{U}$, there exists U realizing (2.5). Consequently, the minimum in (2.5) is attained for some $p \in \mathcal{P}_{j-1}$ and $U \in \mathcal{U}$. Then we have

$$l_j(\Lambda) = \min_{q\in\mathcal{P}_{j-1}} \max_{\lambda\in\Lambda} ||\lambda q(\lambda)| - 1| \leq \max_{\lambda\in\Lambda} ||\lambda p(\lambda)| - 1| \leq \min_{U\in\mathcal{U}} \|Ap(A) - U\| = l_j(A).$$

Suppose A is normal with a unitary diagonalization $A = V\Lambda V^*$. For any $p \in \mathcal{P}_{j-1}$, let λ be such that $|\lambda p(\lambda)| - 1| = \max_{\lambda\in\Lambda} |\lambda p(\lambda)| - 1|$. Then equality holds in (2.8) by the fact that Σ is given by taking the absolute values of the entries of $p(\Lambda)$. ▯

In the nonnormal case the gap between $l_j(\Lambda)$ and $l_j(A)$ can be wide.

EXAMPLE 1. Let $A \in \mathbb{C}^{n\times n}$ be a nonsingular Jordan block. Since the spectrum consists of a single point, we have $l_j(\Lambda) = 0$ already for $j = 1$. Now the spectrum of $Ap(A)$ consists of a single point for any polynomial p. This means attaining a unitary matrix only when $Ap(A) = I$, i.e., when $q(\lambda) = \lambda p(\lambda) - 1$ is a multiple of the characteristic polynomial of A. Therefore $l_{n-1}(A) > 0$.

Although the gap can be wide, the location of spectrum retains its importance. For this, consider a nonnormal but diagonalizable case. If the eigenvalues are located on the unit circle, a mere triangular similarity suffices to transform such matrices to unitary.

PROPOSITION 2.7. *Suppose the eigenvalues of $M \in \mathbb{C}^{n\times n}$ are located on the unit circle. If M is diagonalizable, then there exists a lower triangular matrix $L \in \mathbb{C}^{n\times n}$ such that LML^{-1} is unitary.*

Proof. By the assumptions, there exists an invertible matrix $X \in \mathbb{C}^{n\times n}$ such that XMX^{-1} is unitary. Thus $X^{-*}M^*X^*XMX^{-1} = I$, i.e., $M^*HM = H$ for a positive definite matrix $H = X^*X$. Compute the Cholesky factorization $H = LL^*$ of H. Then repeat the steps backwards with the Cholesky factor L replacing X^*. ▯

Triangular matrices are perfectly suited for preconditioning by the fact that applying the inverse is inexpensive depending, of course, on the sparsity. Suppose thus $l_j(\Lambda) = 0$ for A. Then how to inexpensively compute L approximately so as to reduce the conditioning of $M = Ap(A)$ is a natural problem not considered here.

A necessary condition for being polynomially unitary of degree j is hence the existence of a polynomial $q(\lambda) = \lambda p(\lambda)$ of degree j such that the (polynomial) lemniscate

$$\{\lambda \in \mathbb{C} : \ |q(\lambda)| = 1\} \tag{2.10}$$

contains the spectrum. These are intriguing sets. By the open mapping theorem, a lemniscate cannot possess interior points. In particular, a classical extremal problem is concerned with maximizing the arc length of a lemniscate over $\mathcal{P}_j(\infty)$, the set of monic polynomials of prescribed degree j; see [3]. See also [8] for other motivations. Of course, in connection with analyzing lemniscates, it makes no difference whether one considers monic polynomials or $\mathcal{P}_j(0)$, the set polynomials of degree j at most vanishing at the origin. After a simple affine change of variables, the problems are equivalent. However, a notable difference between these two normalizations is that with (2.10) the arc length is not bounded (as a function of the degree); see Proposition

3.1. This is important since in discretizing, e.g., PDE, the eigenvalues can be very widely spread.

In a certain sense, also monic polynomials can be used to represent the inverse.

EXAMPLE 2. Consider the identity (2.7). By replacing the unitary matrix with a translation of a unitary matrix allows using monic polynomials to factor the inverse. Namely, suppose $U = q(A)$ is unitary for a monic polynomial $q(\lambda) = \lambda^j + c_{j-1}\lambda^{j-1} + \cdots + c_0$. If $|c_0| \neq 1$, then

$$A^{-1} = p(A)M$$

with M given by the Neumann series expansion for the inverse of $U - c_0 I$.

Altogether, lemniscates in the complex plane immediately lead to the following matrix analytic extension. Here, analogously to the polar decomposition of a matrix, the role of the complex sign[5] is naturally taken over by unitary matrices.

DEFINITION 2.8. *Let p be a polynomial. Then*

$$p^{-1}(\mathcal{U}) = \{M \in \mathbb{C}^{n\times n} \;:\; p(M) \;\; \textit{is unitary}\}$$

is said to be a matrix lemniscate.

Of course, this equals $\{M \in \mathbb{C}^{n\times n} \;:\; \max_{\|x\|=1} |\|p(M)x\| - 1| = 0\}$.[6] In particular, the dimension $n = 1$ corresponds to the classical notion of lemniscate such that the problems in [9] can be accordingly posed for matrix lemniscates. (For a less algebraic, somewhat immediate dimensional extension, see [26].)

3. Lemniscate asymptotics and quasilemniscates. Since in practice the dimensions are large, it is not so easy to determine how $l_j(A)$ behaves while j grows. Because of this, and in analyzing iterative methods in general, it is customary to turn the attention to the spectrum to make judgments on how to proceed. Of course, because of Corollary 2.6, in the nearly normal case this is an entirely valid approach. Since the spectrum is rarely available, one typically takes a compact set Λ known to include the eigenvalues. In our case this leads us to study lemniscates and associated asymptotics with respect to Λ.

For a polynomial $q(\lambda) = \alpha_j \prod_{k=1}^{j}(\lambda - \lambda_j)$, its zeros are called the loci of the corresponding lemniscate

$$\{\lambda \in \mathbb{C} \;:\; |q(\lambda)| = 1\}. \tag{3.1}$$

(For basic facts about lemniscates, see [30, p.19].) In what follows, we are concerned with extending the notion of lemniscates having one locus at the origin, so that we have $q(\lambda) = \lambda p(\lambda)$. As opposed to monic polynomials, asymptotically to these correspond a natural limiting family of functions, i.e., functions f analytic in a neighborhood of 0 satisfying the normalization $f(0) = 0$.

EXAMPLE 3. For a famous family of lemniscates having a locus at the origin, set $q_0(\lambda) = \lambda$ and define $q_j(\lambda) = q_{j-1}(\lambda)^2 + \lambda$ for $j = 1, 2, \ldots$. Then the corresponding lemniscates (called the Mandelbrot curves) have the boundary of the Mandelbrot set as the limit set; see, e.g., [19, p.492].

This example illustrates well the flexibility of lemniscates. In particular, see [30, p.248] for Hilbert's theorem on approximating boundaries of compact sets having

[5] If $z \in \mathbb{C}$ is nonzero, its complex sign is defined as $\frac{z}{|z|}$.

[6] The case $p(\lambda) = \lambda$ corresponds to the unitary group and is hence the best understood matrix lemniscate. See [32, p.428] for a collection of basic facts concerning its structure.

connected complement with lemniscates. For the speed of convergence of such approximations, see [1].

As opposed to this, assume a compact $\Lambda \subset \mathbb{C}$ is given. Approximation theoretically now the task in (1.1) is to find a polynomial, with vanishing constant term, whose image of Λ is as close to a subset of the unit circle as possible. Since any positive power of the absolute value of an analytic function is subharmonic, we are concerned with approximating by subharmonic functions. (See [31] for approximating by subharmonic functions.) In particular, it is customary to connect polynomial approximations on the plane with potential theory. For us this means taking the logarithm yielding

$$\max_{\lambda\in\Lambda} |\log|p(\lambda)| + \log|\lambda|| \tag{3.2}$$

which should be minimized over polynomials of degree $j-1$ at most. Algebraically stated, the aim is to minimize

$$\max_{\lambda\in\Lambda} \min_{\theta\in[0,2\pi)} |q(\lambda) - e^{i\theta}| = \max_{\lambda\in\Lambda} |q(\lambda) - \frac{q(\lambda)}{|q(\lambda)|}| \tag{3.3}$$

over the subspace of polynomials q of degree j at most having a zero at the origin.

Consider (1.1). Clearly, if Λ and $\hat{\Lambda}$ are compact such that $\Lambda \subset \hat{\Lambda}$, then

$$l_j(\Lambda) \le l_j(\hat{\Lambda}) \tag{3.4}$$

holds.

Proposition 3.1. *Let $\Lambda \subset \mathbb{C}$ be compact. Then $l_j(z\Lambda) = l_j(\Lambda)$ for any nonzero $z \in \mathbb{C}$.*

Proof. Let $\lambda p(\lambda) = \sum_{k=1}^{j} a_k \lambda^k$ be the polynomial realizing (1.1). For $z\Lambda$ take $z\lambda\tilde{p}(z\lambda) = \sum_{k=1}^{j} \tilde{a}_k z^k \lambda^k$. Therefore choose the coefficients of $\tilde{p}$ according to $\tilde{a}_k z^k = a_k$, yielding the claim. □

For any solution, the following holds for the width of the origin centered annulus containing the image of Λ.

Proposition 3.2. *Let $\Lambda \subset \mathbb{C}$ be compact and assume $q(\lambda) = \lambda p(\lambda)$ solves* (1.1). *Then $\min_{\lambda\in\Lambda} |q(\lambda)| + \max_{\lambda\in\Lambda} |q(\lambda)| = 2$. Moreover, $l_j(\Lambda) = 1$ if and only if $0 \in \Lambda$.*

Proof. This is a matter of scaling, i.e., p must be chosen such that the distances $1-\min_{\lambda\in\Lambda} |q(\lambda)|$ and $\max_{\lambda\in\Lambda} |q(\lambda)|-1$ equal. (Otherwise the solution is not optimal.)

If $0 \in \Lambda$, then $l_j(\Lambda) = 1$. Suppose $0 \notin \Lambda$. Then take p which is nonzero on Λ. Scale with $r > 0$ such that $|\lambda r p(\lambda)| < 2$ for any $\lambda \in \Lambda$. This implies $l_j(\Lambda) < 1$. □

The width of the annulus is zero if Λ is a subset of a lemniscate having a locus at the origin. Otherwise, the limiting behavior of this width is of interest. First of all, if 0 is not in the polynomially convex hull of Λ, then $\lim_{j\to\infty} l_j(\Lambda) = 0$ by the fact that $\frac{1}{\lambda}$ can be approximated with polynomials uniformly on Λ by Mergelyan's theorem. This assumption guarantees the success of the GMRES method. (Hence, the polynomially convex hull of Λ can be regarded to yield the "spectrum" of the GMRES method.)

For the converse, having 0 in the polynomially convex hull of Λ is not an obstruction for us as such. For example, any lemniscate (3.1) having a locus at the origin satisfies this property automatically. However, now the situation is far from being so clear cut.

Theorem 3.3. *Let $U \subset \mathbb{C}$ be an open bounded simply connected set such that its Riemann map has a continuous extension to its boundary Λ_1. Let Λ_2 be compact such that $\Lambda_2 \cap \overline{U} = \emptyset$. If $0 \notin \Lambda = \Lambda_1 \cup \Lambda_2$, then $\lim_{j\to\infty} l_j(\Lambda) = 0$.*

Proof. Suppose first that $\Lambda_2 = \emptyset$. We only need to consider the case of U containing 0. In this case, let us choose the Riemann map f of U onto the open unit disk $\mathbb{D}$ satisfying $f(0) = 0$. Then by Mergelyan's theorem, f can be uniformly approximated by polynomials in the closure of U. These polynomials are of the form $\epsilon_j + \lambda p_{j-1}(\lambda)$ with $\epsilon_j \to 0$ as $j \to \infty$. The polynomials $\lambda p_{j-1}(\lambda)$ yield the claim.

For the general case, set $\frac{f(\lambda)}{\lambda}$ on $\overline{U}$ and $\frac{1}{\lambda}$ on Λ_2. This function satisfies the assumptions of the Mergelyan's theorem and can hence be uniformly approximated with polynomials p_{j-1} on $\Lambda_1 \cup \Lambda_2$. Consider now $\lambda p_{j-1}(\lambda)$ which yield the claim. □

Carathédeory's theorem gives sufficient conditions on the Riemann map to have a continuous extension to the boundary. For instance, if Λ_1 is a Jordan curve, then the Riemann map has a continuous extension to the boundary. Operators with the spectrum satisfying such an inclusion relationship are classical [7, Chapter 4].

Theorem 3.3 can to some extent be extended to allow non-intersecting Jordan arcs to stretch out from the boundary of Λ_1. On these arcs, extend f continuously to have (unit) constant value. The proof goes analogously.

DEFINITION 3.4. *A compact set $\Lambda \subset \mathbb{C}$ not containing 0 is said to be severely indefinite if $\lim_{j\to\infty} l_j(\Lambda) > 0$. The set of severe indefiniteness of Λ consists of those $\mu \in \mathbb{C}$ for which $\Lambda - \mu$ is severely indefinite.*

By Proposition 3.2, the set of severe indefiniteness of Λ includes Λ. We conjecture that a component of Λ^c either belongs to the set of severe indefiniteness of Λ or has an empty intersection with it.[7]

For its full applicability, the following (negative) result should be combined with (3.4) for compact sets $\hat{\Lambda}$ containing Λ as a subset.

THEOREM 3.5. *For a compact $\Lambda \subset \mathbb{C}$, suppose the component of Λ^c containing 0 is bounded and not simply connected. Then Λ is severely indefinite.*

Proof. Assume first $\Lambda = \mathbb{T} \cup \nu$ with $|\nu| < 1$, where $\mathbb{T}$ denotes the unit circle centered at the origin. For p_j solving (1.1), let $\max_{\lambda\in\mathbb{T}} |\lambda p_j(\lambda)| = 1 + \epsilon_j$ with $\epsilon_j \geq 0$. Then the Schwarz lemma yields

$$\frac{|\nu p_j(\nu)|}{1 + \epsilon_j} \leq |\nu|. \tag{3.5}$$

Therefore if ϵ_j converges to zero, $\lim_{j\to\infty} l_j(\Lambda) \geq 1 - |\nu|$.

To prove the general case, let U be the smallest open simply connected set containing the component of Λ^c containing 0. Suppose $\max_{\lambda\in\Gamma} |\lambda p_j(\lambda)| = 1 + \epsilon_j$, where Γ denotes the boundary of U. Let f be the biholomorphic map of from the open unit disk $\mathbb{D}$ onto U provided by the Riemann mapping theorem such that $f(0) = 0$. Consider $r\mathbb{D}$ for $0 < r < 1$. Assume $\nu \in U \cap \Lambda$ and denote $z = f^{-1}(\nu)$. The function

$$\lambda \mapsto \frac{f(r\lambda)p_j(f(r\lambda))}{1 + \epsilon_j}$$

satisfies the assumption of the Schwarz lemma. At $\lambda = z/r$ the Schwarz lemma yields $\frac{|\nu p_j(\nu)|}{1+\epsilon_j} \leq |z/r|$. Therefore if ϵ_j converges to zero, $\lim_{j\to\infty} l_j(\Lambda) \geq 1 - |z/r|$. Choose now $0 < r < 1$ such that $|z/r| < 1$ to have the claim. □

EXAMPLE 4. Let Λ be the boundary of the Mandelbrot set. By Example 3, we know it can be approximated with lemniscates having a locus at the origin. However,

[7] Respective claim for operators would yield a spectral inclusion set analogously, e.g., to the polynomial convex hull of the spectrum.

here Theorem 3.3 cannot be applied since we cannot establish that the Riemann map in this case has a continuous extension to the boundary. In fact, we do not know what $\lim_{j\to\infty} l_j(\Lambda)$ is. Since the Hausdorff dimension of Λ is two, it is not inconceivable either that the limit be strictly positive.

Theorem 3.5 certainly has implications for iterative methods, underscoring that it is not sufficient to concentrate on (2.5) because of the apparent rigidity of polynomial preconditioning for the normal equations. The problems (2.4) and (2.2) are, for moderate values of k, better suited for choosing a preconditioning strategy. Moreover, the GMRES method is a hopeless alternative for solving a linear system with the spectrum including (approximately) a set in Theorem 3.5; preconditioning is absolutely necessary then.

Example 5. Theorem 3.5 is also perturbation theoretically striking in terms of adjoining points. Namely, denote by Λ a lemniscate (3.1) with a locus at the origin. If its complement has a bounded component U containing 0, then take $\nu_1 \in U$. Then the union $\Lambda \cup \nu_1$ is far from being a lemniscate with a locus at the origin in the sense that we have $\lim_{j\to\infty} l_j(\Lambda \cup \nu_1) > 0$, i.e., a severely indefinite set.

In particular, respective bounds for the GMRES method are far more robust under perturbations of this type [25].

As opposed to Theorem 3.3, the following yields tools for more concrete estimates.

Theorem 3.6. *Assume Λ is a lemniscate with a locus at the origin and let ν_k, for $k = 1,\ldots,n$, be in the unbounded component of Λ^c. If $\hat\Lambda = \Lambda \cup \nu_1 \cup \cdots \cup \nu_n$, then $\lim_{j\to\infty} l_j(\hat\Lambda) = 0$.*

Proof. Consider first the case $\Lambda = \mathbb{T}$. Then we have $\hat\Lambda = \mathbb{T} \cup \nu_1 \cup \cdots \cup \nu_n$ with $|\nu_k| > 1$ for $k = 1,\ldots,n$. By interpolation, let $P(\lambda)$ be the polynomial of degree $n-1$ attaining values $\frac{1}{\nu_k}$ at the points ν_k. Let $Q(\lambda) = \prod_{k=1}^n (\lambda - \nu_k)$. Consider then finding a function R satisfying

$$|P(\lambda) + Q(\lambda)R(\lambda)| = 1 \tag{3.6}$$

for $\lambda \in \mathbb{T}$, so that $|\lambda(P(\lambda) + Q(\lambda)R(\lambda))| = 1$ for $\lambda \in \hat\Lambda$. Certainly, (3.6) holds if $R(\lambda) = \frac{1-P(\lambda)}{Q(\lambda)}$ for $\lambda \in \mathbb{T}$. Thereafter R should be approximated with polynomials. To this end, the poles of R are located outside the unit disk and are determined by Q. We have the Maclaurin series expansion

$$\frac{1}{Q(\lambda)} = \prod_{k=1}^{n} (\frac{-1}{\nu_k} \sum_{l=0}^{\infty} (\frac{\lambda}{\nu_k})^l) \tag{3.7}$$

converging uniformly in a neighborhood of the unit disk. Truncate $\frac{1}{Q}$ to have polynomials Q_{j-1}. Consequently, the polynomials $\lambda(P(\lambda) + Q(\lambda)(1 - P(\lambda))Q_{j-1}(\lambda))$ yield the required approximation.

To deal with the general case, assume Λ to be given by (3.1). Consider

$$\lambda \longmapsto \omega = \lambda p(\lambda). \tag{3.8}$$

Then the image of $\hat\Lambda$ under this map is the union of $\mathbb{T}$ and the points $\mu_k = \nu_k p(\nu_k)$ for $k = 1,\ldots,n$. Hence repeat the previous construction in the ω-plane to get polynomials proving the claim by the fact that (3.8) has a zero at 0. ☐

The following should be contrasted with Example 5.

EXAMPLE 6. Assume Λ is a lemniscate with a locus at the origin given by (3.1) and ν_1 is in the unbounded component of Λ^c. Let $\hat{\Lambda} = \Lambda \cup \nu_1$. Then

$$l_{(2+m+1)j-1}(\hat{\Lambda}) \leq \max_{\lambda \in \Lambda} |Q(\lambda p(\lambda))(1 - P(\lambda p(\lambda))| \frac{1}{|\nu_1 p(\nu_1)|^{m+1} |\nu_1 p(\nu_1) - 1|}.$$

Asymptotically we have $\lim_{m\to\infty} l_{(2+m+1)j-1}(\Lambda)^{1/((2+m+1)j-1)} \leq \frac{1}{|\nu_1 p(\nu_1)|^j}$.

This brings up a natural manner to measure how fast the width of the annulus decreases while the degree of the polynomials increases. We also want to have an opposite for being severely indefinite, i.e., we look for a notion of indefiniteness which is far from being severe. When the speed is measured geometrically, we arrive at

$$\eta(\Lambda) = \liminf_{j\to\infty} l_j(\Lambda)^{1/j}.$$

The following definition corresponds to the so-called superlinear convergence of the GMRES method [24, Chapter 5].

DEFINITION 3.7. *A compact set $\Lambda \subset \mathbb{C}$ is said to be a quasilemniscate with a locus at the origin if $\eta(\Lambda) = 0$.*[8]

A compact set Λ is a quasilemniscate with a locus at the origin if and only $z\Lambda$ is for any nonzero $z \in \mathbb{C}$. (See the proof of Proposition 3.1.)

Let us make some potential theoretic remarks. First of all, if the capacity of Λ is zero and $0 \notin \Lambda$, then it follows that Λ is a quasilemniscate. Moreover, assuming Λ does not contain zero, the fact that (2.6) bounds (2.5) gets now expressed as $\eta(\Lambda) \leq e^{-g_\Lambda(0)}$, where $g_\Lambda(z)$ denotes the Green function for Λ; see [24, Chapter 3] or [21].

THEOREM 3.8. *Let $\Lambda = [a, b] \subset \mathbb{R}$ with $0 \leq a < b < \infty$. Then*

$$\left(\frac{\sqrt{b} - \sqrt{a}}{\sqrt{b} + \sqrt{a}}\right)^2 \leq \eta(\Lambda) \leq \frac{\sqrt{b} - \sqrt{a}}{\sqrt{b} + \sqrt{a}}.$$

Proof. For any polynomial p of degree $j - 1$ we have

$$\left||\lambda p(\lambda)|^2 - 1\right| = ||\lambda p(\lambda)| - 1| \, ||\lambda p(\lambda)| + 1| \, .$$

Since $\lambda \in \mathbb{R}$, we have $|\lambda p(\lambda)|^2 = \lambda^2 r(\lambda)$ for a polynomial r of degree $2j - 2$ with real coefficients. Therefore, letting $\lambda r(\lambda)$ to be any complex polynomial of degree $2j - 1$ yields

$$\min_{p \in \mathcal{P}_{2j-1}} \max_{\lambda \in \Lambda} |\lambda p(\lambda) - 1| \leq 2 \min_{p \in \mathcal{P}_{j-1}} \max_{\lambda \in \Lambda} ||\lambda p(\lambda)| - 1| \, .$$

Taking the jth root and limits on both sides yields the claim. Then use the fact that now $g_\Lambda(0) = \log \frac{\sqrt{b}+\sqrt{a}}{\sqrt{b}-\sqrt{a}}$; see, e.g., [21, p.33].
□

COROLLARY 3.9. *If a compact set $\Lambda \subset \mathbb{C}$ contains an interior point, then $\eta(\Lambda) > 0$.*

Proof. Multiply Λ by $e^{i\theta}$ with $\theta \in [0, 2\pi)$ such that the interior of $e^{i\theta}\Lambda$ intersects the real axis. Then use the invariance of $\eta(\Lambda)$ under such rotations. □

[8] For the terminology, in operator theory the so-called quasialgebraic operators generalize algebraic operators [12]. Here, spectrally, this corresponds to an analogous generalization for (2.7).

Regarding the terminology, the preceding and the following properties appear natural for (not being) a quasilemniscate.

THEOREM 3.10. *A quasilemniscate* Λ *with a locus at the origin does not contain a continuum of a line.*

Proof. Suppose Λ contains a line segment with the end points λ_1 and λ_2. The segment cannot contain the origin since it would imply $\eta(\Lambda) = 1$. For any polynomial p we have $\big||\lambda p(\lambda)|^2 - 1\big| = ||\lambda p(\lambda)| - 1|\,||\lambda p(\lambda)| + 1|$ Therefore

$$\min_{\deg(p)\leq j}\max_{t\in[-1,1]}\left||\frac{1-t}{2}\lambda_1 + \frac{(1+t)}{2}\lambda_2|^2|p(\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2)|^2 - 1\right|$$

$$\leq M\min_{\deg(p)\leq j}\max_{t\in[-1,1]}\left||\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2||p(\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2)| - 1\right|$$

for some $0 < M < \infty$. Now $|p(\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2)|^2$ is a real polynomial in t of degree $2j$. Consequently,

$$m\min_{\deg(p)\leq 2j}\max_{t\in[-1,1]}\left|p(t) - \frac{1}{|\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2|^2}\right|$$

$$\leq M\min_{\deg(p)\leq j}\max_{t\in[-1,1]}\left||\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2||p(\frac{1-t}{2}\lambda_1 + \frac{1+t}{2}\lambda_2)| - 1\right|$$

for some $0 < m < \infty$. Now the denominator of $\frac{1}{|t\lambda_1+(1-t)\lambda_2|^2}$ is a two degree polynomial in t and can be analytically extended to the complement of the interval $[-1,1]$. The extended rational function is not entire. Therefore $\frac{1}{|t\lambda_1+(1-t)\lambda_2|^2}$ can be approximated with polynomials on $[-1,1]$ with a speed which is slower than required for a quasilemniscate, see [28, Theorem 3.10]. This yields a strictly positive lower bound. □

So far we have lacked any algebraic structure associated with these extended notions of lemniscates. The following allows us to look at a subalgebra of entire functions yielding quasilemniscates.

THEOREM 3.11. *Suppose* Λ *is compact and* $\Lambda \subset f^{-1}(\mathbb{T})$ *for an entire function* f *with* $f(0) = 0$*. Then* Λ *is a quasilemniscate with a locus at the origin.*

Proof. Take any function f analytic in a neighborhood of the origin such that $f(0) = 0$. Expand f into its Maclaurin series as $f(\lambda) = \sum_{k=1}^{\infty} d_k\lambda^k$ having R as a radius of convergence. Then for $\lambda \in \Lambda$ with $|\lambda| < R$ we have with $p(\lambda) = \sum_{k=1}^{j} d_k\lambda^k$

$$||p(\lambda)| - 1| = ||p(\lambda)| - |f(\lambda)|| \leq |p(\lambda) - f(\lambda)| = |\sum_{k=j+1}^{\infty} d_k\lambda^k| \leq \frac{(\frac{1}{R} + \epsilon_{j+1})^{j+1}|\lambda|^{j+1}}{1 - (\frac{1}{R} + \epsilon_{j+1})|\lambda|} \tag{3.9}$$

for $\epsilon_{j+1} \geq 0$ converging to zero as $j \to \infty$. Consequently, assuming f to be entire yields the claim by the fact that R can be arbitrarily large. □

See [22, p. 521] for a plot in the case of $f(z) = \sin z$. Of course, then the complement of $f^{-1}(\mathbb{T})$ has infinitely many components.

COROLLARY 3.12. *Suppose* f *is an entire function with* $f(0) = 0$*. Then* $f^{-1}(\mathbb{T})$ *does not contain a continuum of a line.*

Without the restriction $f(0) = 0$, the set $f^{-1}(\mathbb{T})$ can contain an entire line as the entire function $f(\lambda) = e^{\lambda}$ illustrates. Then the imaginary axis is mapped onto $\mathbb{T}$.

Occasionally $f^{-1}(\mathbb{T})$ for an analytic function f is called a generalized lemniscate; see [22]. If f is analytic but not necessarily entire, then (3.9) yields a bound on the speed of convergence in case $\Lambda \subset f^{-1}(\mathbb{T})$ such that Λ is located in a disk where the Maclaurin series of f converges.

Example 7. If f is not entire, then $f^{-1}(\mathbb{T})$ can contain a continuum of line. To see this, consider $f(\lambda) = \frac{\lambda}{2-\lambda}$. Then $f^{-1}(\mathbb{T})$ contains the vertical line intersecting the real axis at 1. Since f has a pole at 2, the speed of convergence can be bounded by using (3.9)

The behavior of $\eta(\Lambda)$ is not continuous.

Proposition 3.13. *Assume* $\Lambda = \mathbb{T} \cup t$ *with* $t > 1$. *Then* $\frac{1}{t^2} \leq \eta(\Lambda) \leq \frac{1}{t}$.

Proof. We only need to show that the lower bound holds since the upper bound follows from Example 6. Suppose $\lambda p(\lambda)$ realizes (1.1). Denote by λ_k the zeros of p with $|\lambda_k| < 1$. Then $p(\lambda) = r(\lambda)\lambda^l \prod \frac{\lambda-\lambda_k}{1-\overline{\lambda}_k\lambda}$ with a polynomial r without zeros inside the unit disk. We have $|r(e^{i\theta})| = |p(e^{i\theta})|$ for any $\theta \in \mathbb{R}$ and $|r(\lambda)| \leq |p(\lambda)|$ for $|\lambda| > 1$. Consequently,

$$|r(t)| \leq \frac{1}{t}(1 + l_j(\Lambda)). \tag{3.10}$$

By the maximum modulus principle, $1 - l_j(\Lambda) < |r(\lambda)| < 1 + l_j(\Lambda)$ for $|\lambda| \leq 1$. Let us decompose $r(\lambda) = c + \hat{r}(\lambda)$ with with a constant $|c| = 1$. Using these inequalities, we have $\max_{|\lambda|\leq 1} |\hat{r}(\lambda)| \leq l_j(\Lambda) + \frac{2}{t}\sqrt{j-1}\, l_j(\Lambda)^{1/2}$ obtained by integrating $|r(\lambda)|^2$ over $\mathbb{T}$. Therefore

$$|\hat{r}(t)| \leq \max_{|\lambda|\leq 1} |\hat{r}(\lambda)| t^{j-1} \leq (l_j(\Lambda) + 2\sqrt{j-1}\, l_j(\Lambda)^{1/2}) t^{j-1}.$$

Because of (3.10), we have $|\hat{r}(t)| \geq 1 - \frac{1}{t}(1 + l_j(\Lambda))$. Thus

$$1 - \frac{1}{t}(1 + l_j(\Lambda)) \leq (l_j(\Lambda) + 2\sqrt{j-1}\, l_j(\Lambda)^{1/2}) t^{j-1},$$

so that $\frac{t-1}{t^j} \leq 4\sqrt{j-1}\, l_j(\Lambda)^{1/2}$ which yields the claim. □

To end this section, if Λ is a rectifiable curve, then we have

$$\min_{p \in \mathcal{P}_{j-1}} \left(\int ||\lambda p(\lambda)| - 1|^2 d\lambda\right)^{1/2} \leq L(\Gamma)^{1/2} l_j(\Lambda) \tag{3.11}$$

which thus can be used to bound $\eta(\Lambda)$ from below. How to numerically solve the respective Hilbert space minimization problem is studied in the section that follows.

4. Computing polynomials for mapping sets into an origin centered annulus. This section is concerned with ways to compute polynomials for mapping a given compact set Λ to be near the unit circle, i.e., into an origin centered annulus possessing a small width. In doing so, Theorem 3.5 cannot be ignored in the sense that if Λ surrounds 0, then the corresponding component must in any event be simply connected. Bearing in mind our original problem (2.2), in this connection there are good reasons to allow appropriate relaxations such as the following.

Definition 4.1. *A compact set* $\Lambda \subset \mathbb{C}$ *is said to be an essential lemniscate if it is a union of a lemniscate and a finite set.*

Let us next derive orthogonality conditions for approximating $l_j(\Lambda)$ in the L^2-norm. Assume the set Λ is finite consisting of points λ_l for $l = 1, \ldots, n$. Hence, if Λ

originally contains, e.g., a continuum, it must be appropriately discretized first, possibly taking into account Definition 4.1 through ignoring some points of Λ. Consider then the Hilbert space minimization problem

$$\min_{q\in\mathcal{P}_j(0)} |||q|-1||_2 \tag{4.1}$$

with respect to the discrete measure $\sum_{j=1}^n \delta_{\lambda_j}$ on Λ. Like in the max-norm case, the set of solutions is invariant under multiplications by a constant $e^{i\theta}$ with $\theta \in [0, 2\pi)$.

Instead of applying the gradient method, the following necessary geometric condition can be used in devising a descent method. We find it attractive because it links our problem with orthogonal polynomials. As is well-known, orthogonal polynomials play a central role in iterative methods.

THEOREM 4.2. *Suppose $q(\lambda) = \lambda p(\lambda)$ solves (4.1). Then $p(\lambda)(1 - \frac{1}{|\lambda p(\lambda)|})$ is orthogonal against $\mathcal{P}_{j-1}$ with respect to the discrete measure $\sum_{j=1}^n |\lambda_j|^2\delta_{\lambda_j}$ on Λ.*

Proof. It is immediate that (4.1) is equivalent to solving

$$\min_{q\in\mathcal{P}_j(0),\, g\in\mathcal{G}} ||q-g||_2, \tag{4.2}$$

where $\mathcal{G}$ denotes the group (under multiplication) of continuous functions on Λ having values in the unit circle. Hence we are looking at the distance between a j dimensional polynomial subspace and an infinite group. Moreover, for any fixed polynomial $q(\lambda) = \lambda p(\lambda)$ the nearest $g \in \mathcal{G}$ is given by $\frac{q(\lambda)}{|q(\lambda)|}$. For $q(\lambda)$ to be optimal, i.e., to solve (4.1), the difference $\frac{\lambda p(\lambda)}{|\lambda p(\lambda)|} - \lambda p(\lambda)$ must be orthogonal against $\mathcal{P}_j(0)$. Consequently, for any polynomial $r \in \mathcal{P}_{j-1}$ necessarily

$$(\lambda r(\lambda), \frac{\lambda p(\lambda)}{|\lambda p(\lambda)|} - \lambda p(\lambda))_2 = 0$$

holds, yielding the claim by the fact that $(\lambda r(\lambda), \frac{\lambda p(\lambda)}{|\lambda p(\lambda)|} - \lambda p(\lambda))_2 = (r(\lambda), |\lambda|^2(\frac{p(\lambda)}{|\lambda p(\lambda)|} - p(\lambda)))_2$. □

It is instructive to bear in mind that standard orthogonal polynomials result from orthogonalizing a polynomial p of degree j against $\mathcal{P}_{j-1}$ for $j = 1, 2, \ldots$. (See, e.g., [29, p. 400].) Theorem 4.2 yields a geometric analogy for this. That is, a polynomial $q(\lambda) = \lambda p(\lambda)$ solving (4.1) is such that in the orthogonality condition one missing degree in p is replaced with the multiplying rational function $1 - \frac{1}{|\lambda p(\lambda)|}$. Algorithmically finding this is achieved through converting an infinite dimensional optimization problem (4.2) into a finite dimensional one for which we can compute critical points.

For a descent method, there is a way to find polynomials of Theorem 4.2 by solving differential equations. To this end, equip $\mathcal{P}_{j-1}$ with the inner product with respect to the measure

$$\sum_{j=1}^{n} |\lambda_j|^2 \delta_{\lambda_j} \tag{4.3}$$

on Λ. Then consider the vector field on $\mathcal{P}_{j-1}$ by defining $V_{||}(p)$ at p to be the component of

$$p(\lambda)(\frac{1}{|\lambda p(\lambda)|} - 1) = V_{||}(p) + V_{\perp}(p) \tag{4.4}$$

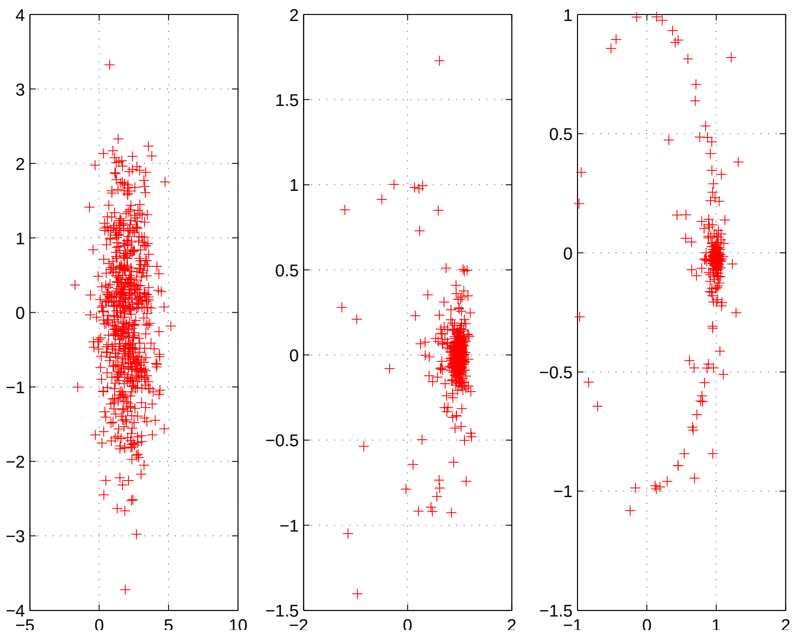

FIG. 4.1. *On the left* Λ. *In the middle, the image of* Λ *with* q *of degree* 30. *On the right, the image of* Λ *with* q *of degree* 60.

in $\mathcal{P}_{j-1}$ while $V_{\perp}(p)$ is the orthogonal component with respect to the inner product used. The respective flow moves points towards polynomials satisfying the orthogonality condition of Theorem 4.2. Once reached, such a point yields a local minimum for (4.1), by construction. (Clearly, if the points belong to the lemniscate determined by $\lambda p(\lambda)$, we are at a critical point.) In particular, if the step length is maximally determined by the condition (4.4), we obtain the following iteration with an initial guess $p_0 \in \mathcal{P}_{j-1}$.

$$\begin{array}{ll} \textbf{for} & k = 1, 2, \ldots \\ & \textbf{compute}\, V_{||}(p_{k-1}) \\ & \textbf{set}\, p_k = p_{k-1} + V_{||}(p_{k-1}) \\ \textbf{end} & \end{array} \tag{4.5}$$

Clearly, this is straightforward to implement such that each step is very inexpensive. (With obvious changes, everything can be repeated to solve $\min_{\deg(q)\leq j} |||q| - 1||_2$.)

EXAMPLE 8. In this numerical example, solved by using Matlab, Λ consisted of 600 randomly generated points produced by setting $\tilde{\Lambda} = \text{randn}(600, 1) + i * \text{randn}(600, 1)$ and then $\Lambda = \tilde{\Lambda} + 2 * \text{ones}(200, 1)$. As can be seen from the left panel in Figure 4.1, it appears challenging to map such a clustered set to be near the unit circle. We executed the iteration (4.5) to have q of various degrees by taking $k = 60$ iterates. The constant polynomial $p_0(\lambda) = 1$ was used as an initial guess. See Figure 4.1 for Λ and its image under q with $\deg(q) = 30$ and $\deg(q) = 60$.

There are other computational approaches for constructing polynomials to map compact sets into an annulus. Let us outline some ideas.

EXAMPLE 9. If Λ is a Jordan curve surrounding the origin, an obvious option is based on approximating the Riemann map. By invoking the proof of Theorem 3.3 with $\Lambda_2 = \emptyset$, approximating the Riemann map f with polynomials leads to a method for computing polynomials of interest. However, f is typically not available. There exists the so-called interpolating polynomial method [27] which can then be applied. It consists of interpolating points of Λ to be on the unit circle. Of course, the critical ingredient of this approach is to cleverly choose the interpolating points. This may not be straightforward.

With the help of the formulation (4.2), the problem can also be converted into first finding g such that thereafter p is determined. To this end, assume the set Λ is finite consisting of points λ_l for $l = 1, \ldots, n$. Set $q_k(\lambda) = \lambda^k$ for $k = 1, \ldots, j$. In terms of the respective discrete formulation of (4.2), consider solving for $(\alpha_1, \ldots, \alpha_k)$ the system

$$\sum_{k=1}^{j} \alpha_k g_k(\lambda_l) = e^{i\theta_l}. \tag{4.6}$$

Assuming $j \leq n$, the task is to choose the angles $\theta_l \in [0, 2\pi)$, for $l = 1, \ldots, n$, so as to attain as small error as possible.

The problem of optimally choosing the angles is nonlinear. To solve it in the least squares sense, form the n-by-j Vandermonde matrix $V = V(\lambda_1, \ldots, \lambda_n) = \{v_{lk}\} = \{g_k(\lambda_l)\}$ and compute its QR factorization $V = QR$. Now the columns of Q yield an orthonormal basis of the columns space of V. Therefore the optimal solution corresponds to maximizing the Fourier coefficients as

$$\max_{\theta_l \in [0,2\pi)\; l=1,\ldots,n} ||Q^* \begin{bmatrix} e^{i\theta_1} \\ \vdots \\ e^{i\theta_n} \end{bmatrix} ||^2 \tag{4.7}$$

which is bounded from above by n. Equality holds in (4.6) if and only if n is attained. Now (4.7) is an unconstrained optimization problem solvable, e.g, by Newton's method. Of course, it suffices to restrict to $[0, 2\pi)^n$.

EXAMPLE 10. The same optimization problem (4.7) can be used to check whether given n points λ_l belong to a lemniscate $\{\lambda \in \mathbb{C} : |p(\lambda)| = 1\}$ for some monic polynomial $p(\lambda) = \lambda^{j-1} + \alpha_{j-2}\lambda^{j-2} + \cdots + \alpha_0$. One simply uses $g_k(\lambda) = \lambda^{k-1}$, for $k = 1, \ldots, j$, in computing $V = QR$. Then replace $\begin{bmatrix} e^{i\theta_1} \\ \vdots \\ e^{i\theta_n} \end{bmatrix}$ with $\begin{bmatrix} e^{i\theta_1} - \lambda_1^{j-1} \\ \vdots \\ e^{i\theta_n} - \lambda_n^{j-1} \end{bmatrix}$.

Acknowledgments. We are very grateful to the referees for the pertinent remarks.

REFERENCES

[1] V. ANDRIEVSKII, *On the approximation of a continuum by lemniscates,* J. Approx. Theory, 105 (2000), pp. 292–304.

[2] G.H. GOLUB M. BENZI AND J. LIESEN, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137.

[3] P. Borwein, *The arc length of the lemniscate* $\{|p(z)| = 1\}$, Proc. Amer. Math. Soc., 123 (1995), pp. 797–799.
[4] V.I. Danchenko, *Lengths of lemniscates. Variations of rational functions,* Sb. Math., 198 (2007), pp. 1111–1117.
[5] K. Davidson, *Essentially normal operators*, Oper. Theory: Adv. Appl., 207 (2010), pp. 209–222.
[6] T. Driscoll, K-C. Toh and L.N. Trefethen, *From potential theory to matrix iterations in six steps,* SIAM Rev., 40 (1998), pp. 547–578.
[7] N. Dunford, *A survey of the theory of spectral operators*, Bull. Amer. Math. Soc., 64 (1958), pp. 217–274.
[8] P. Ebenfelt, D. Khavinson and H.S. Shapiro, Two-dimensional shapes and lemniscates, Complex analysis and dynamical systems IV. Part 1, Contemp. Math., 553, Amer. Math. Soc., Providence, RI, (2011) pp. 45–59.
[9] P. Erdös, F. Herzog and G. Piranian, *Metric properties of polynomials*, J. Analyse Math., 6 (1958), pp. 125–148.
[10] R. Freund, *On polynomial preconditioning and asymptotic convergence factors for indefinite Hermitian matrices*, Linear Algebra Appl., 154–156 (1991), pp. 259–288.
[11] A. Greenbaum and L.N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems,* SIAM J. Sci. Comput., 15(2) (1994), 359–368.
[12] P. Halmos, *Capacity in Banach algebras,* Indiana U. Math. J., 20 (1971), pp. 855–863.
[13] P. Halmos, A Hilbert Space Problem Book, 2nd Edition (Graduate Texts in Mathematics, Volume 19), *Springer-Verlag, New York*, 1982.
[14] R.A. Horn and C.R. Johnson, Matrix Analysis, *Cambridge Univ. Press, Cambridge,* 1987.
[15] M. Huhtanen, *Aspects of nonnormality for iterative methods*, Linear Algebra Appl., 394 (2005), pp. 119–144.
[16] M. Huhtanen, *Energy conservation and unitary approximation numbers,* J. Math. Phys., 48 (2007), 073512.
[17] M. Huhtanen and A. Perämäki, *Orthogonal polynomials of the* $\mathbb{R}$*-linear generalized minimal residual method,* J. Approx. Theory, 167 (2013), pp. 220–239.
[18] M. Huhtanen and O. Seiskari, *Matrix intersection problems for conditioning,* to appear in Banach Center Publ.
[19] V.G. Ivancevic and T.T. Ivancevic, High-Dimensional Chaotic and Attractor Systems: A Comprehensive Introduction, *Springer, Dordrecht,* 2007.
[20] F. Kittaneh, *On the structure of polynomially normal operators,* Bull. Austral. Math. Soc., 30 (1984), pp. 11–18.
[21] A. Kuijlaars, *Convergence analysis of Krylov subspace iterations with methods from potential theory,* SIAM Rev., 48 (2006), pp. 3–40.
[22] O.S. Kuznetsova and V.G. Tkachev, *Length functions on lemniscates*, Manuscripta Math., 112 (2003), pp. 519–538.
[23] R. Lange, *Sets of essentially unitary operators*, Trans. Amer. Math. Soc., 281 (1984), pp. 65–75.
[24] O. Nevanlinna Convergence of iterations for linear equations, Lectures in Mathematics ETH Zürich, *Birkhäuser Verlag, Basel,* 1993.
[25] O. Nevanlinna, Meromorphic functions and linear algebra, Fields Institute Monographs, 18. *American Mathematical Society, Providence, RI,* 2003.
[26] M. Paluszny, G. Montilla and J.R. Ortega, *Lemniscates 3D: a CAGD primitive?*, Numer. Algorithms, 39 (2005), pp. 317–327.
[27] r. M. Porter, *An interpolating polynomial method for numerical conformal mapping,* SIAM J. Sci. Comput., 23 (2001), pp. 1027–1041.
[28] T. Rivlin, The Chebyshev Polynomials, *Wiley-Interscience, New York* 1974 .
[29] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd ed., *SIAM, Philadelphia* 2003.
[30] J.L. Walsh, The Location of Critical Points of Analytic and Harmonic Functions, American Mathematical Society Colloquium Publications, Vol. 34. *Amer. Math. Soc., New York, N. Y.*, 1950.
[31] J.M. Wilson and D. Zwick, *Best approximation by subharmonic functions,* Proc. Amer. Math. Soc., 114 (1992), pp. 897–903.
[32] A. Zalewska-Mitura and J. Zemánek, *The Gerschgorin discs under unitary similarity,* Banach Center Publ., 38, Polish Acad. Sci., Warsaw, 1997, pp. 427–441.