DOUBLE-LAYER POTENTIALS, CONFIGURATION CONSTANTS AND APPLICATIONS TO NUMERICAL RANGES

B. MALMAN, J. MASHREGHI, R. O'LOUGHLIN, AND T. RANSFORD

ABSTRACT. Given a compact convex planar domain Ω with non-empty interior, the classical Neumann's configuration constant $c_{\mathbb{R}}(\Omega)$ is the norm of the Neumann-Poincaré operator K_Ω acting on the space of continuous real-valued functions on the boundary $\partial\Omega$, modulo constants. We investigate the related operator norm $c_{\mathbb{C}}(\Omega)$ of K_Ω on the corresponding space of complex-valued functions, and the norm $a(\Omega)$ on the subspace of analytic functions. This change requires introduction of techniques much different from the ones used in the classical setting. We prove the equality $c_{\mathbb{R}}(\Omega) = c_{\mathbb{C}}(\Omega)$, the analytic Neumann-type inequality $a(\Omega) < 1$, and provide various estimates for these quantities expressed in terms of the geometry of Ω. We apply our results to estimates for the holomorphic functional calculus of operators on Hilbert space of the type $\|p(T)\| \leq K \sup_{z \in \Omega} |p(z)|$, where p is a polynomial and Ω is a domain containing the numerical range of the operator T. Among other results, we show that the well-known Crouzeix-Palencia bound $K \leq 1 + \sqrt{2}$ can be improved to $K \leq 1 + \sqrt{1 + a(\Omega)}$. In the case that Ω is an ellipse, this leads to an estimate of K in terms of the eccentricity of the ellipse.

1. INTRODUCTION

1.1. Double-layer potential.

Throughout this article, Ω will denote a compact convex planar domain with non-empty interior. If $C(\partial\Omega)$ is the space of continuous functions on the boundary $\partial\Omega$ and $f \in C(\partial\Omega)$, then its *double-layer potential* u is the harmonic function

$$(1) \qquad u(z) = \frac{1}{\pi} \int_{\partial\Omega} f(\sigma) \, d\arg(\sigma - z) = \frac{1}{\pi} \int_{\partial\Omega} f(\sigma) \operatorname{Re}\left(\frac{N(\sigma)}{\sigma - z}\right) ds, \quad z \in \Omega^o.$$

Here $ds = |d\sigma|$ is the arclength measure on the rectifiable curve $\partial\Omega$, Ω^o is the interior of Ω, and $N(\sigma)$ is the outer-pointing normal at the boundary point σ. The equality between the two expressions for $u(z)$ above follows from an elementary computation in the case that $\partial\Omega$ is sufficiently smooth. In the general case, we interpret $N(\sigma)(\sigma - z)^{-1}$ as a Borel measurable function on $\partial\Omega$. By convexity of the domain, both the tangent $T(\sigma)$ and the normal $N(\sigma)$ exist and are continuous at all but a countable number of points σ which we will call *corners*, at which the discontinuity of T and N amounts to a jump in the argument. In Appendix A we include more details regarding boundaries of planar convex domains, and other facts mentioned below.

The Neumann–Poincaré operator appears in connection with the study of boundary behaviour of the double-layer potential. It is known that u given by (1) has a continuous extension to $\partial\Omega$, and we have the representation

$$(2) \qquad u(\zeta) = f(\zeta) + K_\Omega f(\zeta), \quad \zeta \in \partial\Omega$$
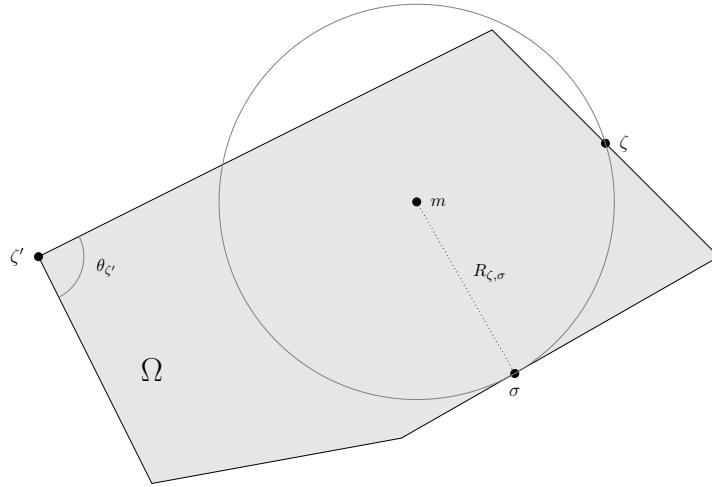
FIGURE 1. Example domain Ω with corner of angle $\theta_{\zeta'}$ at ζ', and a circle of radius $R_{\zeta,\sigma}$ with center m, tangent to $\partial\Omega$ at σ and passing through ζ.

where K_Ω denotes the *Neumann-Poincaré integral operator*

$$K_\Omega f(\zeta) := \frac{1}{\pi} \int_{\partial\Omega} f(\sigma) \, d\mu_\zeta(\sigma), \quad \zeta \in \partial\Omega.$$

Here μ_ζ is the probability measure

$$(3) \qquad d\mu_\zeta = (1 - \theta_\zeta/\pi)d\delta_\zeta + \rho_\zeta ds$$

where θ_ζ can be interpreted as the angle of the aperture at the possible corner at ζ of $\partial\Omega$, δ_ζ is a unit mass at the point ζ, and ρ_ζ is the Radon-Nikodym derivative

$$(4) \qquad \rho_\zeta(\sigma) := \frac{d\mu_\zeta}{ds}(\sigma) := \frac{1}{\pi} \operatorname{Re}\left(\frac{N(\sigma)}{\sigma - \zeta}\right) = \frac{1}{\pi} \operatorname{Im}\left(\frac{T(\sigma)}{\sigma - \zeta}\right).$$

It is natural to use the convention that $\theta_\zeta = \pi$ if ζ is not a corner. This occurs precisely when μ_ζ assigns no mass to the singleton $\{\zeta\}$. We will say that the collection of measures $\{\mu_\zeta\}_{\zeta \in \partial\Omega}$ is the *Neumann-Poincaré kernel* of Ω.

The density ρ_ζ has the following useful geometric interpretation. If $\sigma \in \partial\Omega \setminus \{\zeta\}$ is not a corner, and $R_{\zeta,\sigma}$ is the radius of the unique circle passing through ζ which is tangent to $\partial\Omega$ at σ, then the equality

$$(5) \qquad \rho_\zeta(\sigma) = \frac{1}{2\pi R_{\zeta,\sigma}}$$

holds. The radius $R_{\zeta,\sigma}$ may degenerate to ∞ if ζ is contained in the tangent line to $\partial\Omega$ passing through σ. In that case we see easily that $\rho_\zeta(\sigma) = 0$, so (5) still holds. To establish the formula, note that the center m of the circle in question is of the form $m = \sigma - RN(\sigma)$, where the radius $R = R_{\zeta,\sigma} > 0$ of the circle satisfies $|m - \zeta|^2 = |(\sigma - \zeta) - RN(\sigma)|^2 = R^2$. Expanding the squares and solving for R leads to (5).

1.2. **Neumann's configuration constant.**

1.2.1. *Real configuration constant.* Historically, the Neumann–Poincaré operator has been used to solve the *Dirichlet problem* of finding a harmonic extension to Ω^o of a given continuous function u on $\partial\Omega$. The extension can be obtained by finding $f \in C(\partial\Omega)$ which solves (2). Indeed, if such an f is found, then the extension of u to Ω^o is given by the double-layer potential in (1). This naturally leads to questions of invertibility of the operator $I + K_\Omega$ appearing on the right-hand side of (2), and consequently to the introduction of the *Neumann's configuration constant*, which we shall soon define as the operator norm of K_Ω acting on an appropriate space. Note that if $\mathbf{1}$ is the constant function, then we have that $K_\Omega \mathbf{1} = \mathbf{1}$, since each μ_ζ is a probability measure. Thus K_Ω can be naturally defined as a linear mapping on the quotient space $C(\partial\Omega)/\mathbb{C}\mathbf{1}$. The classical approach is to instead consider K_Ω as acting on the space of *real-valued* continuous functions $C_\mathbb{R}(\partial\Omega)$, in which case the corresponding quotient space $C_\mathbb{R}(\partial\Omega)/\mathbb{R}\mathbf{1}$ is endowed with the norm

$$(6) \qquad \|g + \mathbb{R}\mathbf{1}\|_{\partial\Omega} := \max_{\zeta,\zeta' \in \partial\Omega} \frac{|g(\zeta) - g(\zeta')|}{2} = \min_{r \in \mathbb{R}} \max_{\zeta \in \partial\Omega} |g(\zeta) - r|.$$

It is not hard to see that the two above expressions for the norm of the coset $g + \mathbb{R}\mathbf{1}$ are equivalent: they are both equal to half of the length of the interval $g(\partial\Omega) := \{g(\partial\Omega) : \zeta \in \partial\Omega\}$, the image of g. The right-most expression is minimized by choosing r to be the mid-point of the image interval. Neumann's (real) configuration constant $c_\mathbb{R}(\Omega)$ is defined as the operator norm of K_Ω acting on the quotient space $C_\mathbb{R}(\partial\Omega)/\mathbb{R}\mathbf{1}$:

$$(7) \qquad c_\mathbb{R}(\Omega) := \|K_\Omega : C_\mathbb{R}(\partial\Omega)/\mathbb{R}\mathbf{1} \to C_\mathbb{R}(\partial\Omega)/\mathbb{R}\mathbf{1}\|.$$

It is not hard to see that we may let K_Ω instead act from $C_\mathbb{R}(\partial\Omega)$ into the quotient $C_\mathbb{R}(\partial\Omega)/\mathbb{R}\mathbf{1}$ without affecting the operator norm. Since each measure μ_ζ is of unit mass, we have $0 \leq c_\mathbb{R}(\Omega) \leq 1$. If

$$\|f\|_{\partial\Omega} := \sup_{\zeta \in \partial\Omega} |f(\zeta)| \leq 1,$$

then

$$|K_\Omega f(\zeta) - K_\Omega f(\zeta')| \leq \|\mu_\zeta - \mu_{\zeta'}\|,$$

where we use the total variation norm (functional norm) on the right-hand side. By varying f over the unit ball of $C_\mathbb{R}(\partial\Omega)$ and ζ, ζ' over $\partial\Omega$, we obtain the important relation

$$(8) \qquad c_\mathbb{R}(\Omega) := \sup_{\zeta,\zeta' \in \partial\Omega} \frac{\|\mu_\zeta - \mu_{\zeta'}\|}{2}.$$

This expression for $c_\mathbb{R}(\Omega)$ will play a fundamental role in our study.

1.2.2. *Neumann's lemma.* From (8) we can immediately deduce that $c_\mathbb{R}(\Omega) = 1$ in the case that Ω is a triangle or a convex quadrilateral. Indeed, in those cases one sees from (3) and (4) that if ζ_1 and ζ_2 are corners of Ω (opposing, in the case of the quadrilateral) then μ_{ζ_1} and μ_{ζ_2} are mutually singular, and so $\|\mu_{\zeta_1} - \mu_{\zeta_2}\| = 2$, implying $c_\mathbb{R}(\Omega) = 1$. *Neumann's lemma*, which appears initially in Neumann's book [13], states that the cases of the triangle and quadrilateral are exceptional. For any other type of domain we have the strict inequality $c_\mathbb{R}(\Omega) < 1$. See [16] for a proof of this claim by Schober, and the curious history of incomplete attempts at a valid proof in full generality. Neumann's lemma implies the invertibility of $I + K_\Omega$ on $C_\mathbb{R}(\partial\Omega)/\mathbb{R}\mathbf{1}$, and thus the solvability of the Dirichlet problem on a convex domain Ω which is not one of the two exceptional cases. The remaining cases can be handled by considering instead powers of K_Ω. See, for instance, [12, Theorem 3.8], [5, Proposition 7], or the article [15], which contains also an exposition of the double-layer potential and Neumann's lemma.

At the other extreme, we have $c_\mathbb{R}(\Omega) = 0$ if and only if Ω is a disk. This result will be proved in Section 5.

1.3. Complex and analytic configuration constants.

1.3.1. *Two new configuration constants.* In the present article we will discuss certain applications of the double-layer potential to operator theory which motivate the definition of the *complex configuration constant*

(9) $$c_{\mathbb{C}}(\Omega) := \|K_\Omega : C(\partial\Omega)/\mathbb{C}\mathbf{1} \to C(\partial\Omega)/\mathbb{C}\mathbf{1}\|.$$

The difference between (7) and (9) is that the latter is the norm of K_Ω on the larger space of complex-valued functions. As a consequence, we have $c_{\mathbb{R}}(\Omega) \leq c_{\mathbb{C}}(\Omega) \leq 1$. There is a principal difference between the geometric interpretations of the norms in the quotient spaces $C_{\mathbb{R}}(\partial\Omega)/\mathbb{R}\mathbf{1}$ and $C(\partial\Omega)/\mathbb{C}\mathbf{1}$. In the former case, as we have already noted, the norm (6) of the coset represented by the real-valued function g is equal to half of the length of the image of g, this image being an interval on the real line $\mathbb{R}$. In the case of complex-valued g, the quotient norm

(10) $$\|g + \mathbb{C}\mathbf{1}\|_{\partial\Omega} := \min_{\lambda\in\mathbb{C}} \max_{\zeta\in\partial\Omega} |g(\zeta) - \lambda|$$

can instead be interpreted as the *radius of the smallest disk containing the image of g*. A crucial difference is that we lose the ability to estimate the norm of the coset $g + \mathbb{C}\mathbf{1}$ by considering the quantities $|g(\zeta) - g(\zeta')|$ only. This is the essence of why new tools are required to treat this case.

We will also study an analogous analytic constant, which is the norm of the operator K_Ω restricted to the subspace of analytic functions in $C(\partial\Omega)$. More precisely, we let $\mathcal{A}(\Omega)$ be the space of functions which are continuous in Ω and analytic in Ω^o. Each function in $\mathcal{A}(\Omega)$ has a unique restriction to $\partial\Omega$, and thus $\mathcal{A}(\Omega)$ can be naturally identified with a subspace of $C(\partial\Omega)$. We define the *analytic configuration constant* as

(11) $$a(\Omega) := \|K_\Omega : \mathcal{A}(\Omega)/\mathbb{C}\mathbf{1} \to C(\partial\Omega)/\mathbb{C}\mathbf{1}\|.$$

The space $\mathcal{A}(\Omega)$ is not invariant under K_Ω, but we do have that $K_\Omega f$ is the complex conjugate of a function in $\mathcal{A}(\Omega)$ (in [4, proof of Lemma 2.1] this claim is established for Ω with smooth boundary, but the same argument works in general). Clearly, we have the inequality $a(\Omega) \leq c_{\mathbb{C}}(\Omega)$. We note also that if $\widetilde{\Omega}$ is the image of Ω under an affine transformation of the plane, then the configuration constants of the two domains are equal. We shall verify this claim in Section 6.

1.3.2. *An application to functional calculi.* Given an operator T on a Hilbert space $\mathcal{H}$ with *numerical range*

$$W(T) := \{\langle Tx, x\rangle_{\mathcal{H}} : x \in \mathcal{H}, \|x\|_{\mathcal{H}} = 1\},$$

we are interested in the optimal constant $K > 0$ in the inequality

(12) $$\|p(T)\| \leq K \cdot \sup_{z\in W(T)} |p(z)| = K\|p\|_{W(T)},$$

where p is an analytic polynomial, and the left-hand side is the operator norm of $p(T)$ acting on $\mathcal{H}$. More generally, if $W(T)$ in (12) is replaced by an arbitrary domain Ω, and if the corresponding inequality holds for some K, then we say that Ω is a K-*spectral set* for T. Von Neumann's inequality says that the unit disk is a 1-spectral set for any contraction T, and a result of Okubo-Ando from [14] says that any disk containing $W(T)$ is a 2-spectral set for T.

The numerical range $W(T)$ is a bounded convex subset of the plane, its closure $\overline{W(T)}$ contains the spectrum $\sigma(T)$ of T, and it has non-empty interior in the case that T is not a normal operator (see, for instance, [9, Chapter 1]). For normal operators, the bound (12) with constant $K = 1$ is a consequence of the spectral theorem, and it suffices to take the supremum on the right-hand side over the smaller set $\sigma(T)$. For general T, even establishing the existence of a bound as in (12) is a non-trivial task. A result of Deylon-Deylon from [5, Theorem 3] establishes the existence of the bound, and shows that K can be chosen depending only on the area and the diameter of $W(T)$.
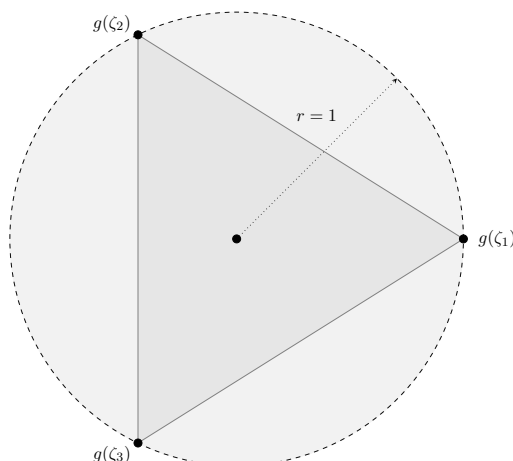
FIGURE 2. A triangular image of a complex-valued function g contained in a disk of radius 1, with three points on the boundary of a disk.

The remarkable work of Crouzeix in [3] establishes that (12) holds with $K \leq 11.08$. A subsequent work of Crouzeix and Palencia in [4] improves the estimate to $K \leq 1+\sqrt{2}$. The Neumann–Poincaré operator appears as an essential tool in all of the mentioned works. The standing conjecture of Crouzeix from [2] is that the bound holds with $K = 2$. This bound is presently known to hold in the case $\mathcal{H}$ being of dimension 2, and has been established by Crouzeix in [2].

Our interest in the new notions of configuration constants is inspired by a recent work of Schwenninger and de Vries in [17], where bounds for general homomorphisms between uniform algebras and the algebras of bounded linear operators are studied. In Section 6 we will combine their arguments with the methods of Crouzeix-Palencia to obtain the following estimate:

$$(13) \qquad \|p(T)\| \leq \left(1 + \sqrt{1 + a(W)}\right)\|p\|_W, \quad W := \overline{W(T)}.$$

For instance, if W is a disk, then $a(W) = 0$, which gives the Okubo-Ando result mentioned above. In [17], Schwenninger and de Vries recovered this result also. The estimate (13) is our motivation for the following investigation of the configuration constants $c_{\mathbb{R}}(\Omega), c_{\mathbb{C}}(\Omega)$ and $a(\Omega)$, and the relations between them.

1.4. Main results.

1.4.1. *Relation between the real and complex constants.* Consider the situation in Figure 2, where the triangular image of the complex-valued function $g : \partial\Omega \to \mathbb{C}$ is contained in a disk of radius 1, and intersects the boundary circle of the disk in three distinct points. The three-point set $\{g(\zeta_1), g(\zeta_2), g(\zeta_3)\}$ is not contained in any open half-circle of the boundary, and it follows from a simple geometric argument (which we shall present in the proofs below) that $\|g + \mathbb{C}\mathbf{1}\|_{\partial\Omega} = 1$. However, the sides of the triangular image of g are all of lengths strictly less than 2, and this implies that

$$\|g + \mathbb{C}\mathbf{1}\|_{\partial\Omega} = 1 > \max_{\zeta,\zeta'\in\partial\Omega} \frac{|g(\zeta) - g(\zeta')|}{2}.$$

If such a function g lies in the image of the unit ball of $C(\partial\Omega)$ under the Neumann–Poincaré operator K_Ω for some domain Ω which satisfies $c_{\mathbb{R}}(\Omega) < 1$, then a strict inequality $c_{\mathbb{R}}(\Omega) < c_{\mathbb{C}}(\Omega)$ occurs. Our first main result excludes this possibility, and so establishes the simplest possible relation between the real and complex configuration constants.

Theorem 1. *The equality*

$$c_{\mathbb{R}}(\Omega) = c_{\mathbb{C}}(\Omega)$$

holds for every compact convex domain Ω with non-empty interior.

It follows that every considered domain has a well-defined *configuration constant* $c(\Omega)$, which is equal to the operator norm of K_Ω on $C(\partial\Omega)/\mathbb{C}\mathbf{1}$, and which can be computed according to the right-hand side of (8). An important consequence of this result is the inequality

$$(14) \qquad\qquad a(\Omega) \le c(\Omega) = \sup_{\zeta,\zeta' \in \partial\Omega} \frac{\|\mu_\zeta - \mu_{\zeta'}\|}{2},$$

which, as we shall soon see, has some interesting consequences.

Theorem 1 doesn't appear nearly as straightforward to prove as it is to state, and the proof takes up a large portion of the article. However, the only property of the Neumann-Poincaré operator used in the proof is that its integral kernel $\{\mu_\zeta\}_{\zeta \in \partial\Omega}$ consists of real-valued measures. In fact, the theorem will be deduced as a corollary of a result which we call the *Three-measures theorem*, and which is a general statement regarding the geometry of the space $C(X)$ of continuous functions on a compact Hausdorff space X. This result, which we discuss and prove in Section 2, puts a restriction on the possible configurations of point sets in the plane which arise as values of a collection of real-valued functionals on $C(X)$.

1.4.2. *Analytic Neumann's lemma.* Note that the above estimate in (14), together with Neumann's lemma, implies that $a(\Omega) < 1$ whenever Ω is not a triangle or a quadrilateral. This can be improved, for we have an analytic version of Neumann's lemma, in which no exceptional cases occur.

Theorem 2. *The strict inequality*

$$a(\Omega) < 1$$

holds for every compact convex domain Ω with non-empty interior.

Our proof of Theorem 2 is much different from the one given by Schober in his proof of the *real* Neumann's lemma in [16], but it works also in the real context. At the end of Section 4 we show how our technique leads to a different proof of Neumann's lemma.

Together with the estimate (13), Theorem 2 shows that for any fixed operator $T : \mathcal{H} \to \mathcal{H}$, the optimal constant in (12) is always strictly smaller than $1+\sqrt{2}$. This improves the Crouzeix-Palencia bound, although by an indefinite amount. More precisely, we have

$$\|p(T)\| \le K_W \|p\|_W$$

with a constant

$$K_W < 1 + \sqrt{2}$$

which depends only on the shape of $W = \overline{W(T)}$, and not on the operator T itself. We show in Section 5 that no better universal bound can be obtained by means of the analytic configuration constant: for any $\epsilon > 0$ there exists a "thin" quadrilateral Ω_ϵ for which we have $a(\Omega_\epsilon) > 1 - \epsilon$.

1.4.3. *Estimates for the configuration constants.* In Section 5 we present also other computations and estimates for the configuration constants. Surprisingly, in the case of an elliptical domain, the configuration constant is computable exactly, and we obtain

$$c(\Omega_{a,b}) = \frac{2}{\pi} \arctan\Big(\frac{1}{2}\Big|\frac{b}{a} - \frac{a}{b}\Big|\Big) = \frac{2}{\pi} \arctan\Big(\frac{1}{2}\frac{e^2}{\sqrt{1-e^2}}\Big)$$

where a and b are lengths of the semi-axes of the ellipse $\Omega_{a,b}$, and e is the eccentricity of the ellipse, given by $e := \sqrt{1 - b^2/a^2}$ in case that $a \ge b$. This fact, together with Theorem 1, estimate (13), and the inequality $a(\Omega) \le c(\Omega)$, has the following consequence.

Corollary 3. *Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator on a Hilbert space $\mathcal{H}$ with numerical range contained in (or equal to) the ellipse $\Omega_{a,b}$. Then, for every polynomial p, we have*

$$\|p(T)\| \le K(a,b)\|p\|_{\Omega_{a,b}}.$$

where

$$K(a,b) := 1 + \sqrt{1 + \frac{2}{\pi}\arctan\left(\frac{1}{2}\left|\frac{b}{a} - \frac{a}{b}\right|\right)}.$$

Note that the function $a \mapsto K(a,1)$ is continuous and increasing for $a \ge 1$, and we have

$$\lim_{a \to \infty} K(a,1) = 1 + \sqrt{2}, \quad \lim_{a \to 1} K(a,1) = 2.$$

Hence the estimate in Corollary 3 gets worse as the eccentricity of the ellipse $\Omega_{a,b}$ grows, and approaches the Crouzeix-Palencia bound in the limit $a \to \infty$. On the other hand, as $a \to 1$, the eccentricity of the ellipse $\Omega_{a,1}$ tends to 0. The estimate is then close to the conjectured optimal bound $K = 2$ and coincides with the Okubo-Ando bound for $a = 1$, in which case the domain is a disk. From this perspective, Corollary 3 may be interpreted as an elliptical generalization of the Okubo-Ando estimate.

For many other types of domains, the exact value of $c(\Omega)$ is inaccessible. To help the situation, we establish an integral estimate which gives an upper bound on $c(\Omega)$ in terms of the curvature of $\partial\Omega$, roughly speaking. For a fixed σ which is not a corner of $\partial\Omega$, recall the definition of $R_{\zeta,\sigma}$ in (5), and consider

$$(15) \qquad R_\Omega(\sigma) := \sup_{\zeta \in \partial\Omega} R_{\zeta,\sigma}.$$

If $\kappa(\sigma)$ is the curvature of $\partial\Omega$ at σ, then $R_\Omega(\sigma)$ is at least as large as the *radius of curvature*

$$(16) \qquad 1/\kappa(\sigma) = \lim_{\zeta \to \sigma} R_{\zeta,\sigma},$$

which is also the radius of the osculating circle at σ. Geometrically, $R_\Omega(\sigma)$ is the radius of the smallest disk tangent to $\partial\Omega$ at σ which contains Ω, if such a disk exists, and it is equal to ∞ otherwise. The latter case occurs, for instance, if σ lies on a straight line segment contained in $\partial\Omega$. However, if $\partial\Omega$ is sufficiently curved on a segment of $\partial\Omega$, then R_Ω will be bounded above there. We obtain in such a situation a non-trivial upper bound on $c(\Omega)$.

Theorem 4. *With the above notation, we have the estimate*

$$c(\Omega) \le 1 - \frac{1}{2\pi}\int_{\partial\Omega} \frac{ds}{R_\Omega}.$$

The result implies spectral constant estimates similar to the one in Corollary 3 above. It also generalizes some similar results in the literature. See Section 5 for further details and examples.

1.4.4. *An unresolved matter.* We have mentioned above that $c(\Omega) = 0$ if and only if Ω is a disk. With some additional effort, we will show in Section 5 that the condition $a(\Omega) = 0$ also characterizes disks. In this case, we have the equality $a(\Omega) = c(\Omega)$. It is natural to ask whether other domains exist for which the equality occurs, or if the case of the disk is exceptional.

Question. *Do we always have the strict inequality*

$$a(\Omega) < c(\Omega)$$

whenever Ω is not a disk?

As a consequence of Theorem 2 and the exceptional cases of Neumann's lemma, we see that the strict inequality holds whenever Ω is a triangle or a quadrilateral. The authors have not been able to confirm that the inequality holds in any other examples.

1.5. **Notations.** Some of our notation has already been introduced above. For a continuous function f defined on a set X, we denote by $\|f\|_X$ the supremum of $|f|$ over X. For cosets of the form $f + \mathbb{C}\mathbf{1}$ we use the convention

$$\|f + \mathbb{C}\mathbf{1}\|_X := \inf_{\lambda \in \mathbb{C}} \|f + \lambda\|_X,$$

with similar convention for real-valued f and cosets $f + \mathbb{R}\mathbf{1}$. A norm $\|\cdot\|$ without a subscript usually denotes a linear functional norm or a total variation norm of a measure. The distinction will be unimportant and should anyway be easy to deduce from context. We use boldface letters, such as $\mathbf{x}$, to denote vectors in $\mathbb{R}^n$, and plain letters, such as x_j, to denote the coordinates.

2. The three-measures theorem

2.1. **Definitions of relevant spaces and operators.** Theorem 1 will be proved as a corollary of our analysis of three-point configurations

$$\big(\ell_1(x), \ell_2(x), \ell_3(x)\big) \in \mathbb{C}^3,$$

where x is an element of a given normed space $\mathcal{N}$, and $\ell_1, \ell_2, \ell_3 \in \mathcal{N}^*$ are three bounded linear functionals on $\mathcal{N}$. A point configuration of this type has to satisfy certain conditions. For instance, we must have the distance bound

$$|\ell_j(x) - \ell_k(x)| \leq \|\ell_j - \ell_k\|_{\mathcal{N}^*} \|x\|_{\mathcal{N}}, \quad 1 \leq j, k \leq 3.$$

Our principal interest will be in estimating the radius of the smallest disk which contains such a three-point set.

In order to use the tools of functional analysis, we will formulate our problem as one of estimating the norm of an operator between normed spaces. To this end, we use the space $\mathbb{C}^3$ of triples of complex numbers, and we equip it with the following norm:

$$(17) \qquad\qquad \|(a, b, c)\|_\infty := \max\{|a|, |b|, |c|\}.$$

Similarly to our previous notational conventions, we shall set $\mathbf{1} := (1, 1, 1) \in \mathbb{C}^3$. The quotient norm in the quotient space $\mathbb{C}^3/\mathbb{C}\mathbf{1}$ satisfies

$$\|(a, b, c) + \mathbb{C}\mathbf{1}\|_\infty := \min_{\lambda \in \mathbb{C}} \max\{|a - \lambda|, |b - \lambda|, |c - \lambda|\},$$

and it has the geometric interpretation adequate to our problem: it is the radius of the smallest disk containing the three point set $\{a, b, c\}$. Given a normed space $\mathcal{N}$ and three linear functionals $\ell_1, \ell_2, \ell_3 \in \mathcal{N}^*$, we introduce the linear operator $\mathcal{L} : \mathcal{N} \to \mathbb{C}^3/\mathbb{C}\mathbf{1}$ defined by

$$(18) \qquad\qquad \mathcal{L}x := \big(\ell_1(x), \ell_2(x), \ell_3(x)\big) + \mathbb{C}\mathbf{1}.$$

With these conventions, each three-point configuration $\big(\ell_1(x), \ell_2(x), \ell_3(x)\big)$ is contained in a disk of radius at most $\|\mathcal{L}\|_{\mathcal{N} \to \mathbb{C}^3/\mathbb{C}\mathbf{1}} \cdot \|x\|_{\mathcal{N}}$. We want to estimate the operator norm $\|\mathcal{L}\|_{\mathcal{N} \to \mathbb{C}^3/\mathbb{C}\mathbf{1}}$.

2.2. **Statement of the theorem.** Without any information regarding the space $\mathcal{N}$ or the functionals ℓ_1, ℓ_2, ℓ_3, the optimal estimate is

$$(19) \qquad\qquad \|\mathcal{L}\|_{\mathcal{N} \to \mathbb{C}^3/\mathbb{C}\mathbf{1}} \leq \frac{1}{\sqrt{3}} \max_{j,k} \|\ell_j - \ell_k\|.$$

Indeed, we see that we cannot do better by choosing $\mathcal{N} = \mathbb{C}$, $x = 1$, and the functionals (scalars) to be the vertices of an equilateral triangle inscribed in the unit circle. For instance,

$$\ell_1 = 1, \quad \ell_2 = -1/2 + i\sqrt{3}/2, \quad \ell_3 = -1/2 - i\sqrt{3}/2.$$

The sides of the triangle have the common length equal to $|\ell_i - \ell_j| = \sqrt{3}$, and the smallest disk containing the three points $\ell_i(x) = \ell_i$ is the unit disk itself. Thus, in this case, (19) holds with

equality. The estimate holds in general as a consequence of *Jung's theorem*, which appeared first in [10], and which in the context of the plane says that any set of diameter d is contained in a disk of radius $d/\sqrt{3}$. In our setting $d \leq \max_{j,k} \|\ell_j - \ell_k\|_{\mathcal{N}^*}$, and so the estimate (19) follows from Jung's theorem.

In our intended application, the role of the space $\mathcal{N}$ is played by $C(X)$, the Banach space of continuous functions on a compact Hausdorff space X, and the functionals are given by integration against *real-valued* measures

$$f \mapsto \mu_j(f) := \int_X f \, d\mu_j.$$

It turns out that the three-point configurations which arise in this way are contained in disks of radius smaller than predicted by Jung's theorem. The main result of the section is the following.

Theorem 5. *Let $C(X)$ be the space of continuous functions on a compact Hausdorff space X, and $\mathcal{L} : C(X) \to \mathbb{C}^3/\mathbb{C}\mathbf{1}$ be the operator in (18) defined by three functionals induced by three finite real-valued Borel measures μ_1, μ_2, μ_3. Then*

$$\tag{20} \|\mathcal{L}\|_{C(X) \to \mathbb{C}^3/\mathbb{C}\mathbf{1}} = \frac{1}{2} \max_{j,k} \|\mu_j - \mu_k\|.$$

It is the "$\leq$" estimate in (20) that is the critical one. The lower bound "$\geq$" follows from the definition of the functional norm

$$\frac{1}{2} \|\mu_j - \mu_k\| = \frac{1}{2} \sup_{f : \|f\|_X = 1} |\mu_j(f) - \mu_k(f)| \leq \|\mathcal{L}\|_{C(X) \to \mathbb{C}^3/\mathbb{C}\mathbf{1}}.$$

We will spend the rest of the section on proving Theorem 5. The outline of the proof is as follows. We will first use duality to formulate the problem in terms of the adjoint operator $\mathcal{L}^*$ between the dual spaces. Next, a discretization will help us reduce the dual problem to a finite-dimensional optimization problem. Finally, we will solve the finite-dimensional problem by the use of techniques of convex analysis.

Before proceeding, we remark that the natural generalization of the above theorem to an arbitrary n-tuple of real-valued measures is valid. See Theorem 15 below.

2.3. Dual problem.
Let us denote by Y the space $\mathbb{C}^3/\mathbb{C}\mathbf{1}$ equipped with the norm in (17). Then the dual space Y^* is the two-dimensional space of three-tuples (α, β, γ) of complex numbers which satisfy

$$\alpha + \beta + \gamma = 0,$$

and the norm on Y^* is given by

$$\|(\alpha, \beta, \gamma)\|_1 := |\alpha| + |\beta| + |\gamma|.$$

In the case $\mathcal{N} = C(X)$, the dual space $(C(X))^*$ is just the space of finite Borel measures on X. The adjoint operator $\mathcal{L}^* : Y^* \to \mathcal{N}^*$ is then given by

$$\mathcal{L}^* : (\alpha, \beta, \gamma) \mapsto \alpha\mu_1 + \beta\mu_2 + \gamma\mu_3$$

and the estimate (20) is equivalent to

$$\tag{21} \|\alpha\mu_1 + \beta\mu_2 + \gamma\mu_3\| \leq \frac{|\alpha| + |\beta| + |\gamma|}{2} \max_{j,k} \|\mu_j - \mu_k\|.$$

Since $\alpha + \beta = -\gamma$ and $(\alpha + \beta + \gamma)\mu_3 = 0$, we may rewrite the above inequality into

$$\tag{22} \|\alpha\nu_1 + \beta\nu_2\| \leq \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\left\{ \|\nu_1\|, \|\nu_2\|, \|\nu_1 - \nu_2\| \right\}$$

where

$$\nu_1 := \mu_1 - \mu_3, \quad \nu_2 := \mu_2 - \mu_3.$$

Note that ν_1 and ν_2 are real-valued if μ_1, μ_2, μ_3 are real-valued. Theorem 5 is thus a consequence of the following slightly more general statement in which the topological structure of X does not play a role.

Proposition 6. *Let ν_1 and ν_2 be two finite real-valued measures on a measurable space X. Then for any complex numbers α, β we have the inequality*

$$(23) \qquad \|\alpha\nu_1 + \beta\nu_2\| \leq \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\Big\{ \|\nu_1\|, \|\nu_2\|, \|\nu_1 - \nu_2\| \Big\},$$

where the norm on the right-hand side is the total variation norm $\|\mu\| := |\mu|(X)$.

In our next step, we shall simplify the problem further, and show that Proposition 6 can be established by considering finite sets X only.

2.4. **Discretization.** With notations as in Proposition 6, set $\sigma := |\nu_1| + |\nu_2|$. Then σ is a positive finite measure on X, and by the Radon–Nikodym theorem we have $d\nu_1 = f \, d\sigma$ and $d\nu_2 = g \, d\sigma$, where f, g are bounded real measurable functions on X. For a moment, let $\|\cdot\|_{\sigma,1}$ denote the norm

$$\|f\|_{\sigma,1} := \int_X |f| \, d\sigma.$$

Then Proposition 6 is equivalent to the inequality

$$(24) \qquad \int_X |\alpha f + \beta g| \, d\sigma \leq \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\Big\{ \|f\|_{\sigma,1}, \|g\|_{\sigma,1}, \|f - g\|_{\sigma,1} \Big\}.$$

We will say that a function is *simple* if it only takes on a finite number of distinct values. By standard measure theory, there exist simple measurable real functions f_m, g_m on X such that $f_m \to f$ and $g_m \to g$ uniformly on X. Clearly $\|\alpha f_m + \beta g_m\|_{\sigma,1} \to \|\alpha f + \beta g\|_{\sigma,1}$. Likewise $\|f_m\|_{\sigma,1} \to \|f\|_{\sigma,1}$ and $\|g_m\|_{\sigma,1} \to \|g\|_{\sigma,1}$ and $\|f_m - g_m\|_{\sigma,1} \to \|f - g\|_{\sigma,1}$. Thus, if the inequality (24) holds for each pair of simple functions, then it holds for f, g. So it suffices to establish (24) when f, g are simple measurable real functions.

Hence, suppose that f, g are simple measurable real functions on X. We can write them as $f = \sum_{j=1}^n a_j 1_{X_j}$ and $g = \sum_{j=1}^n b_j 1_{X_j}$, where $\{X_1, \ldots, X_n\}$ is a measurable partition of X, and $a_j, b_j \in \mathbb{R}$ for all j. The inequality in (24) becomes

$$\sum_{j=1}^n |\alpha a_j + \beta b_j| \sigma(X_j) \leq \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\Big\{ \sum_{j=1}^n |a_j| \sigma(X_j), \sum_{j=1}^n |b_j| \sigma(X_j), \sum_{j=1}^n |a_j - b_j| \sigma(X_j) \Big\}.$$

Writing

$$x_j := a_j \sigma(X_j), \quad \mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$$

and

$$y_j := b_j \sigma(X_j), \quad \mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$$

we see that this becomes

$$\|\alpha\mathbf{x} + \beta\mathbf{y}\|_1 \leq \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\Big\{ \|\mathbf{x}\|_1, \|\mathbf{y}\|_1, \|\mathbf{x} - \mathbf{y}\|_1 \Big\},$$

where now $\mathbf{x}, \mathbf{y}$ are vectors in $\mathbb{R}^n$ and $\|\cdot\|_1$ denotes the usual ℓ^1-norm on $\mathbb{R}^n$ given by

$$(25) \qquad \|\mathbf{x}\|_1 := \sum_{j=1}^n |x_j|.$$

To summarize, to prove Proposition 6 and consequently to prove Theorem 5, it suffices to establish the following discrete result.

Proposition 7. *Let $n \geq 1$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then for all complex numbers α, β we have the inequality*

$$(26) \qquad \|\alpha\mathbf{x} + \beta\mathbf{y}\|_1 \leq \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\left\{\|\mathbf{x}\|_1, \|\mathbf{y}\|_1, \|\mathbf{x} - \mathbf{y}\|_1\right\}.$$

This reduction of the problem to the finite-dimensional setting allows us to use the tools of convex analysis.

2.5. **Optimization over a convex set.** Consider the set

$$(27) \qquad C_n := \left\{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n : \|\mathbf{x}\|_1 \leq 1, \|\mathbf{y}\|_1 \leq 1, \|\mathbf{x} - \mathbf{y}\|_1 \leq 1\right\}.$$

Thus C_n is a compact convex polytope in $\mathbb{R}^n \times \mathbb{R}^n$, and so it has a finite number of extreme points. That is, points of C_n which do not lie in the interior of any line segment in C_n. A well-known theorem of Carathéodory says that each point of a compact convex polytope is a convex combination of its extreme points.

Lemma 8. *In order to establish Proposition 7, it suffices to show that the inequality (26) holds for every extreme point of C_n.*

Proof. Let us fix $\alpha, \beta \in \mathbb{C}$ and $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n$. By the homogeneity of the inequality in (26), we may assume that

$$(28) \qquad \max\left\{\|\mathbf{x}\|_1, \|\mathbf{y}\|_1, \|\mathbf{x} - \mathbf{y}\|_1\right\} = 1.$$

Then $(\mathbf{x}, \mathbf{y}) \in C_n$ and so we may express it as a convex combination of the extreme points of C_n, namely

$$\mathbf{x} = \sum_{k=1}^m t_k \mathbf{e^k}, \quad \mathbf{y} = \sum_{k=1}^m t_k \mathbf{f^k}$$

where $\mathbf{e^k} \in \mathbb{R}^n$, $\mathbf{f^k} \in \mathbb{R}^n$, the pairs $(\mathbf{e^k}, \mathbf{f^k})$ are extreme points of C_n, $t_k > 0$, and $\sum_{k=1}^m t_k = 1$. Note that since $(\mathbf{e^k}, \mathbf{f^k})$ is an extreme point of C_n, we must have

$$\max\left\{\|\mathbf{e^k}\|_1, \|\mathbf{f^k}\|_1, \|\mathbf{e^k} - \mathbf{f^k}\|_1\right\} = 1.$$

Since we are assuming that (26) holds for extreme points, we can estimate

$$\begin{aligned}
\|\alpha\mathbf{x} + \beta\mathbf{y}\|_1 &\leq \sum_{k=1}^m t_k \|\alpha\mathbf{e^k} + \beta\mathbf{f^k}\|_1 \\
&\leq \sum_{k=1}^m t_j \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \max\left\{\|\mathbf{e^k}\|_1, \|\mathbf{f^k}\|_1, \|\mathbf{e^k} - \mathbf{f^k}\|_1\right\} \\
&= \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2} \sum_{k=1}^m t_k \\
&= \frac{|\alpha| + |\beta| + |\alpha + \beta|}{2}.
\end{aligned}$$

Recalling our normalization in (28), this is the desired estimate in (26). $\qquad\square$

From the above lemma and our sequence of reductions above, it follows that in order to prove Theorem 5 it suffices show that the inequality (26) holds at every extreme point of the polytope C_n. Proposition 9 below characterizes these extreme points by partitioning them into three equivalence classes.

Note that C_n is invariant under the following linear symmetries:

$$(29) \qquad \begin{cases} x'_j & := x_{\pi(j)} \\ y'_j & := y_{\pi(j)} \end{cases}$$

where π is any permutation of $\{1, 2, \dots, n\}$,

$$(30) \qquad \begin{cases} x'_j & := \epsilon_j x_j \\ y'_j & := \epsilon_j y_j \end{cases}$$

for any choice of $\epsilon_1, \dots, \epsilon_n \in \{-1, 1\}$,

$$(31) \qquad \begin{cases} \mathbf{x}' & := \mathbf{y} \\ \mathbf{y}' & := \mathbf{x} \end{cases}$$

and

$$(32) \qquad \begin{cases} \mathbf{x}' & := \mathbf{x} \\ \mathbf{y}' & := \mathbf{x} - \mathbf{y} \end{cases}$$

Denote by G_n the group generated by these symmetries. As these symmetries are linear automorphisms of $\mathbb{R}^n \times \mathbb{R}^n$, it is clear that G_n leaves invariant the set of extreme points C_n. We say that two extreme points of C_n are G_n-*equivalent* if there is an element of G_n mapping one of them to the other. Thus the action of G_n on C_n partitions the set of extreme points of C_n into a finite number of equivalence classes. Note that if the inequality (26) holds for some $(\mathbf{x}, \mathbf{y}) \in C_n$, then it holds also for any point of C_n in the orbit of $(\mathbf{x}, \mathbf{y})$ under the group action of G_n on C_n.

The extreme points of C_n are identified in the following proposition.

Proposition 9. *If $n \geq 3$, then every extreme point $(\mathbf{x}, \mathbf{y})$ of C_n is G_n-equivalent to one of the pairs*

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad and \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad and \quad \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

One can readily check that each of the three above pairs really is an extreme point of C_n. We omit the proof, since we do not actually need this fact. In the case that $n = 1$, the same result holds, but only the first kind of pair can arise. Likewise, if $n = 2$, the same result holds, but only the first two types of pairs can arise.

We will prove Proposition 9 in Section 2.6. For now let us see how Theorem 5 follows. In order to verify (26) for all extreme points of C_n, it suffices to verify the inequality for the three pairs of vectors appearing in Proposition 9. This is an easy task. For instance, if $(\mathbf{x}, \mathbf{y})$ is the second pair in Proposition 9, then we have

$$\begin{aligned} \|\alpha \mathbf{x} + \beta \mathbf{y}\|_1 &= |\alpha + \beta/2| + |\beta|/2 \\ &= |\alpha/2 + \alpha/2 + \beta/2| + |\beta|/2 \\ &\leq |\alpha + \beta|/2 + |\alpha|/2 + |\beta|/2. \end{aligned}$$

The inequality for the other two pairs is verified similarly. Then from Lemma 8 we conclude that Proposition 7 holds, from which Theorem 5 follows by the earlier reduction.

It remains to prove Proposition 9.

2.6. **Extreme points of the polytope.** In the proof of Proposition 9, the group G_n generated by the symmetries (29)-(32) will be extensively used. In particular we will use the property that $(\mathbf{x}, \mathbf{y})$ is an extreme point of C_n if and only if some extreme point of C_n is G_n-equivalent to it. Moreover, the following two observations will be useful to single out.

Lemma 10. *If for a pair* $(\mathbf{x}, \mathbf{y}) \in C_n$ *there exists two distinct indices* j, k *such that*

$$x_j > 0, \quad y_j > 0, \quad x_k > 0, \quad y_k > 0$$

then $(\mathbf{x}, \mathbf{y})$ *is not an extreme point of* C_n.

   More generally, if for two distinct indices j, k *we have that two of the quantities* $x_j x_k$, $y_j y_k$ *and* $(x_j - y_j)(x_k - y_k)$ *are non-zero and have the same sign, then* $(\mathbf{x}, \mathbf{y})$ *is not an extreme point of* C_n.

Proof. Using the symmetry (29) we may suppose that $j = 1$, $k = 2$. Note that $x_1 < 1$, $x_2 < 1$, since $\|\mathbf{x}\|_1 \leq 1$. The same is true for the corresponding coordinates of $\mathbf{y}$. Let $\mathbf{d} = (1, -1, 0, \ldots, 0)^T \in \mathbb{R}^n$. It is easy to verify that if t is a real number, and $|t|$ is sufficiently small, then we have

$$(\mathbf{x}, \mathbf{y}) + t(\mathbf{d}, \mathbf{d}) = (\mathbf{x} + t\mathbf{d}, \mathbf{y} + t\mathbf{d}) \in C_n.$$

Thus $(\mathbf{x}, \mathbf{y})$ lies on a line segment inside C_n, and so is not an extreme point of C_n.

   The more general statement follows by applications of a sequence of symmetries in (29)-(32) to transform $(\mathbf{x}, \mathbf{y})$ satisfying the more general assumption into a point $(\mathbf{x}', \mathbf{y}')$ where the first two coordinates of the vectors $\mathbf{x}'$ and $\mathbf{y}'$ are positive. □

Lemma 11. *If for a pair* $(\mathbf{x}, \mathbf{y}) \in C_n$ *the vector* $\mathbf{x}$ *or* $\mathbf{y}$ *has at least three non-zero coordinates, then* $(\mathbf{x}, \mathbf{y})$ *is not an extreme point of* C_n.

Proof. By using symmetries (29)-(31) we may suppose that coordinates x_1, x_2, x_3 are non-zero and positive. If two of the coordinates y_1, y_2, y_3 are positive, then by Lemma 10 we conclude that $(\mathbf{x}, \mathbf{y})$ is not an extreme point of C_n. In the contrary case, two of the coordinates y_1, y_2, y_3 are non-negative. Then again by Lemma 10 and the symmetry (32) the pair $(\mathbf{x}, \mathbf{x} - \mathbf{y}) \in C_n$ is not extreme, and thus neither is $(\mathbf{x}, \mathbf{y})$, since these two pairs are G_n-equivalent. □

   We are ready to prove Proposition 9. We denote by ℓ_n^1 the space $\mathbb{R}^n$ equipped with the norm $\| \cdot \|_1$ given by (25). Recall that the extreme points of the unit ball $B := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_1 \leq 1\}$ are the vectors with precisely one non-zero coordinate, this coordinate being equal to ± 1.

Proof of Proposition 9. We will split up the proof into three cases, each case corresponding to one of the pairs in the statement of the proposition.

   Case 1: At least one of the norms $\|\mathbf{x}\|_1, \|\mathbf{y}\|_1, \|\mathbf{x} - \mathbf{y}\|_1$ is strictly less than 1. We will show that in this case $(\mathbf{x}, \mathbf{y})$ is G_n-equivalent to the first pair in the statement of the proposition.

   By applying a suitable combination of symmetries (29)-(32), we may suppose that in fact $\|\mathbf{x} - \mathbf{y}\|_1 < 1$. We claim that $\mathbf{x}$ must be an extreme point of the unit ball of ℓ_n^1. For if not, then it lies at the midpoint of a line segment I such that $\|\mathbf{x}'\|_1 \leq 1$ for all $\mathbf{x}' \in I$. Since $\|\mathbf{x} - \mathbf{y}\|_1 < 1$, by shrinking I if necessary, we also have $\|\mathbf{x}' - \mathbf{y}\|_1 < 1$ for all $\mathbf{x}' \in I$. Thus $I \times \{\mathbf{y}\}$ is a line segment in C_n with interior point $(\mathbf{x}, \mathbf{y})$, contradicting the fact that $(\mathbf{x}, \mathbf{y})$ is extreme.

   Likewise, $\mathbf{y}$ is extreme in the unit ball of ℓ_n^1. Applying a suitable symmetry, we may suppose that $x_1 = 1$ and $y_j = \pm 1$ for some j, all the other entries of $\mathbf{x}$ and $\mathbf{y}$ being 0. Since we must have $\|\mathbf{x} - \mathbf{y}\|_1 < 1$, this implies that actually $j = 1$ and $y_1 = 1$. Thus $(\mathbf{x}, \mathbf{y})$ is equivalent to the first pair of vectors listed in the statement of the proposition. This concludes Case 1.

   Case 2: We have $\|\mathbf{x}\|_1 = \|\mathbf{y}\|_1 = \|\mathbf{x} - \mathbf{y}\|_1 = 1$, and one of the vectors $\mathbf{x}$, $\mathbf{y}$ or $\mathbf{x} - \mathbf{y}$ has only one non-zero coordinate. In this case, $(\mathbf{x}, \mathbf{y})$ will be now shown to be G_n-equivalent to the second pair in the statement of the proposition.

Using our symmetries, we may suppose that $\mathbf{x} = (1, 0, \dots, 0)^T$. Note that

$$\|\mathbf{x} - \mathbf{y}\|_1 = |1 - y_1| + |y_2| + \dots + |y_n| = 1$$

and

$$\|\mathbf{y}\|_1 = |y_1| + |y_2| + \dots + |y_n| = 1$$

force

$$|1 - y_1| = |y_1|,$$

the unique real solution y_1 to this equation being $y_1 = 1/2$. By Lemma 11, $\mathbf{y}$ has only one other non-zero coordinate, and $\|\mathbf{y}\|_1 = 1$ forces this coordinate to be equal to $\pm 1/2$. Applying symmetries (29) and (30) we conclude that $(\mathbf{x}, \mathbf{y})$ is G_n-equivalent to the second pair in the statement. This concludes Case 2.

Case 3: We have $\|\mathbf{x}\|_1 = \|\mathbf{y}\|_1 = \|\mathbf{x} - \mathbf{y}\|_1 = 1$, and all of the vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{x} - \mathbf{y}$ have exactly two non-zero coordinates. We will show that $(\mathbf{x}, \mathbf{y})$ is G_n-equivalent to the third pair in the statement of the proposition.

This case is slightly more complicated than the previous two. As before, we may suppose that $x_1 > 0$ and $x_2 > 0$. We claim that y_1 and y_2 cannot both be equal to zero. If they were, then $\mathbf{x} - \mathbf{y}$ has four non-zero coordinates, contrary to the assumption. In fact, precisely one of y_1 and y_2 must be non-zero. If both were non-zero, then since $\mathbf{x} - \mathbf{y}$ has exactly two non-zero coordinates, we would have $x_1 - y_1 \neq 0$ and $x_2 - y_2 \neq 0$. Then the three quantities $x_1 x_2$, $y_1 y_2$ and $(x_1 - y_1)(x_2 - y_2)$ would be non-zero, and Lemma 10 would imply that $(\mathbf{x}, \mathbf{y})$ is not an extreme point.

By an application of symmetries we may, in addition to $x_1 > 0$ and $x_2 > 0$, suppose that $y_1 \neq 0$, $y_2 = 0$ and $y_3 = s > 0$. Since $x_1 + x_2 = 1$, we have $x_1 = t$, $x_2 = 1 - t$ for some $t \in (0, 1)$. Our vectors thus have the following structure:

$$x = \begin{pmatrix} t \\ 1 - t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ 0 \\ s \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad x - y = \begin{pmatrix} t - y_1 \\ 1 - t \\ -s \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Recall that $\mathbf{x} - \mathbf{y}$ has only two non-zero coordinates. Since $1 - t \neq 0$ and $s \neq 0$, we conclude from the above that $t = y_1$. But then $\|\mathbf{x} - \mathbf{y}\|_1 = 1 - t + s = 1$, and so $t = s$. Finally, $1 = \|\mathbf{y}\|_1 = t + s = 2s$ shows that $s = t = 1/2$, and so $(\mathbf{x}, \mathbf{y})$ is G_n-equivalent to the third pair in the statement of the proposition. $\qquad \square$

3. Proof of Theorem 1

In addition to Theorem 5 from Section 2, we will also need some facts from plane geometry in order to prove Theorem 1. In particular, we will need to discuss the *minimum enclosing disk problem* appearing in computational geometry.

3.1. **Minimal enclosing disk.** Let K be a compact subset of $\mathbb{C}$ containing at least two points. Among all closed disks which contain K there exists a unique one of minimal radius. We will denote this disk by $\mathbb{D}_K$ and call it the *minimal disk* for K. The radius of $\mathbb{D}_K$ will be denoted by $R(K)$.

If $\mathbb{D}_K$ is minimal for K, then the intersection $K \cap \partial \mathbb{D}_K$ must obviously be non-empty. In fact, this intersection must contain at least two points, and there is also a restriction on the locations of the points in $K \cap \partial \mathbb{D}_K$.
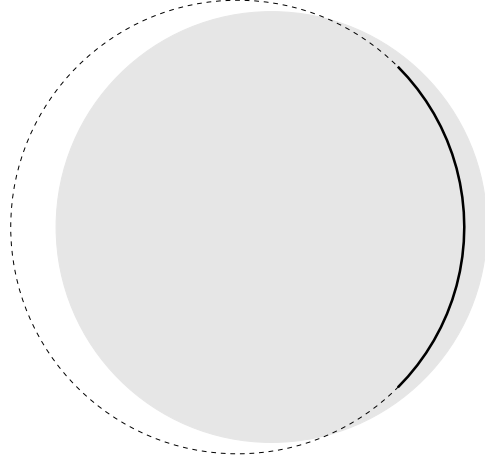
FIGURE 3. The initial disk $\mathbb{D}_K$ is the dashed circle, and we assume that $\partial\mathbb{D}_K \cap K$ is contained in the black thick arc. Then K will be contained in the grey disk which is obtained from $\mathbb{D}_K$ by first translating $\mathbb{D}_K$ in the direction of the positive real axis, and then slightly shrinking the translated disk. This contradicts the minimality of $\mathbb{D}_K$.

Lemma 12. *Let K be a compact subset of $\mathbb{C}$ which contains at least two points. Then the intersection $\partial\mathbb{D}_K \cap K$ is not contained in any arc of $\partial\mathbb{D}_K$ which has length strictly smaller than half of the circumference of $\mathbb{D}_K$. In particular, if $K \cap \partial\mathbb{D}_K = \{a, b\}$ is a two-point set, then a and b are antipodal on $\partial\mathbb{D}_K$*

Proof. Seeking a contradiction, assume that $\partial\mathbb{D}_K \cap K$ is contained in an arc of length strictly less than half of the circumference of $\mathbb{D}_K$. By translation, rescaling and rotation of the setting, we may assume that $\mathbb{D}_K$ is the unit disk, and that $\partial\mathbb{D}_K \cap K$ is contained in some half-space

$$\{z \in \mathbb{C} : \operatorname{Re} z > \delta\}, \quad \delta > 0.$$

By compactness, the distance between the compact sets K and $\partial\mathbb{D}_K \cap \{z \in \mathbb{C} : \operatorname{Re} z \leq \delta/2\}$ is positive. It follows that we may translate the disk $\mathbb{D}_K$ in the positive direction of the real axis, and then shrink the radius of the translated disk slightly, and the resulting disk will still contain K, yet be of strictly smaller radius than R_K. See Figure 3. This contradiction establishes Lemma 12. $\square$

Lemma 13. *Let $T = \{a, b, c\}$ be a three-point set. If D is a closed disk for which $T \subset \partial D$, and T is not contained in any arc of ∂D which is strictly smaller than half of the circumference of D, then $D = \mathbb{D}_T$.*

Proof. Assume, seeking a contradiction, that $D \neq \mathbb{D}_T$, and so that $R(T)$ is strictly smaller than the radius of D. Since ∂D is the unique circle passing through the three points a, b, c, we must have that $T \cap \partial\mathbb{D}_T$ contains precisely two points. Say $a, b \in \partial\mathbb{D}_T$ but $c \notin \partial\mathbb{D}_T$. Lemma 12 implies that a and b are antipodal on $\mathbb{D}_T$. By translation, rescaling and rotation, we may assume that $\mathbb{D}_T$ is the unit disk, $a = i, b = -i$, c has non-negative real part and $|c| < 1$. After these operations, we have that $R(T) = 1$ and the circumference of D is larger than 2π. Thus by hypothesis, surely T is not contained in any arc of ∂D of length strictly smaller than π. But the shorter of the arcs of ∂D which contains T is then contained in $\{z \in \mathbb{C} : 0 \leq \operatorname{Re} z, |z| \leq 1\}$, and so this arc must have a length smaller than π. This is a contradiction, and the lemma follows. $\square$
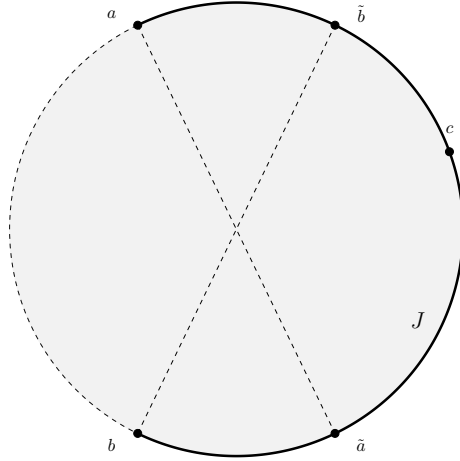
FIGURE 4. The thick arc J between a and b is the smallest containing the compact set K. It follows that the shorter arc between the antipodal points $\tilde{a}$ an $\tilde{b}$ must contain points of K.

3.2. **Reduction to three-point sets.** The following simple result on minimal disks makes it possible to apply Theorem 5 to more than three measures.

Lemma 14. *Let K be a compact subset of $\mathbb{C}$ containing at least two points. There exists a subset $T \subset K$ which contains at most three points and for which $\mathbb{D}_K = \mathbb{D}_T$. In particular, $R(K) = R(T)$.*

It may be convenient to refer to Figure 4 during the reading of the proof.

Proof. If there are two points in K which are antipodal on $\partial \mathbb{D}_K$, then we take T to consist of those two points. Clearly $\mathbb{D}_K = \mathbb{D}_T$. In the case that no pair of antipodal points of $\partial \mathbb{D}_K$ are contained in K, let J be the shortest closed arc of $\partial \mathbb{D}_K$ which contains K, and let $a, b \in J \cap K$ be the end-points of J. By Lemma 12, the length of J is strictly larger than half of the circumference of $\partial \mathbb{D}_K$, and so J is the longer of the two arcs between a and b. Let $\tilde{a}$ and $\tilde{b}$ be points on $\partial \mathbb{D}_K$ which are antipodal to a and b, respectively. By assumption, $\tilde{a} \notin K, \tilde{b} \notin K$. We claim that the shorter of the two open arcs between $\tilde{a}$ and $\tilde{b}$ must contain points of K. If not, then the longer of the two arcs between $\tilde{a}$ and $\tilde{b}$ would contain K in its interior, and this arc has the same length as J. A routine compactness argument would lead to a contradiction to the minimality of J.

Let $T = \{a, b, c\}$, where $c \in K$ is any point contained in the shorter open arc between $\tilde{a}$ and $\tilde{b}$. Note that any arc containing T must contain either $\tilde{a}$ or $\tilde{b}$. Then such an arc contains two antipodal points on $\mathbb{D}_K$, and so it has a length which is at least half of the circumference of $\mathbb{D}_K$. By Lemma 13 we conclude that $\mathbb{D}_K = \mathbb{D}_T$. $\qquad\square$

3.3. **Finalizing the proof.** We are finally ready to give a proof of the equality $c_{\mathbb{R}}(\Omega) = c_{\mathbb{C}}(\Omega)$.

Proof of Theorem 1. Since $c_{\mathbb{R}}(\Omega) \leq c_{\mathbb{C}}(\Omega)$, it will suffice to show the reverse inequality. To this end, we need to show that given $f \in C(\partial\Omega)$ satisfying $\|f\|_{\partial\Omega} \leq 1$, we have that $\|K_\Omega f + \mathbb{C}\mathbf{1}\|_{\partial\Omega} \leq c_{\mathbb{R}}(\Omega)$. Since $K_\Omega f$ is continuous, the image $K = K_\Omega f(\partial\Omega)$ is a compact subset of $\mathbb{C}$. If K consists of a single point, then $\|K_\Omega f + \mathbb{C}\mathbf{1}\|_{\partial\Omega} = 0$, and the proof is complete. In other case, let $\mathbb{D}_K$ be the minimal disk for K. We use Lemma 14 to obtain a three-point set $T = \{a, b, c\} \subset K$ for which $R(T) = R(K)$ (note that if $K \cap \partial\mathbb{D}_K$ contains only two points $\{a, b\}$, then we may pick $c \in K$ arbitrarily to complete T to a three-point set). The geometric interpretation of the quotient norm

in $C(\partial\Omega)/\mathbb{C}\mathbf{1}$ implies that $\|K_\Omega f + \mathbb{C}\mathbf{1}\|_{\partial\Omega} = R(K) = R(T)$. Since T is contained in the image of $K_\Omega f$, there exists $\zeta_1, \zeta_2, \zeta_3 \in \partial\Omega$ such that

$$(a, b, c) = \big(K_\Omega f(\zeta_1), K_\Omega f(\zeta_2), K_\Omega f(\zeta_3)\big).$$

Since $K_\Omega f(\zeta_j) = \int_{\partial\Omega} f \, d\mu_{\zeta_j}$, we may apply Theorem 5 to $X = \partial\Omega$, $\mu_j = \mu_{\zeta_j}$ for $j = 1, 2, 3$, and conclude that the operator $\mathcal{L} : C(\partial\Omega) \to \mathbb{C}^3/\mathbb{C}\mathbf{1}$ defined by

$$\mathcal{L} : f \mapsto \big(K_\Omega f(\zeta_1), K_\Omega f(\zeta_2), K_\Omega f(\zeta_3)\big) + \mathbb{C}\mathbf{1}$$

has a norm satisfying the bound (20). With $\|\cdot\|_\infty$ denoting the norm on $\mathbb{C}^3/\mathbb{C}\mathbf{1}$ given in (17), we obtain

$$\begin{aligned}
\|K_\Omega f + \mathbb{C}\mathbf{1}\|_{\partial\Omega} = R(T) &= \|(a, b, c) + \mathbb{C}\mathbf{1}\|_\infty \\
&= \|\mathcal{L}f\|_\infty \\
&\leq \|\mathcal{L}\|_{C(\partial\Omega) \to \mathbb{C}^3/\mathbb{C}\mathbf{1}} \\
&\leq \frac{1}{2} \max_{j,k} \|\mu_{\zeta_j} - \mu_{\zeta_k}\| \\
&\leq \frac{1}{2} \sup_{\zeta, \zeta' \in \partial\Omega} \|\mu_\zeta - \mu_{\zeta'}\| \\
&= c_\mathbb{R}(\Omega).
\end{aligned}$$

$\square$

The earlier mentioned extension of Theorem 5 to an *n-measures theorem* is obtained by employing the same argument as in the above proof. The normed space $\mathbb{C}^n/\mathbb{C}\mathbf{1}$ appearing below is defined analogously to the case $n = 3$ treated in Section 2.1.

Theorem 15. *Let $C(X)$ be the space of continuous functions on a compact Hausdorff space X, $n \geq 3$ an integer, and $\mathcal{L} : C(X) \to \mathbb{C}^n/\mathbb{C}\mathbf{1}$ the operator defined by*

$$\mathcal{L}f = \big(\mu_1(f), \ldots, \mu_n(f)\big) + \mathbb{C}\mathbf{1}$$

where $\mu_1, \ldots, \mu_n$ are finite real-valued Borel measures on X. Then

$$\|\mathcal{L}\|_{C(X) \to \mathbb{C}^n/\mathbb{C}\mathbf{1}} = \frac{1}{2} \max_{j,k} \|\mu_j - \mu_k\|.$$

Proof. We use Lemma 14 to pick a three-point subset T of $K = \{\mu_j(f)\}_{j=1}^n$ for which we have $R(K) = R(T)$, and apply Theorem 5 as in the preceeding proof. $\square$

4. Proof of Theorem 2

4.1. **Exploiting subsequences.** We will argue by contradiction in order to prove Theorem 2. That is, we will assume that there exists a convex domain Ω with $a(\Omega) = 1$, and so that there exists a sequence of functions (f_n) in $\mathcal{A}(\Omega)$ which satisfy

$$\|f_n + \mathbb{C}\mathbf{1}\|_\Omega = 1$$

and

(33) $$\lim_{n \to \infty} \|K_\Omega f_n + \mathbb{C}\mathbf{1}\|_\Omega = 1.$$

We shall see that this leads to a contradiction. The proof technique below is different from the one employed by Schober in [16] in his proof of Neumann's lemma, and analyticity is used only at the very end of the proof. In fact, we shall remark at the end of the section how our arguments lead to a new proof of Neumann's lemma which is different from the one in [16].

Thus, for now, we assume merely that $f_n \in C(\partial\Omega)$, and we will derive certain consequences of (33). In the course of the proof we shall replace the sequence (f_n) by a subsequence multiple times, and for convenience we will not be changing the subscripts. We may suppose that $\|f_n\|_\Omega = 1$, and consequently that the images

$$K_\Omega f_n(\partial\Omega) := \{K_\Omega f_n(\zeta) : \zeta \in \partial\Omega\}$$

are contained in a closed disk of radius 1 centred at the origin. For large n, this observation and (33) forces there to be points of the image of $K_\Omega f_n$ outside of any disk centred at the origin of radius strictly less than 1. By exchanging f_n for a unimodular multiple of itself, we may thus assume that there exists a sequence of points (ζ_n) in $\partial\Omega$ for which we have

$$(34) \qquad \lim_{n\to\infty} K_\Omega f_n(\zeta_n) = \lim_{n\to\infty} \int_{\partial\Omega} f_n \, d\mu_{\zeta_n} = 1.$$

Using that the functions f_n are bounded by 1 in modulus, and the positive measures $d\mu_\zeta$ are of unit mass, we obtain

$$\lim_{n\to\infty} \int_{\partial\Omega} |f_n - 1|^2 d\mu_{\zeta_n} = \lim_{n\to\infty} \int_{\partial\Omega} \left(|f_n|^2 - 2\operatorname{Re} f_n + 1\right) d\mu_{\zeta_n}$$
$$\leq \lim_{n\to\infty} \left(2 - 2\operatorname{Re} \int_{\partial\Omega} f_n \, d\mu_{\zeta_n}\right) = 0.$$

Recall from (3) that ρ_{ζ_n} denotes the ds-absolutely continuous part of μ_{ζ_n}. The above computation implies that

$$(35) \qquad \lim_{n\to\infty} \int_{\partial\Omega} |f_n - 1|^2 \rho_{\zeta_n} ds = 0.$$

Compactness of the boundary $\partial\Omega$ implies that we may assume convergence of the sequence (ζ_n) to some points $\zeta \in \partial\Omega$. The following lemma shows that we may replace in (35) the densities ρ_{ζ_n} with the density ρ_ζ.

Lemma 16. *With notations as above, we have*

$$(36) \qquad \lim_{n\to\infty} \int_{\partial\Omega} |f_n - 1|^2 \rho_\zeta ds = 0.$$

Consequently, after passing to a subsequence, we can ensure that

$$\lim_{n\to\infty} f_n(\sigma) = 1$$

for almost every $\sigma \in \partial\Omega$ with respect to the measure $\rho_\zeta ds$.

Proof. Note that whenever σ is not a corner of $\partial\Omega$ or any of the points ζ_n or ζ, we have

$$\rho_{\zeta_n}(\sigma) - \rho_\zeta(\sigma) = (\zeta_n - \zeta) \cdot \operatorname{Re} \frac{N(\sigma)}{\pi(\sigma - \zeta)(\sigma - \zeta_n)}.$$

If $B = B(\zeta, \delta)$ is a disk around ζ of small radius $\delta > 0$, then for large enough n the second factor on the right-hand side above is bounded uniformly for $\sigma \in \partial\Omega \setminus B$, with exception of a countable set. This shows uniform convergence of $\rho_{\zeta_n}(\sigma)$ to $\rho_\zeta(\sigma)$ for $\sigma \in \partial\Omega \setminus B$, again with exception of
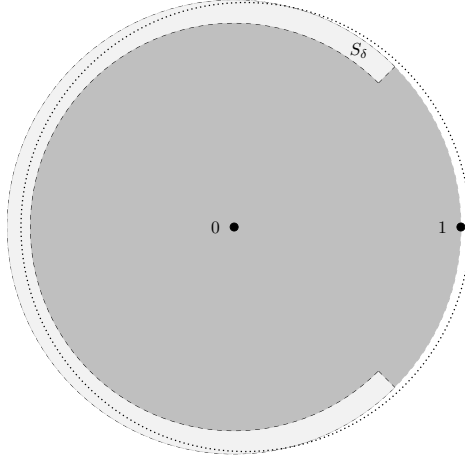
FIGURE 5. The unit disk in dark grey with the strip S_δ removed. The dotted circle containing the dark grey area has a radius slightly smaller than 1.

an at most countable set. Since $|f_n - 1|^2 \le 4$, we obtain from (35) that

$$\limsup_{n\to\infty} \int_{\partial\Omega} |f_n - 1|^2 \rho_\zeta ds \le \limsup_{n\to\infty} \int_{\partial\Omega\cap B} |f_n - 1|^2 \rho_\zeta ds$$
$$+ \limsup_{n\to\infty} \int_{\partial\Omega\setminus B} |f_n - 1|^2 \rho_\zeta ds$$
$$\le 4 \int_{\partial\Omega\cap B} \rho_\zeta ds.$$

Since $\partial\Omega \cap B$ is an arc of length which tends to 0 as the radius δ of B tends to 0, the last quantity above can be made arbitrarily small by choosing δ small enough. This establishes (36). Basic measure theory now implies that we may pass again to a subsequence and ensure the pointwise convergence $f_n \to 1$ almost everywhere with respect to $\rho_\zeta ds$. □

Out next observation extracts more information from (33). Consider the strips

$$S_\delta = \{z = re^{it} : 1 - \delta < r < 1, |t| \in [\pi/4, \pi]\}, \quad \delta > 0.$$

These strips have a fixed large "length" but shrinking "width". One such strip is marked in Figure 5. We claim that each one of the strips S_δ intersects the images $K_\Omega f_n(\Omega)$ non-trivially for infinitely many indices n. For if not, then for some fixed $\delta > 0$, we would have that $S_\delta \cap K_\Omega f_n(\partial\Omega) = \emptyset$ for all sufficiently large n, which means that the images $K_\Omega f_n(\partial\Omega)$ are entirely contained in $B(0,1) \setminus S_\delta$, where $B(0,1)$ denotes the closed disk of radius 1 centred at the origin. But if ϵ_1 and ϵ_2 are sufficiently small positive numbers, then $B(0,1) \setminus S_\delta \subset B(\epsilon_1, 1 - \epsilon_2)$, a disk of radius $1 - \epsilon_2$ centred at the point $\epsilon_1 \in \mathbb{R}$. See Figure 5. Recalling the geometric interpretation of the norm $\|K_\Omega f_n + \mathbb{C}\mathbf{1}\|_{\partial\Omega}$ as the radius of the smallest disk containing the image of $K_\Omega f_n$, we would arrive at a contradiction to (33). Thus every strip S_δ contains points in the image of $K_\Omega f_n$ for infinitely many n.

Lemma 17. *With notations as above, we may pass to a subsequence again, and obtain a new sequence (ζ_n') which converges to some point $\zeta' \in \partial\Omega$, and such that*

$$\lim_{n\to\infty} f_n(\sigma) = \alpha$$

for some unimodular constant $\alpha \ne 1$ and for almost every $\sigma \in \partial\Omega$ with respect to the measure $\rho_{\zeta'} ds$.

Proof. Since each strip S_δ intersects the images of $K_\Omega f_n$ for infinitely many n, passing to a subsequence and a routine compactness argument produces a sequence (ζ_n') convergent to some $\zeta' \in \partial\Omega$, for which $K_\Omega f_n(\zeta_n') \to \alpha$, with α unimodular and lying in the closure of each of the strips S_δ. Thus $\alpha \neq 1$. We therefore merely need to repeat the previous arguments to see that, after passing to a subsequence, we will have $f_n(\sigma) \to \alpha$ for almost every σ with respect to the measure $\rho_{\zeta'} ds$. $\square$

4.2. **Proof of Theorem 2.** The above arguments are valid for $f_n \in C(\partial\Omega)$. However, under the assumption of analyticity, the sequence (f_n) cannot converge to two different constants on two different sets of positive arclength measure. To make this statement precise, we appeal to the classical theory of analytic functions in the (open) unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. Here [8, Chapter II] is an excellent reference for the claims made in the following proof.

Proof of Theorem 2. Let $H^\infty = H^\infty(\mathbb{D})$ be the space of bounded analytic functions in $\mathbb{D}$, identified as usual through boundary function correspondence with a weak-star closed subspace of the space $L^\infty(\partial\mathbb{D}) = (L^1(\partial\mathbb{D}))^*$ of bounded measurable functions on $\partial\mathbb{D}$, the dual of the Lebesgue space $L^1(\partial\mathbb{D})$ of functions integrable on $\partial\mathbb{D}$ with respect to the Lebesgue measure (arclength measure) on $\partial\mathbb{D}$. It is well known that a function $\tilde{f} \in H^\infty$ which vanishes on a subset of positive Lebesgue measure on $\partial\mathbb{D}$ must vanish identically.

Fix some conformal mapping $\phi : \mathbb{D} \to \Omega$. Under the assumption that $f_n \in \mathcal{A}(\Omega)$, $\|f_n\|_\Omega \leq 1$, the functions
$$\tilde{f}_n := f_n \circ \phi \in H^\infty, \quad n \geq 1$$
are bounded in modulus by 1 in $\mathbb{D}$. By Carathéodory's classical theorem (see, for instance, [8, Chapter I.3]), ϕ extends to a homeomorphism between $\partial\mathbb{D}$ and $\partial\Omega$. If $\|K_\Omega f_n + \mathbb{C}\mathbf{1}\|_\Omega \to 1$, then Lemmas 16 and 17 show that there exist two sets $E, E' \subset \partial\Omega$ which have positive arclength measure, such that
$$\lim_{n\to\infty} \tilde{f}_n(\lambda) = 1, \quad \lambda \in \phi^{-1}(E)$$
and
$$\lim_{n\to\infty} \tilde{f}_n(\lambda) = \alpha, \quad \lambda \in \phi^{-1}(E').$$
Since Ω is convex, the curve $\partial\Omega$ is rectifiable, and general theory of harmonic measures tells us that the sets $\phi^{-1}(E)$ and $\phi^{-1}(E')$ have positive Lebesgue measure (see [8, Chapter VI]). Since $L^1(\partial\mathbb{D})$ is separable and the functions $\tilde{f}_n$ are uniformly bounded by 1 in modulus, the usual Helly-type selection process will produce a subsequence of $(\tilde{f}_n)$ which converges in the weak-star topology to some function $\widetilde{f} \in H^\infty$. By the above pointwise convergence, we must have $\widetilde{f} \equiv 1$ on $\phi^{-1}(E)$ and $\tilde{f} \equiv \alpha$ on $\phi^{-1}(E')$. Then the non-zero function $\widetilde{f} - 1$ vanishes on the subset $\phi^{-1}(E)$ of positive Lebesgue measure on $\partial\mathbb{D}$. This is a contradiction, which shows that our assumption $\|K_\Omega f_n + \mathbb{C}\mathbf{1}\|_\Omega \to 1$ must be false. Theorem 2 follows. $\square$

4.3. **A proof of Neumann's lemma.** We indicate how one may proceed to use our above arguments to obtain a proof of Neumann's lemma, stating that $c(\Omega) = 1$ if and only if Ω is a triangle or a quadrilateral. We need only the following simple geometric observation regarding the densities ρ_ζ.

Lemma 18. *Fix $\zeta \in \partial\Omega$. Any $\sigma \in \partial\Omega \setminus \{\zeta\}$ which is not a corner of $\partial\Omega$ and which satisfies $\rho_\zeta(\sigma) = 0$ is contained in the union of at most two line segments of $\partial\Omega$ containing ζ.*

Proof. It will suffice to show that all σ satisfying the above conditions are contained in at most two different tangent lines to Ω. To see this, recall formula (5). The condition $\rho_\zeta(\sigma) = (2\pi R_{\zeta,\sigma})^{-1} = 0$ gives $R_{\zeta,\sigma} = \infty$, and so ζ is contained in the tangent line to Ω at σ. The tangent line divides the plane $\mathbb{C}$ into two half-planes, one of which contains Ω. Assume that two different tangent lines, at σ and σ', intersect at ζ. They divide the plane $\mathbb{C}$ into four sectors, and by convexity precisely one

of those sectors contains Ω. Now, any line which passes through ζ and the open sector containing Ω must separate $\sigma, \sigma' \in \partial\Omega$. Therefore, it is not a tangent to Ω. $\qquad\square$

Neumann's lemma is established as follows. Assume that $c(\Omega) = 1$. From Lemmas 16 and 17 we see that two points ζ, ζ' exist for which the measures $\rho_\zeta ds$ and $\rho_{\zeta'} ds$ are mutually singular. From Lemma 18 we deduce that the support of $\rho_\zeta ds$ is the union of at most two line segments containing ζ', and the complement of the support of $\rho_\zeta ds$ is also a union of at most two line segments. Thus $\partial\Omega$ is the union of at most four line segments.

5. Examples

In this section, we compute and estimate the configuration constants for some types of domains.

5.1. Configuration constant of an ellipse. For $a, b > 0$, let

$$\Omega_{a,b} := \left\{ x + iy \in \mathbb{C} : \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1 \right\}$$

be the ellipse centred at the origin with semi-axes of lengths a and b, respectively. It is quite remarkable that the configuration constant can in this case be computed explicitly.

Proposition 19. *With the above notation, we have*

$$c(\Omega_{a,b}) = \frac{2}{\pi} \arctan\left(\frac{1}{2} \left| \frac{b}{a} - \frac{a}{b} \right| \right).$$

In order to prove the proposition, our first step is to derive an expression for the density of the Neumann-Poincaré kernel of $\Omega_{a,b}$. The boundary $\partial\Omega_{a,b}$ is parametrized by

(37)
$$\gamma(t) := a \cos t + ib \sin t, \quad t \in [0, 2\pi].$$

Here $[0, 2\pi]$ can be replaced by any interval of length 2π. Recalling formula (4) for μ_ζ and setting $\zeta = \gamma(s)$, $\sigma = \gamma(t)$, we obtain

(38)
$$d\mu_{\gamma(s)}(\gamma(t)) = \rho_{\gamma(s)}(\gamma(t)) \, ds(\gamma(t))$$

$$= \frac{1}{\pi} \operatorname{Im}\left(\frac{T(\gamma(t))}{\gamma(t) - \gamma(s)} \right) |\gamma'(t)| \, dt$$

$$= \frac{1}{\pi} \operatorname{Im} \frac{\gamma'(t)}{\gamma(t) - \gamma(s)} \, dt.$$

Using (37), this formula can be greatly simplified.

Lemma 20. *With the notation above, we have*

(39)
$$d\mu_{\gamma(s)}(\gamma(t)) = \frac{1}{2\pi} \frac{A}{1 + B \cos(t + s)} \, dt, \quad s, t \in [0, 2\pi],$$

where

$$A := \frac{2ab}{a^2 + b^2} \quad and \quad B := \frac{b^2 - a^2}{b^2 + a^2}.$$

The lemma is established by combining (37) and (38), and then using elementary trigonometric identities to simplify the resulting expression.

With this formula in hand, we now evaluate the configuration constant of the ellipse $\Omega_{a,b}$.

Proof of Proposition 19. Using the formulas (8) and (39), we obtain

$$c(\Omega_{a,b}) = \sup_{s_1,s_2 \in [0,2\pi]} \frac{1}{2} \frac{1}{2\pi} \int_{[-\pi,\pi]} \left| \frac{A}{1 + B\cos(t+s_1)} - \frac{A}{1 + B\cos(t+s_2)} \right| dt.$$

By the periodicity of cos, this last expression simplifies to

$$c(\Omega_{a,b}) = \sup_{s \in (0,2\pi)} \frac{1}{2} \frac{1}{2\pi} \int_{[-\pi,\pi]} \left| \frac{A}{1 + B\cos(t+s)} - \frac{A}{1 + B\cos(t)} \right| dt.$$

For the time being, let us assume that $b \geq a$, so $B \geq 0$. Using the fact that (39) is the density of a probability measure for each $s \in [0, 2\pi]$, we have

$$\frac{1}{2} \frac{1}{2\pi} \int_{[-\pi,\pi]} \left| \frac{A}{1 + B\cos(t+s)} - \frac{A}{1 + B\cos(t)} \right| dt$$
$$= \frac{1}{2\pi} \int_{\{t : \cos t \geq \cos(t+s)\}} \left(\frac{A}{1 + B\cos(t+s)} - \frac{A}{1 + B\cos(t)} \right) dt.$$

We readily verify that $\cos(t) \geq \cos(t+s)$ if and only if $t \in [-s/2, \pi - s/2]$. Therefore

$$\frac{1}{2} \frac{1}{2\pi} \int_{[-\pi,\pi]} \left| \frac{A}{1 + B\cos(t+s)} - \frac{A}{1 + B\cos(t)} \right| dt$$
$$= \frac{1}{2\pi} \int_{-s/2}^{\pi-s/2} \left(\frac{A}{1 + B\cos(t+s)} - \frac{A}{1 + B\cos(t)} \right) dt$$
$$= \frac{1}{2\pi} \int_{s/2}^{\pi+s/2} \frac{A}{1 + B\cos(t)} dt - \frac{1}{2\pi} \int_{-s/2}^{\pi-s/2} \frac{A}{1 + B\cos(t)} dt$$
$$= \frac{1}{2\pi} \int_{\pi-s/2}^{\pi+s/2} \frac{A}{1 + B\cos(t)} dt - \frac{1}{2\pi} \int_{-s/2}^{s/2} \frac{A}{1 + B\cos(t)} dt$$
$$= \frac{1}{2\pi} \int_{-s/2}^{s/2} \frac{A}{1 - B\cos(t)} dt - \frac{1}{2\pi} \int_{-s/2}^{s/2} \frac{A}{1 + B\cos(t)} dt$$
$$= \frac{1}{2\pi} \int_{-s/2}^{s/2} \frac{2AB\cos(t)}{1 - B^2\cos^2(t)} dt.$$

It is clear that this last integral is maximized over $s \in [0, 2\pi]$ when $s = \pi$. Putting everything together, we deduce that, if $b \geq a$, then

$$c(\Omega_{a,b}) = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \frac{2AB\cos(t)}{1 - B^2\cos^2(t)} dt.$$

All that remains is to evaluate the integral. Making the substitution $x = \sin t$, and exploiting the fact that $A^2 + B^2 = 1$, we have

$$\frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \frac{2AB\cos(t)}{1 - B^2\cos^2(t)} dt = \frac{1}{2\pi} \int_{-1}^{1} \frac{2AB}{1 - B^2(1 - x^2)} dx$$
$$= \frac{1}{\pi} \int_{-1}^{1} \frac{AB}{A^2 + B^2 x^2} dx$$
$$= \frac{2}{\pi} \arctan\left(\frac{B}{A} \right)$$
$$= \frac{2}{\pi} \arctan\left(\frac{1}{2}\left(\frac{b}{a} - \frac{a}{b} \right) \right).$$
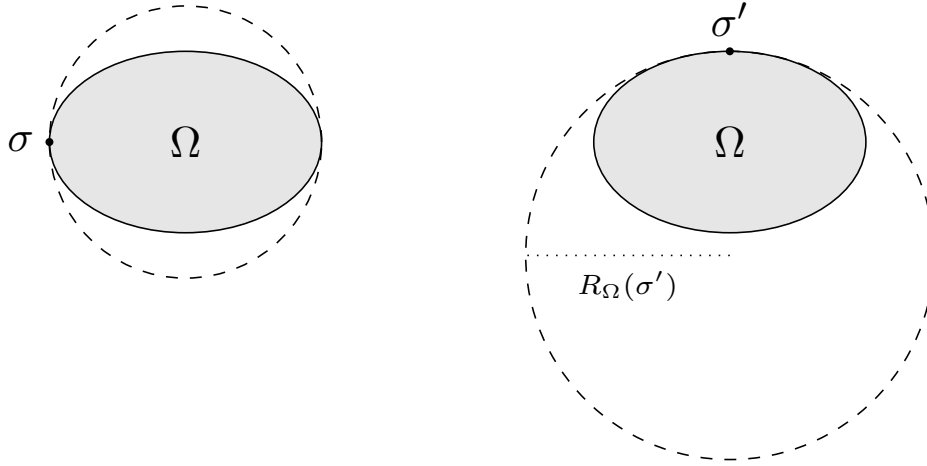
FIGURE 6. A domain Ω with two circles corresponding to values $R_\Omega(\sigma)$ and $R_\Omega(\sigma')$

This proves the result in the case when $b \geq a$. The remaining case is obtained by exchanging the roles of a and b. □

5.2. **Integral estimates.** For a general domain, the exact value of $c(\Omega)$ is often inaccessible. In this section, we will present a simple estimate which is applicable to domains Ω with a non-flat part of the boundary which leads to an upper bound on $c(\Omega)$.

Assume that we find a Borel measure ν on $\partial\Omega$ such that

$$k_\Omega^\nu := \sup\{\|\mu_\zeta - \nu\| : \zeta \in \partial\Omega\} < 1.$$

If so, then, for every $\phi \in C(\partial\Omega)$ with $\|\phi\|_{\partial\Omega} \leq 1$, we have

$$\left| K_\Omega \phi(\zeta) - \int_{\partial\Omega} \phi \, d\nu \right| \leq k_\Omega^\nu, \quad \zeta \in \partial\Omega,$$

which shows that the image of $K_\Omega\phi$ is contained in a disk of radius k_Ω^ν centred at $\int_{\partial\Omega} \phi \, d\nu$. Thus,

$$c_\mathbb{R}(\Omega) = c_\mathbb{C}(\Omega) = \|K_\Omega : C(\partial\Omega)/\mathbb{C}\mathbf{1} \to C(\partial\Omega)/\mathbb{C}\mathbf{1}\| \leq k_\Omega^\nu.$$

One approach is to seek a positive measure ν on $\partial\Omega$ satisfying $\mu_\zeta \geq \nu$ for all $\zeta \in \partial\Omega$. Then

$$\|\mu_\zeta - \nu\| = (\mu_\zeta - \nu)(\partial\Omega) = 1 - \nu(\partial\Omega),$$

and so $k_\Omega^\nu = 1 - \nu(\partial\Omega)$.

We will construct the largest non-negative Borel measure ν on $\partial\Omega$ which satisfies $\mu_\zeta - \nu \geq 0$. The construction is based on the geometric interpretation of the density $\rho_\zeta(\sigma)$ and has already been mentioned in (15). In order to avoid the need to establish Borel measurability of R_Ω defined as a supremum of an uncountable family as in (15), we proceed to define it in a slightly different but equivalent way. Namely, it is easy to see that, given $\sigma \in \partial\Omega$, if there exists a closed disk Δ such that $\Omega \subset \Delta$ and $\sigma \in \partial\Delta$, then there exists one of smallest radius. We denote this radius by $R_\Omega(\sigma)$. Note that if σ is not a corner of $\partial\Omega$, then the corresponding disk must be tangent to $\partial\Omega$ at σ. If no disk passing through σ exists which contains Ω, then we set $R_\Omega(\sigma) := \infty$. This happens, for instance, if σ is contained in the interior of a line segment in $\partial\Omega$. In particular, $R_\Omega(\sigma) = \infty$ for all but a finite number of points of any polygonal domain.

Lemma 21. *The function $R_\Omega : \partial\Omega \to (0, \infty]$ is lower semicontinuous. In particular, it is Borel measurable.*

Proof. Let $\sigma \in \partial\Omega$ and let (σ_n) be a sequence in $\partial\Omega$ such that $\sigma_n \to \sigma$. We need to show that $\liminf_{n\to\infty} R_\Omega(\sigma_n) \geq R_\Omega(\sigma)$. We can suppose that $L := \liminf_{n\to\infty} R_\Omega(\sigma_n) < \infty$, otherwise there is nothing to prove. Let $L' > L$. Then, replacing (σ_n) by a subsequence, we can suppose that $R_\Omega(\sigma_n) < L'$ for all n. Thus, for each n, there exists a closed disk Δ_n of radius L' such that $\Omega \subset \Delta_n$ and $\sigma_n \in \partial\Delta_n$. The sequence of centres (c_n) of the disks Δ_n is bounded, so there exists a convergent subsequence $c_{n_j} \to c$. Let Δ be the closed disk with centre c and radius L'. Then we have $\Omega \subset \Delta$ and $\sigma \in \partial\Delta$. It follows that $R_\Omega(\sigma) \leq L'$. As this last inequality holds for all $L' > L$, we deduce that $R_\Omega(\sigma) \leq L$. This completes the proof. $\square$

We set

(40)
$$d\nu := \frac{ds}{2\pi R_\Omega}.$$

By the above lemma, ν is a non-negative Borel measure on $\partial\Omega$. For any $\zeta \in \partial\Omega$, we have

(41)
$$\mu_\zeta \geq \nu.$$

To see this, note that if σ is not a corner and $R_{\zeta,\sigma}$ is the radius of the unique circle tangent to $\partial\Omega$ at σ and passing through ζ, then $R_\Omega(\sigma) \geq R_{\zeta,\sigma}$. Therefore, according to (5),

$$\frac{1}{2\pi R_\Omega(\sigma)} \leq \frac{1}{2\pi R_{\zeta,\sigma}} = \rho_\zeta(\sigma)$$

for almost every σ with respect to arclength measure on $\partial\Omega$. Inequality (41) follows. Although we shall skip a formal proof, we mention also that ν is in fact the largest measure satisfying $\mu_\zeta \geq \nu$ for all $\zeta \in \partial\Omega$. This maximality property of ν is to be interpreted in the following sense: if ν' is any measure satisfying $\mu_\zeta \geq \nu'$ for all ζ, then $\nu \geq \nu'$.

By our earlier discussion, we obtain the following upper estimate for the configuration constant:

$$c(\Omega) \leq 1 - \frac{1}{2\pi} \int_{\partial\Omega} \frac{ds}{R_\Omega}.$$

Note that this is precisely the assertion of Theorem 4 stated in Section 1.

We will now mention some consequences. Recall that if γ is a plane curve of class C^2, then the *radius of curvature* of γ is the reciprocal of its curvature.

Corollary 22. *If Ω has a C^2-boundary of length L, whose radius of curvature is everywhere at most ρ, then*

$$c(\Omega) \leq 1 - \frac{L}{2\pi\rho}.$$

Proof. In this case, one sees from (15) and (16) that $R_\Omega(\sigma) \leq \rho$ for all $\sigma \in \partial\Omega$, from which the result follows. $\square$

This last result was already known. See for example [6, pp.45–46] and [11, pp.128–129]. However the proofs in these references are quite different from the one above.

Corollary 23. *Consider a convex circular sector*

$$\Omega = \{z \in \mathbb{C} : 0 \leq |z| \leq r, \, 0 \leq \arg(z) \leq \theta\},$$

where $r > 0$ and $0 < \theta \leq \pi$. Then

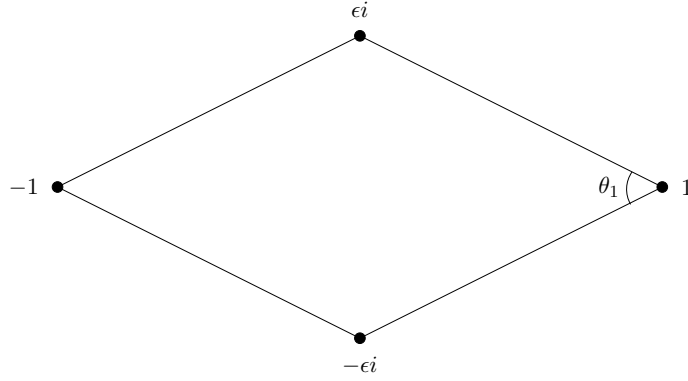$$c(\Omega) \leq 1 - \frac{\theta}{2\pi}.$$

FIGURE 7. A quadrilateral domain Ω_ϵ with $a(\Omega_\epsilon) > 1 - (4/\pi)\epsilon$.

Proof. It is obvious that $R_\Omega(\sigma) = r$ for σ in the curved part of $\partial\Omega$, and that $R_\Omega(\sigma) = \infty$ elsewhere. Hence

$$\frac{1}{2\pi}\int_{\partial\Omega}\frac{ds}{R_\Omega} = \frac{1}{2\pi}\frac{r\theta}{r} = \frac{\theta}{2\pi}.$$

The result now follows from Theorem 4. $\square$

5.3. **Analytic configuration constants of quadrilaterals.** Theorem 2 shows that $a(\Omega) < 1$ for every Ω. Here we show by example that $a(\Omega)$ may be arbitrarily close to 1. Figure 7 shows a narrow quadrilateral domain for which this phenomenon occurs.

Proposition 24. *For $\epsilon > 0$, let Ω_ϵ be the convex hull of $\{\pm 1, \pm\epsilon i\}$. Then*

$$a(\Omega_\epsilon) \geq 1 - (4/\pi)\epsilon.$$

Proof. Let f be a conformal mapping of the interior of Ω_ϵ onto the unit disk $\mathbb{D}$. By Carathéodory's theorem, f extends to a homeomorphism of Ω_ϵ onto $\overline{\mathbb{D}}$, and so clearly $f \in A(\Omega)$. Post-composing with a suitable automorphism of $\overline{\mathbb{D}}$, we may further suppose that $f(1) = 1$ and $f(-1) = -1$.

Consider $\zeta = 1$. Recalling (3), we have $\mu_1 = (1 - \theta_1/\pi)\delta_1 + (\theta_1/\pi)\nu$, where θ_1 is the angle of the aperture of $\partial\Omega_\epsilon$ at 1, and ν is a probability measure on $\partial\Omega_\epsilon \setminus \{1\}$. It follows that

$$\mathrm{Re}(K_{\Omega_\epsilon}f)(1) = \int_{\partial\Omega_\epsilon}(\mathrm{Re}\,f)\,d\mu_1$$
$$= (1 - \theta_1/\pi)\,\mathrm{Re}\,f(1) + (\theta_1/\pi)\int_{\partial\Omega_\epsilon\setminus\{1\}}(\mathrm{Re}\,f)\,d\nu$$
$$\geq (1 - \theta_1/\pi)(1) + (\theta_1/\pi)(-1)$$
$$= 1 - 2\theta_1/\pi.$$

Likewise

$$\mathrm{Re}(K_{\Omega_\epsilon}f)(-1) \leq -(1 - 2\theta_1/\pi).$$

It follows that the diameter of $(K_{\Omega_\epsilon}f)(\Omega)$ is at least $2(1 - 2\theta_1/\pi)$, whence $a(\Omega) \geq (1 - 2\theta_1/\pi)$.

Finally, by trigonometry, θ_1 is related to ϵ by $\tan(\theta_1/2) = \epsilon$, whence $\theta_1 = 2\arctan\epsilon \leq 2\epsilon$. The result follows. $\square$

5.4. **Configuration constants equal to zero.** Recall the estimate (13) from Section 1, which will be proved in the next section. The estimate is strongest if $a(W) = 0$, and in this case we reach the conjectured bound $K = 2$. Unfortunately, the only domain W for which we have $a(W) = 0$ is a disk, and in this case (13) reduces to the well-known Okubo-Ando bound from [14]. For completeness we give a proof of the statement that $a(\Omega) = 0$ if and only if Ω is a disk. More precisely, we have the following.

Proposition 25. *Let Ω be a compact convex domain with non-empty interior. The following are equivalent:*

  (i) Ω is a disk,
  (ii) $c(\Omega) = 0$,
 (iii) $a(\Omega) = 0$.

Proof. In the case that Ω is a disk, then a short computation shows directly that $\rho_\zeta(\sigma)$ in (4) is a constant independent of ζ, and so for every $\zeta \in \partial\Omega$, the measure μ_ζ is a normalized arclength measure on the circular boundary $\partial\Omega$. Then it follows from the definition that $K_\Omega f$ is a constant function, and consequently $\|K_\Omega f + \mathbb{C}\mathbf{1}\|_{C(\partial\Omega)} = 0$, so $c(\Omega) = a(\Omega) = 0$. This shows that the implications $(i) \Rightarrow (ii)$ and $(ii) \Rightarrow (iii)$ hold.

It remains to prove $(iii) \Rightarrow (i)$. Fix a conformal mapping $\phi : \mathbb{D} \to \Omega^o$, where $\mathbb{D}$ is the open unit disk and Ω^o is the interior of Ω. The mapping ϕ extends to a homeomorphism of $\partial\mathbb{D}$ and $\partial\Omega$, and so it makes sense to define the probability measures μ_ζ^ϕ on $\partial\mathbb{D}$ by the equation

$$\mu_\zeta^\phi(E) := \mu_\zeta(\phi(E)),$$

where E is a Borel subset of $\partial\mathbb{D}$, and $\{\mu_\zeta\}_{\zeta \in \partial\Omega}$ is the Neumann-Poincaré kernel of Ω. Since $a(\Omega) = 0$, it follows that for every $f \in \mathcal{A}(\Omega)$ and every pair of points $\zeta, \zeta' \in \partial\Omega$ we have, by the change of variables formula, that

$$0 = \int_{\partial\Omega} f \, d\mu_\zeta - \int_{\partial\Omega} f \, d\mu_{\zeta'}$$
$$= \int_{\partial\mathbb{D}} f \circ \phi \, d\mu_\zeta^\phi - \int_{\partial\mathbb{D}} f \circ \phi \, d\mu_{\zeta'}^\phi.$$

As f varies over $\mathcal{A}(\Omega)$, $f \circ \phi$ varies over $\mathcal{A}(\mathbb{D}) := \mathcal{A}(\overline{\mathbb{D}})$, and it follows that $\mu_\zeta^\phi - \mu_{\zeta'}^\phi$ annihilates $\mathcal{A}(\mathbb{D})$. Then the theorem of brothers Riesz (see, for instance, [7, Exercise 1, Chapter III]) implies that

$$\mu_\zeta^\phi - \mu_{\zeta'}^\phi = h \, ds_{\partial\mathbb{D}}$$

where h is a function with vanishing positive Fourier coefficients. Note that h is real-valued, so the negative Fourier coefficients also vanish, and consequently $h \equiv 0$. Since ζ, ζ' were arbitrary, we conclude that the hypothesis $a(\Omega) = 0$ implies that all the measures μ_ζ are equal.

The conclusion that Ω is a disk is now a consequence of the geometric formula for $\rho_\zeta(\sigma)$ in (4). For any three different $\zeta_1, \zeta_2, \zeta_3 \in \partial\Omega$ we have that $\rho_{\zeta_1}(\sigma) = \rho_{\zeta_2}(\sigma) = \rho_{\zeta_3}(\sigma)$ for any σ which is not a corner of $\partial\Omega$. Then by the geometric interpretation in (4), the unique circle C passing through $\zeta_1, \zeta_2, \zeta_3$ is tangent to $\partial\Omega$ at σ. Hence $\partial\Omega \subset C$, and we conclude that Ω is a disk. □

6. Application to numerical ranges

6.1. **Spectral constant estimate.** Our principal motivation for the introduction of the analytic configuration constant is the following result which was mentioned in the Section 1 and which we will now prove.

Proposition 26. *Let T be a bounded linear operator on a Hilbert space $\mathcal{H}$, and $W = \overline{W(T)}$ the closure of the numerical range of T. If W has non-empty interior, then for every $f \in \mathcal{A}(W)$ we have*

$$\|f(T)\| \leq \left(1 + \sqrt{1 + a(W)}\right)\|f\|_W,$$

where $a(W)$ is the analytic configuration constant in (11), *and $\mathcal{A}(W)$ is the space of continuous functions on W which are analytic in the interior of W.*

Of course, if W has no interior, then its convexity forces it to be a line segment. In that case T is a normal operator, and the spectral theorem gives us the better estimate $\|f(T)\| \leq \|f\|_{\sigma(T)}$, where f may be any Borel measurable function on the spectrum $\sigma(T)$. In what follows, we will assume that W has non-empty interior.

Let us make some initial remarks before going into the proof of Proposition 26. In the case $\sigma(T)$ is contained in the interior of W, then $f(T)$ is defined, as usual, through the Dunford-Riesz holomorphic functional calculus. If $\partial W \cap \sigma(T) \neq \emptyset$, then this definition does not work. Nevertheless, if $f \in \mathcal{A}(W)$, then it is a standard result of approximation theory that a sequence of analytic polynomials (p_n) exists which converges to f uniformly on W. In the presence of *any* uniform bound of the form $\|p(T)\| \leq K\|p(T)\|_W$ for polynomials p, we may then define $f(T)$ as the limit of the sequence $(p_n(T))$ in the operator norm. Such bounds are known to exists, the strongest known bound $K \leq 1 + \sqrt{2}$ being due to Crouzeix and Palencia. Proposition 26 improves this estimate given information about the numerical range of T.

Our proof of Proposition 26 combines the argument of Crouzeix and Palencia from [4] with ideas of Schwenninger and de Vries from [17], where bounds for various functional calculi are derived as a consequence of the existence of *extremal functions* and *extremal vectors*. Let U be an open set in the plane, and $H^\infty(U)$ be the algebra of holomorphic functions on U. Given an operator $T : \mathcal{H} \to \mathcal{H}$ with $\sigma(T)$ contained in U, it is elementary that the quantity

$$\sup\left\{\|f(T)\| : f \in H^\infty(U), \|f\|_U \leq 1\right\}$$

is finite. A normal-families argument shows that an $f \in H^\infty(U)$ exists with $\|f\|_U = 1$ for which the supremum above is attained. Any such f will be called for an *extremal function*. If, moreover, a vector $x \in \mathcal{H}$ with $\|x\|_{\mathcal{H}} = 1$ exists for which

$$\sup\left\{\|f(T)\| : f \in H^\infty(U), \|f\|_U \leq 1\right\} = \|f(T)x\|_{\mathcal{H}}$$

then we will say that x is an *extremal vector*, and (f, x) is an *extremal pair*. Unless $\dim \mathcal{H} < \infty$, an extremal vector may not exist, but we will be able to reduce the proof to the finite-dimensional case. The importance of the concept of extremal pairs (f, x) stems from the following result. We refer the reader to [1, Theorem 4.5] for a proof (see also [17, Proposition 3]).

Lemma 27. *Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator, and U be an open neighbourhood of $\sigma(T)$. Let (f, x) be a corresponding extremal pair. If $\|f(T)\| > 1$, then $f(T)x$ is orthogonal to x in $\mathcal{H}$:*

$$\langle f(T)x, x\rangle_{\mathcal{H}} = 0.$$

The next two lemmas will reduce our task to consideration of finite-dimensional Hilbert spaces, in which extremal vectors exist, and will dispose of the problematic set $\sigma(T) \cap \partial W$. The first observation is essentially contained in [17, Proposition 9].

Lemma 28. *If for some $K > 0$ the estimate*

$$\|p(T)\| \leq K\|p\|_\Omega$$

holds for every polynomial p and every operator T on a finite-dimensional Hilbert space, then the same estimate with the same constant K holds also for operators T on infinite-dimensional Hilbert spaces.

Proof. Let $T : \mathcal{H} \to \mathcal{H}$ be as above, with $\dim \mathcal{H} = \infty$. It suffices to show that

$$\|p(T)x\|_{\mathcal{H}} \leq K\|p\|_{\Omega}\|x\|_{\mathcal{H}}$$

holds for every analytic polynomial p and every $x \in \mathcal{H}$. Note that $p(T)x$ is contained in the finite-dimensional subspace $\mathcal{K}$ spanned by $\{x, Tx, \ldots, T^d x\}$, where d is the degree of the polynomial p. If $\Pi : \mathcal{H} \to \mathcal{K}$ is the orthogonal projection, then $p(T)x = \Pi p(T)x = p(\Pi T)x$, where ΠT can be considered as an operator on the finite-dimensional Hilbert space $\mathcal{K}$. Our hypothesis implies that

$$\|p(T)x\|_{\mathcal{H}} = \|p(\Pi T)x\|_{\mathcal{K}} \leq K\|p\|_{\Omega}\|x\|_{\mathcal{K}} = K\|p\|_{\Omega}\|x\|_{\mathcal{H}}.$$

The lemma follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The proof of the next lemma will use affine invariance of the configuration constants. Let us fix $\alpha, \beta \in \mathbb{C}$, $\alpha \neq 0$, and an affine mapping $A(z) := \alpha z + \beta$. Then A is a conformal transformation of $\mathbb{C}$ with the additional property of taking a line segment of length L to a line segment of length $|\alpha|L$, and a circle of radius R to a circle of radius $|\alpha|R$. Let $\widetilde{\Omega} = A(\Omega)$ be the affine image of Ω under A, and recall the formula for the Neumann-Poincaré kernel in (3) and its geometric interpretation. If $\zeta, \sigma \in \partial\Omega$, E is a Borel subset of $\partial\Omega$, and $s, \widetilde{s}$ are the arclength measures on $\partial\Omega$ and $\partial\widetilde{\Omega}$ respectively, then it follows from the properties of A listed above that

  (i) $\theta_\zeta = \theta_{A(\zeta)}$,
  (ii) $|\alpha|s(E) = \widetilde{s}(A(E))$,
  (iii) $|\alpha|R_{\zeta,\sigma} = R_{A(\zeta),A(\sigma)}$.

A consequence is that the Neumann-Poincaré kernels $\{\mu_\zeta\}_{\zeta \in \partial\Omega}$ and $\{\widetilde{\mu}_{A(\zeta)}\}_{A(\zeta) \in \partial\Omega}$ of the respective domains satisfy

$$\widetilde{\mu}_{A(\zeta)}(A(E)) = \mu_\zeta(E), \quad E \text{ a Borel subset of } \partial\Omega.$$

Then a change of variables shows that $K_\Omega(\widetilde{f} \circ A) = K_{\widetilde{\Omega}}\widetilde{f}$ for any $\widetilde{f} \in C(\partial\widetilde{\Omega})$, and it follows that

$$a(\Omega) = a(\widetilde{\Omega}), \quad c(\Omega) = c(\widetilde{\Omega}).$$

Armed with these equalities, we make our second observation.

Lemma 29. *Assume that the estimate*

$$(42) \qquad\qquad\qquad \|p(T)\| \leq \left(1 + \sqrt{1 + a(\Omega)}\right)\|p\|_{\Omega}$$

holds for every polynomial, every compact convex domain Ω, and every operator T for which $W(T)$ is contained in the interior of Ω. Then Proposition 26 holds.

Proof. Replacing T by an operator $T + \beta I$ for some $\beta \in \mathbb{C}$, we may assume that 0 lies in the interior of $W(T)$. Let $W = \overline{W(T)}$, and

$$W_r = \{rz : z \in W\}, \quad r > 1.$$

Then W_r is a convex domain which contains W in its interior. By our assumption, for any analytic polynomial p we have

$$\|p(T)\| \leq \left(1 + \sqrt{1 + a(W_r)}\right)\|p\|_{W_r}.$$

Since W_r is an affine image of W, we have $a(W_r) = a(W)$. Since this holds for all $r > 1$, and since $\lim_{r \to 1} \|p\|_{W_r} = \|p\|_W$, we may let $r \to 1$ to obtain the desired estimate whenever p is an analytic polynomial. The estimate for $f \in \mathcal{A}(W)$ follows by density of polynomials in $\mathcal{A}(W)$. $\qquad\square$

Proof of Proposition 26. By Lemma 29, it will be sufficient to establish the estimate (42) whenever Ω contains $W(T)$ in its interior Ω^o. Moreover, by Lemma 28, we may assume that T is an operator on a finite-dimensional Hilbert space $\mathcal{H}$. Let $U = \Omega^o$ and (f, x) be an extremal pair corresponding to U and the operator T. If $\|f(T)\| \leq 1$, then (42) certainly holds, so we may assume that $\|f(T)\| > 1$.

Let (f_n) be a sequence in $A(\Omega)$ such that $\|f_n\|_\Omega \leq 1$ and $f_n \to f$ locally uniformly in Ω. Then $f_n(T) \to f(T)$ in operator norm. Set $g_n := K_\Omega \overline{f}_n$. It is shown in [4, Lemmas 2.1 and 2.3] that $g_n \in \mathcal{A}(\Omega)$ and

$$(43) \qquad \|f_n(T) + g_n(T)^*\| \leq 2.$$

For each n, we may choose $\lambda_n \in \mathbb{C}$ such that

$$\|g_n + \lambda_n \mathbf{1}\|_\Omega = \inf_{\lambda \in \mathbb{C}} \|g_n + \lambda \mathbf{1}\|_\Omega \leq a(\Omega).$$

We now have the following identity:

$$(44) \qquad \begin{aligned} \langle f_n(T)x, \, f_n(T)x \rangle_\mathcal{H} = \; &\langle f_n(T)x, (f_n(T) + g_n(T)^*)x \rangle_\mathcal{H} \\ &- \langle f_n(T)x, (g_n + \lambda_n \mathbf{1})(T)^* x \rangle_\mathcal{H} \\ &+ \lambda_n \langle f_n(T)x, x \rangle_\mathcal{H}. \end{aligned}$$

Let us consider each of the terms in this identity. By the choice of x, we have

$$\langle f_n(T)x, f_n(T)x \rangle_\mathcal{H} = \|f_n(T)x\|^2 \xrightarrow[n\to\infty]{} \|f(T)x\|^2 = \|f(T)\|^2.$$

Also, from (43) and the Cauchy–Schwarz inequality,

$$\left| \langle f_n(T)x, (f_n(T) + g_n(T)^*)x \rangle_\mathcal{H} \right| \leq 2\|f_n(T)\| \xrightarrow[n\to\infty]{} 2\|f(T)\|.$$

By Lemma 27, we have

$$\begin{aligned} \left| \langle f_n(T)x, (g_n + \lambda_n \mathbf{1})(T)^* x \rangle \right| &= \left| \langle (f_n(g_n + \lambda_n \mathbf{1}))(T)x, x \rangle \right| \\ &\leq \|f_n(g_n + \lambda_n \mathbf{1})\|_\Omega \\ &\leq \|g_n + \lambda_n \mathbf{1}\|_\Omega \\ &\leq a(\Omega). \end{aligned}$$

By Lemma 27 again, $\langle f(T)x, x \rangle_\mathcal{H} = 0$. Since the sequence (λ_n) is certainly bounded (indeed $|\lambda_n| \leq 2$), we deduce that

$$\lambda_n \langle f_n(T)x, x \rangle_\mathcal{H} \xrightarrow[n\to\infty]{} 0.$$

Thus, letting $n \to \infty$ in (44), we deduce that

$$\|f(T)\|^2 \leq 2\|f(T)\| + a(\Omega).$$

Hence

$$\|f(T)\| \leq 1 + \sqrt{1 + a(\Omega)}.$$

In particular, for every polynomial p with $\|p\|_\Omega = 1$ we have

$$\|p(T)\| \leq \|f(T)\| \leq 1 + \sqrt{1 + a(\Omega)},$$

since f is extremal. This is equivalent to (42), and so the proof is complete. $\qquad \square$

Appendix A. Double-layer potential on a general convex domain

A.1. **Convex domains.** Let Ω be a compact convex domain in the plane $\mathbb{C}$ with non-empty interior Ω^o. We will be making no assumptions regarding smoothness of the boundary $\partial\Omega$. However, convexity itself implies that $\partial\Omega$ is a rectifiable simple closed curve with some additional properties.

The orientation of $\partial\Omega$ is to be counter-clockwise (that is, positive), and we use $\sigma' \uparrow \sigma$ and $\sigma' \downarrow \sigma$ to denote, respectively, the counter-clockwise and clockwise one-sided convergence of σ' to σ within $\partial\Omega$. As a consequence of convexity of Ω, the one-sided tangent angles exist at every point $\sigma \in \partial\Omega$, are locally given by

$$\alpha_+(\sigma) := \lim_{\sigma' \downarrow \sigma} \arg(\sigma' - \sigma), \quad \alpha_-(\sigma) := \lim_{\sigma' \uparrow \sigma} \arg(\sigma - \sigma'),$$

and satisfy

$$\alpha_-(\sigma) \leq \alpha_+(\sigma).$$

Strict inequality may occur at most at a countable subset of $\partial\Omega$. If it occurs at σ, then we say that $\partial\Omega$ has a *corner* at σ. At any point which is not a corner, the tangent angle

$$\alpha(\sigma) := \alpha_+(\sigma) = \alpha_-(\sigma)$$

is well-defined, and so is the tangent $T(\sigma) := e^{i\alpha(\sigma)}$ itself. If $t \mapsto \gamma(t)$ is any (positively-oriented) parametrization of $\partial\Omega$, and we set $\alpha(\sigma) = \alpha_+(\sigma)$ at the corners, then the locally defined function $\alpha(\gamma(t))$ is increasing in t, and consequently the tangent T is continuous at every point which is not a corner of $\partial\Omega$. At a corner, the discontinuity of T amounts to a jump of the argument of T. We denote by $N(\sigma) := -iT(\sigma)$ the outward-pointing normal at $\sigma \in \partial\Omega$.

A.2. **Double-layer potential.** Let Ω^o denote the interior of Ω. To each $z \in \Omega^o$ we associate the measure μ_z on $\partial\Omega$, which for any arc $J \subset \partial\Omega$ satisfies

$$(45) \qquad \mu_z(J) = \frac{1}{\pi} \int_J d\arg(\sigma - z) = \frac{1}{\pi}\big(\text{angle subtended at } z \text{ by } J\big).$$

Here $\arg(\sigma - z)$ is any locally defined continuous determination of the argument function on $\partial\Omega$. Non-negativity of μ_z follows from convexity of Ω and our choice of positive orientation of $\partial\Omega$. With respect to this orientation, every arc $J = (a, b) \subset \partial\Omega$ has a start-point a and an end-point b, and it is easy to see that

$$\mu_z(J) = \frac{\arg(b - z) - \arg(a - z)}{\pi}.$$

In particular, $\mu_z(\partial\Omega) = 2$.

The measure μ_z is absolutely continuous with respect to arclength s on $\partial\Omega$. Indeed, if $\sigma_0 \in \partial\Omega$, $J_n = (a_n, b_n)$ is a sequence of arcs of $\partial\Omega$ which are shrinking to σ_0, and $|J_n|$ are the corresponding arclengths, then

$$\frac{\pi\mu_z(J_n)}{|J_n|} = \frac{1}{|J_n|} \int_{J_n} d\arg(\sigma - z)$$

$$= \frac{\arg(b_n - z) - \arg(a_n - z)}{|J_n|}$$

$$= \operatorname{Im}\left(\frac{\log(b_n - z) - \log(a_n - z)}{b_n - a_n} \cdot \frac{b_n - a_n}{|J_n|}\right).$$

We use above an appropriate locally defined holomorphic branch of the logarithm. As $n \to \infty$, the first factor inside the brackets satisfies

$$\lim_{n \to \infty} \frac{\log(b_n - z) - \log(a_n - z)}{b_n - a_n} = \frac{1}{\sigma_0 - z},$$

while the second factor stays bounded as a consequence of the inequality $|b_n - a_n| \le |J_n|$. Thus

$$\limsup_{n\to\infty} \frac{\mu_z(J_n)}{|J_n|} < \infty$$

and from elementary measure theory we obtain that μ_z is absolutely continuous with respect to s. If moreover σ_0 is not a corner, then it can be shown that

$$\lim_{n\to\infty} \frac{|J_n|}{|b_n - a_n|} = 1,$$

and so in additional to boundedness we even have the convergence

$$\lim_{n\to\infty} \frac{b_n - a_n}{|J_n|} = \lim_{n\to\infty} \frac{b_n - a_n}{|b_n - a_n|} = T(\sigma_0) = iN(\sigma_0).$$

Thus the Radon-Nikodym derivative satisfies

(46) $$\rho_z(\sigma) := \frac{d\mu_z}{ds}(\sigma) = \frac{1}{\pi} \operatorname{Im}\left(\frac{T(\sigma)}{\sigma - z}\right) = \frac{1}{\pi} \operatorname{Re}\left(\frac{N(\sigma)}{\sigma - z}\right)$$

at every $\sigma \in \partial\Omega$ which is not a corner.

A.3. **Boundary kernel.** The Neumann-Poincaré kernel is the boundary version of the family of measures $\{\mu_z\}_{z\in\Omega^o}$ introduced above. To each point $\zeta \in \partial\Omega$ we associate the Borel probability measure on $\partial\Omega$ defined by (45) for arcs $J \subset \partial\Omega$ not containing the point ζ. Because $\zeta \in \partial\Omega$, this definition implies that $\mu_\zeta(\partial\Omega \setminus \{\zeta\}) = \theta_\zeta/\pi$, where $\theta_\zeta = \pi - \alpha_+(\zeta) + \alpha_-(\zeta)$ can be interpreted as the angle of the aperture at ζ. Indeed, θ_ζ is equal to the increase in the argument of $\sigma - \zeta$ as we traverse one loop around $\partial\Omega$ starting and ending at the point ζ, and since μ_ζ is a probability measure, we must have

$$\mu_\zeta(\{\zeta\}) = 1 - \frac{\theta_\zeta}{\pi}.$$

With the exception of this possible point mass, μ_ζ is otherwise absolutely continuous with respect to arclength. The corresponding Radon-Nikodym derivative is given by

(47) $$\rho_\zeta(\sigma) := \frac{d\mu_\zeta}{ds}(\sigma) = \frac{1}{\pi} \operatorname{Im}\left(\frac{T(\sigma)}{\sigma - \zeta}\right) = \frac{1}{\pi} \operatorname{Re}\left(\frac{N(\sigma)}{\sigma - \zeta}\right).$$

The formula (47) is established analogously to (46). All in all, the measure μ_ζ can be decomposed as

$$d\mu_\zeta = (1 - \theta_\zeta/\pi)d\delta_\zeta + \rho_\zeta ds,$$

where δ_ζ is a unit mass at $\zeta \in \partial\Omega$, θ_ζ is the angle of the aperture at ζ (with the convention that $\theta_\zeta = \pi$ if ζ is not a corner), and where the density ρ_ζ is given by (47).

A.4. **Weak-star convergence.** We establish now that

$$\lim_{z\to\zeta} \mu_z = \delta_\zeta + \mu_\zeta$$

in the sense of the weak-star topology on measures. Note that if $B = B(\zeta, \delta)$ is a ball of radius $\delta > 0$ centred at $\zeta \in \partial\Omega$, then expressions (46) and (47) for the densities of μ_z and μ_ζ show that

(48) $$\lim_{z\to\zeta} \int_{\partial\Omega\setminus B} f \, d\mu_z = \int_{\partial\Omega\setminus B} f \, d\mu_\zeta$$

for every $f \in C(\partial\Omega)$. In particular, choosing $f = \mathbf{1}$, we obtain

$$2 = \mu_\zeta(\partial\Omega \setminus B) + \lim_{z\to\zeta} \mu_z(B).$$

Since

$$\lim_{\delta \to 0} \mu_\zeta(\partial\Omega \setminus B) = \mu_\zeta(\partial\Omega \setminus \{\zeta\}) = \theta_\zeta/\pi$$

we see that given $\epsilon > 0$ for all sufficiently small $\delta > 0$ we will have

$$\limsup_{z \to \zeta} |\mu_z(B) - 2 + \theta_\zeta/\pi| \le \epsilon.$$

Returning to general $f \in C(\partial\Omega)$, we have

$$\begin{aligned}
\int_{\partial\Omega} f \, d\mu_z - \int_{\partial\Omega} f \, d[\delta_\zeta + \mu_\zeta] &= \int_{\partial\Omega\setminus B} f \, d\mu_z - \int_{\partial\Omega\setminus B} f \, d\mu_\zeta \\
&+ \int_B \big(f - f(\zeta)\big) \, d\mu_z \\
&+ f(\zeta)\big(\mu_z(B) - 2 + \theta_\zeta/\pi\big) \\
&- \int_{B\setminus\{\zeta\}} f \, d\mu_\zeta.
\end{aligned}$$

On the right-hand side, the first term tends to zero as $z \to \zeta$, the second can be made arbitrarily small by continuity of f, the crude estimate $\mu_z(B) \le 2$ and choice of sufficiently small δ, the third is dominated in modulus by $\|f\|_{\partial\Omega} \cdot \epsilon$ for z sufficiently close to ζ, and the fourth is dominated by $\|f\|_{\partial\Omega} \cdot \mu_\zeta(B \setminus \{\zeta\})$ which also can be made arbitrarily small by choice of sufficiently small δ. The desired weak-star convergence follows.

References

[1] K. Bickel, P. Gorkin, A. Greenbaum, T. Ransford, F. L. Schwenninger, and E. Wegert. Crouzeix's conjecture and related problems. *Comput. Methods Funct. Theory*, 20(3-4):701–728, 2020.

[2] M. Crouzeix. Bounds for analytical functions of matrices. *Integral Equations Oper. Theory*, 48(4):461–477, 2004.

[3] M. Crouzeix. Numerical range and functional calculus in Hilbert space. *J. Funct. Anal.*, 244(2):668–690, 2007.

[4] M. Crouzeix and C. Palencia. The numerical range is a $(1 + \sqrt{2})$-spectral set. *SIAM J. Matrix Anal. Appl.*, 38(2):649–655, 2017.

[5] B. Delyon and F. Delyon. Generalization of von Neumann's spectral sets and integral representation of operators. *Bull. Soc. Math. Fr.*, 127(1):25–41, 1999.

[6] D. Gaier. *Konstruktive Methoden der konformen Abbildung.*, volume 3 of *Springer Tracts Nat. Philos.* Springer-Verlag, New York, NY, 1964.

[7] J. Garnett. *Bounded analytic functions*, volume 236 of *Grad. Texts Math.* New York, NY: Springer, revised 1st ed. edition, 2006.

[8] J. Garnett and D. Marshall. *Harmonic measure*, volume 2 of *New Math. Monogr.* Cambridge: Cambridge University Press, 2005.

[9] K. Gustafson and D. Rao. *Numerical range. The field of values of linear operators and matrices.* Universitext. New York, NY: Springer, 1996.

[10] H. Jung. Über den kleinsten Kreis, der eine ebene Figur einschließt. *J. Reine Angew. Math.*, 137:310–313, 1910.

[11] L. Kantorovich and V. Krylov. Approximate methods of higher analysis. Groningen: P. Noordhoff Ltd. xii, 1958.

[12] J. Kral. *Integral operators in potential theory*, volume 823 of *Lect. Notes Math.* Springer, Cham, 1980.

[13] C. Neumann. Untersuchungen über das logarithmische und Newton'sche Potential. Leipzig. Teubner, 1877.

[14] K. Okubo and T. Ando. Constants related to operators of class C_ρ. *Manuscr. Math.*, 16:385–394, 1975.

[15] M. Putinar and S. Sandberg. A skew normal dilation on the numerical range of an operator. *Math. Ann.*, 331(2):345–357, 2005.

[16] G. Schober. Neumann's Lemma. *Proc. Am. Math. Soc.*, 19:306–311, 1968.

[17] F. Schwenninger and J. de Vries. On abstract spectral constants. In *Operator and Matrix Theory, Function Spaces, and Applications. IWOTA 2022. Operator Theory: Advances and Applications*, volume 295. Birkhäuser, Cham.

Division of Mathematics and Physics, Mälardalen University, Västerås, Sweden
Email address: bartosz.malman@mdu.se

Département de mathématiques et de statistique, Université Laval, Québec (QC), G1V 0A6, Canada
Email address: javad.mashreghi@mat.ulaval.ca

Département de mathématiques et de statistique, Université Laval, Québec (QC), G1V 0A6, Canada
Email address: ryan.oloughlin.1@ulaval.ca

Département de mathématiques et de statistique, Université Laval, Québec (QC), G1V 0A6, Canada
Email address: thomas.ransford@mat.ulaval.ca