

## RANDOM POLYNOMIALS WITH PRESCRIBED NEWTON POLYTOPE

BERNARD SHIFFMAN AND STEVE ZELDITCH

### CONTENTS

|   |     |
|---|-----|
| 1. Introduction   | 50  |
| 1.1. Statement of results                                 | 50  |
| 1.2. Examples   | 58  |
| 1.3. Methods of proof                                     | 63  |
| 1.4. Brief outline and remarks                            | 64  |
| 2. Background   | 66  |
| 2.1. Fans   | 66  |
| 2.2. Szegő kernels  | 67  |
| 3. Polytope character                                     | 69  |
| 3.1. Proof of Proposition 3.1                             | 70  |
| 4. Mass asymptotics: Proof of Theorem 1.2                 | 76  |
| 4.1. The classically allowed region                       | 79  |
| 4.2. The classically forbidden region                     | 80  |
| 5. Distribution of zeros                                  | 94  |
| 5.1. Expected zero current and the Szegő kernel           | 94  |
| 5.2. Asymptotics of zeros: Proofs of Theorems 1.3 and 1.4 | 96  |
| 5.3. Proof of Theorem 1.1                                 | 99  |
| 5.4. Vanishing of $\psi_P$ along the normal flow          | 100 |
| 5.5. Amoebas in the plane                                 | 101 |
| 6. Appendix: Approach using toric geometry                | 103 |
| 6.1. Relationship to toric geometry                       | 103 |
| 6.2. Geometric proof of Theorem 4.1                       | 104 |
| Index of notation   | 106 |
| References  | 106 |

---

Received by the editors March 12, 2002 and, in revised form, May 17, 2003.

2000 *Mathematics Subject Classification*. Primary 12D10, 60D05; Secondary 14Q99, 32H99, 52B20.

*Key words and phrases*. Random polynomial system, distribution of zeros, Newton polytope, Szegő kernel, polytope character, Bernstein-Kouchnirenko Theorem, amoeba, zero current, complex stationary phase.

Research partially supported by NSF grant DMS-0100474 (first author) and DMS-0071358 (second author).

## 1. INTRODUCTION

It is well known that the Newton polytope  $P$  of a polynomial  $f(z_1, \dots, z_m)$  of degree  $p$  has a crucial influence on its value distribution and in particular on its zero set. Even the number of simultaneous zeros in  $(\mathbb{C}^*)^m := (\mathbb{C} \setminus \{0\})^m$  of  $m$  generic polynomials  $f_1, \dots, f_m$  depends on their Newton polytopes [Be, Ko1, Ko2]. Our purpose in this paper is to demonstrate that the Newton polytope of a polynomial  $f$  also has a crucial influence on its *mass density*  $|f(z)|^2 dV$  and on the *spatial distribution of zeros*  $\{f = 0\}$ . We will show that there is a *classically allowed region* where the mass almost surely concentrates and a *classically forbidden region* where it almost surely is exponentially decaying. The classically allowed region is the inverse image  $\mu^{-1}(\frac{1}{p}P^\circ)$  of the (scaled, open) polytope  $\frac{1}{p}P^\circ$  under the standard moment map  $\mu$  of complex projective space  $\mathbb{C}\mathbb{P}^m$ . The simultaneous zeros in  $(\mathbb{C}^*)^m$  of  $m$  generic polynomials  $f_1, \dots, f_m$  with Newton polytope  $P$  tend to concentrate (in the limit of high degrees) in the classically allowed region, giving a kind of quantitative localized version of the Bernstein-Kouchnirenko Theorem [Be, Ko1, Ko2]. Polynomials with a given Newton polytope  $P$  are often called *sparse* in the literature, and methods of algebraic (including toric) geometry have recently been applied to the computational problem of locating zeros of systems of such sparse polynomials (e.g., see [HS, MR, St, Ro, Ve]). Our results give information on the expected location of zeros when the polynomials are given the conditional measure described below. To our knowledge, this kind of precise asymptotic concentration of zeros of sparse systems of random polynomials has not been observed before.

The Newton polytope has an equally strong (though different) impact on the common zeros of  $k < m$  polynomials  $f_1, \dots, f_k$ . The image of the zero set of  $f_1, \dots, f_k$  under the moment map is (up to a logarithmic re-parameterization) known as an *amoeba* in the sense of [GKZ, Mi2]. Results on the expected distribution of amoebas can be obtained from our results on the expected zero current; for example, for a polytope  $\frac{1}{p}P$  with vertices in the interior of the standard unit simplex  $\Sigma \subset \mathbb{R}^m$ , there is also a *very forbidden region* which the amoeba almost surely avoids (see Corollary 1.6).

**1.1. Statement of results.** The patterns in zeros discussed here are statistical—they hold for random polynomials with prescribed Newton polytope—and are asymptotic as the degree of the polynomials tends to infinity. To state our problems and results precisely, let us recall some definitions. Let

$$(1) \quad f(z_1, \dots, z_m) = \sum_{\alpha \in \mathbb{N}^m: |\alpha| \leq p} a_\alpha \chi_\alpha(z_1, \dots, z_m), \quad \chi_\alpha(z) = z_1^{\alpha_1} \cdots z_m^{\alpha_m}$$

( $\alpha = (\alpha_1, \dots, \alpha_m)$ ,  $|\alpha| = \alpha_1 + \cdots + \alpha_m$ ) be a polynomial of degree  $p$  in  $m$  complex variables. By the *support* of  $f$  we mean the set

$$(2) \quad S_f = \{\alpha \in \mathbb{Z}^m : a_\alpha \neq 0\},$$

and by its *Newton polytope*  $P_f$  we mean the set

$$(3) \quad P_f := \text{the convex hull in } \mathbb{R}^m \text{ of } S_f.$$

Our aim is to study the statistical patterns in the space

$$(4) \quad \text{Poly}(P) = \{f \in \mathbb{C}[z_1, \dots, z_m] : P_f \subset P\}$$

of polynomials with support contained in a fixed Newton polytope  $P$ . We therefore need to find a natural measure on this space of polynomials. Since our purpose is to compare zero sets and masses as the polytope  $P$  varies, we view the polytope  $P$  as placing a condition on the Gaussian ensemble  $\text{Poly}(p\Sigma)$  of all polynomials of degree  $\leq p$  (where  $p\Sigma$  is the dilation by  $p$  of the standard unit simplex  $\Sigma$ ), and we give  $\text{Poly}(P)$  the resulting conditional probability measure. Here,  $p$  is chosen such that  $p\Sigma \supset P$ .

To describe this Gaussian ensemble, we first identify  $\text{Poly}(p\Sigma)$  with the space of homogeneous polynomials of degree  $p$  in  $m + 1$  variables in the usual way, i.e., by identifying  $f \in \text{Poly}(p\Sigma)$  with the homogeneous polynomial  $F$  such that  $F(1, z_1, \dots, z_m) = f(z_1, \dots, z_m)$ . Using this identification, we give the space  $\text{Poly}(p\Sigma)$  the  $\text{SU}(m + 1)$ -invariant inner product

$$(5) \quad \langle f_1, \bar{f}_2 \rangle := \langle F_1, \bar{F}_2 \rangle_{S^{2m+1}} = \frac{1}{m!} \int_{S^{2m+1}} F_1 \bar{F}_2 d\nu,$$

where  $d\nu$  is Haar probability measure on the  $(2m + 1)$ -sphere  $S^{2m+1} \subset \mathbb{C}^{m+1}$ . The monomials  $\chi_\alpha$  are orthogonal, so an orthonormal basis of  $\text{Poly}(p\Sigma)$  is given by  $\{\|\chi_\alpha\|^{-1}\chi_\alpha\}_{|\alpha| \leq p}$ , where  $\|\cdot\|$  denotes the norm in  $\text{Poly}(p\Sigma)$  given by (5). We emphasize here that the norms  $\{\|\chi_\alpha\|\}$  of the monomials  $\{\chi_\alpha\}$  do not depend on the constraining polytope  $P$ —they are given by the  $\text{SU}(m + 1)$ -invariant inner product (5) on  $\text{Poly}(p\Sigma)$ . (Of course they do depend on the choice of the integer  $p$ .)

The  $\text{SU}(m + 1)$ -invariant Gaussian measure  $\gamma_p$  corresponding to the inner product (5) is defined by

$$(6) \quad d\gamma_p(f) = \frac{1}{\pi^{k_p}} e^{-|c|^2} dc, \quad f = \sum_{|\alpha| \leq p} c_\alpha \frac{\chi_\alpha}{\|\chi_\alpha\|},$$

where  $k_p = \#\{\alpha \in \mathbb{N}^m : |\alpha| \leq p\} = \binom{m+p}{p}$ . Thus, the coefficients  $c_\alpha$  are independent complex Gaussian random variables with mean zero and variance one. We then endow the space  $\text{Poly}(P)$  with the associated *conditional probability measure*  $\gamma_{p|P}$ :

$$(7) \quad d\gamma_{p|P}(f) = \frac{1}{\pi^{\#P}} e^{-|c|^2} dc, \quad f = \sum_{\alpha \in P} c_\alpha \frac{\chi_\alpha}{\|\chi_\alpha\|},$$

where the coefficients  $c_\alpha$  are again independent complex Gaussian random variables with mean zero and variance one. (By an abuse of notation, we let  $\sum_{\alpha \in P}$  denote the sum over the lattice points  $\alpha \in P \cap \mathbb{Z}^m$ ;  $\#P$  denotes the cardinality of  $P \cap \mathbb{Z}^m$ .) We observe that  $\text{Poly}(P)$  inherits the inner product  $\langle f_1, \bar{f}_2 \rangle$  from  $\text{Poly}(p\Sigma)$ , and that  $\gamma_{p|P}$  is the induced Gaussian measure. Probabilities relative to  $\gamma_{p|P}$  can be considered as conditional probabilities; i.e. for any event  $E$ ,

$$\text{Prob}_\gamma\{f \in E | P_f = P\} = \text{Prob}_{\gamma_{p|P}}(E).$$

The expected distribution of mass and zeros of polynomials with fixed Newton polytope turns out to involve the moment map  $\mu : (\mathbb{C}^*)^m \rightarrow \mathbb{R}^m$  given by

$$(8) \quad \mu(z) = \left( \frac{|z_1|^2}{1 + \|z\|^2}, \dots, \frac{|z_m|^2}{1 + \|z\|^2} \right).$$

The image of  $\mu$  is the interior of the standard unit simplex  $\Sigma$  in  $\mathbb{R}^m$  with vertices at the points

$$(0, \dots, 0), (1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 0, 1).$$

The map  $\mu$  is the moment map of  $\mathbb{C}\mathbb{P}^m \supset (\mathbb{C}^*)^m$  (with the Fubini-Study symplectic form  $\omega_{\text{FS}}$ ) and plays a role in the geometric approach in [SZ2, STZ1, STZ2], where we regard polynomials with a fixed polytope as sections of a holomorphic line bundle on a toric variety (see §6).

By a *convex integral polytope*  $P$ , we mean the convex hull in  $\mathbb{R}^m$  of a finite set of points in  $\mathbb{Z}^m$ . We use the moment map  $\mu$  to describe our classically allowed regions:

**Definition.** Let  $P$  be a convex integral polytope in  $\mathbb{R}^m$  such that  $P \subset p\Sigma$ . The *classically allowed region* for polynomials in  $\text{Poly}(P)$  is the set

$$(9) \quad \mathcal{A}_P := \mu^{-1} \left( \frac{1}{p} P^\circ \right) \subset (\mathbb{C}^*)^m$$

(where  $P^\circ$  denotes the interior of  $P$ ), and the *classically forbidden region* is its complement  $(\mathbb{C}^*)^m \setminus \mathcal{A}_P$ . (If  $P$  has empty interior, i.e., if  $\dim P < m$ , then  $\mathcal{A}_P = \emptyset$ .)

All of our results on the asymptotic expected distribution of zeros of systems of random polynomials with prescribed Newton polytopes follow from our asymptotic result (Theorem 1.2) on the pointwise expected values of random polynomials with given Newton polytopes. Our results on zeros have a somewhat different flavor as the codimension (or number of polynomials in the system) varies. We first describe the results for the maximum codimension  $m$  where the zero set is almost surely discrete (Theorem 1.1). Then after describing the expected value distribution, we discuss the codimension 1 case, where the zero set is a hypersurface, and we finish with a result for general codimension that encompasses the previous results.

1.1.1. *Distribution of zeros: The point case.* Let us first consider the simultaneous zero set of  $m$  independent random polynomials in  $m$  variables. Bézout's theorem tells us that  $m$  generic homogeneous polynomials  $F_1, \dots, F_m$  of degree  $p$  have exactly  $p^m$  simultaneous zeros in  $(\mathbb{C}^*)^m$ . In fact, one immediately sees (by uniqueness of Haar probability measure on  $\mathbb{C}\mathbb{P}^m$ ) that the expected distribution of zeros is uniform with respect to the  $\text{SU}(m+1)$ -invariant projective volume form, when the ensemble is given the  $\text{SU}(m+1)$ -invariant measure  $d\gamma_p$ .

According to the Bernstein-Kouchnirenko Theorem [Be, Ko1, Ko2], the number of common zeros in  $(\mathbb{C}^*)^m$  of  $m$  generic polynomials  $\{f_1, \dots, f_m\}$  with given Newton polytope  $P$  equals  $m! \text{Vol}(P)$ , where  $\text{Vol}(P)$  is the Euclidean volume of  $P$ . For example, if  $P = p\Sigma$ , where  $\Sigma$  is the standard unit simplex in  $\mathbb{R}^m$ , then  $\text{Vol}(p\Sigma) = p^m \text{Vol}(\Sigma) = \frac{p^m}{m!}$ , and we get Bézout's theorem. (More generally, the  $f_j$  may have different Newton polytopes, in which case the number of zeros is given by the Bernstein-Kouchnirenko formula as a 'mixed volume' [Be, Ko1, Ko2].)

Now consider the ensemble of  $m$  independent random polynomials with Newton polytope  $P$ , equipped with the conditional probability (7). We let  $\mathbf{E}_{|P}(Z_{f_1, \dots, f_m})$  denote the expected density of their simultaneous zeros. More generally, for  $N \in \mathbb{Z}^+$  we define the expected zero density measure  $\mathbf{E}_{|NP}(Z_{f_1, \dots, f_m})$  by

$$(10) \quad \mathbf{E}_{|NP}(Z_{f_1, \dots, f_m})(U) = \int d\gamma_{Np|NP}(f_1) \cdots \int d\gamma_{Np|NP}(f_m) [\#\{z \in U : f_1(z) = \cdots = f_m(z) = 0\}],$$

for  $U \subset (\mathbb{C}^*)^m$ , where the integrals are over  $\text{Poly}(NP)$ . In fact,  $\mathbf{E}_{|NP}(Z_{f_1, \dots, f_m})$  is an absolutely continuous measure given by a  $\mathcal{C}^\infty$  density (see Proposition 5.1).

Our first result shows that, as the polytope  $P$  expands, these zeros are expected to concentrate in the classically allowed region and have (asymptotically) uniform density there. We measure the density of zeros with respect to the projective volume  $d \text{Vol}_{\mathbb{C}\mathbb{P}^m} = \frac{1}{m!} \omega_{\text{FS}}^m$ , where  $\omega_{\text{FS}} = \frac{i}{2\pi} \partial \bar{\partial} \log(1 + \|z\|^2)$  is the Fubini-Study Kähler form on  $\mathbb{C}^m \subset \mathbb{C}\mathbb{P}^m$ . (Note that with this normalization, the volume of  $\mathbb{C}\mathbb{P}^m$  is  $\frac{1}{m!}$ .)

**Theorem 1.1.** *Suppose that  $P \subset p\Sigma \subset \mathbb{R}^m$  is a convex integral polytope with nonempty interior. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{(Np)^m} \mathbf{E}_{|NP}(Z_{f_1, \dots, f_m}) = \begin{cases} \omega_{\text{FS}}^m & \text{on } \mathcal{A}_P, \\ 0 & \text{on } (\mathbb{C}^*)^m \setminus \mathcal{A}_P \end{cases}$$

in the measure sense; i.e., for any Borel set  $B \subset (\mathbb{C}^*)^m$ , we have

$$\frac{1}{(Np)^m} \mathbf{E}_{|NP}(\#\{z \in B : f_1(z) = \dots = f_m(z) = 0\}) \rightarrow m! \text{Vol}_{\mathbb{C}\mathbb{P}^m}(B \cap \mathcal{A}_P).$$

Theorem 1.1 is a special case of our general result on zeros (Theorem 1.4). In fact, our results imply that the convergence of the zero current on the classically allowed region is exponentially fast in the sense that

$$\mathbf{E}_{|NP}(Z_{f_1, \dots, f_m}) = (Np)^m \omega_{\text{FS}}^m + O(e^{-\lambda N}) \quad \text{on } \mathcal{A}_P,$$

for some positive continuous function  $\lambda$  on  $\mathcal{A}_P$ .

Note that we can choose any integer  $p$  such that  $p\Sigma \supset P$ . Theorem 1.1 (as well as our other results) depends quite strongly on the choice of  $p$ , since the classically allowed region  $\mathcal{A}_P$  shrinks as  $p$  is increased. In order to minimize notational clutter,  $p$  does not appear in  $\mathcal{A}_P$ ,  $\mathbf{E}_{|NP}$ ,  $|f(z)|_{\text{FS}}$ , and some other expressions to be defined below.

Figure 1 shows the classically allowed region (shaded) and the classically forbidden region (unshaded) when  $P$  is the unit square in  $\mathbb{R}^2$  (and  $p = 2$ ):

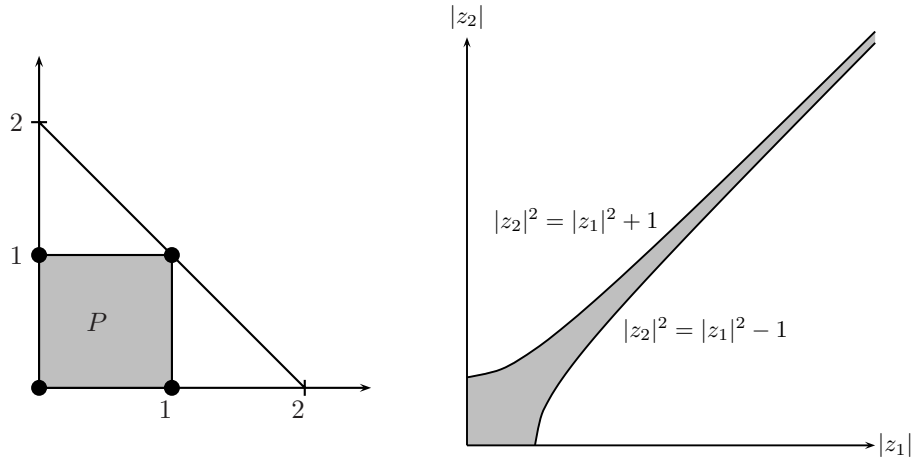


FIGURE 1. The classically allowed region for  $P = [0, 1] \times [0, 1]$

1.1.2. *Mass distribution.* By the mass density of a polynomial  $f \in \text{Poly}(p\Sigma)$  at a point  $z \in (\mathbb{C}^*)^m$ , we mean the square  $|f(z)|_{\text{FS}}^2$  of the Fubini-Study norm

$$(11) \quad |f(z)|_{\text{FS}} := \frac{|f(z)|}{(1 + \|z\|^2)^{p/2}} = |F(\zeta)|, \quad \zeta = \frac{1}{\sqrt{1 + \|z\|^2}}(1, z_1, \dots, z_m) \in S^{2m+1},$$

where  $F$  is the homogenization of  $f$  described above. Our next result describes the asymptotics of the expected mass density of  $\mathcal{L}^2$ -normalized polynomials with Newton polytope  $NP$ , i.e., the expected density  $\mathbf{E}_{\nu_{NP}}(|f(z)|_{\text{FS}}^2)$  with respect to the Haar probability measure, denoted by  $\nu_{NP}$ , on the unit sphere in  $\text{Poly}(NP)$ . We shall show that  $\mathbf{E}_{\nu_{NP}}(|f(z)|_{\text{FS}}^2)$  is asymptotically uniform with respect to Fubini-Study measure in the classically allowed region (as if there were no constraint at all); while in the forbidden region the mass decays exponentially. Thus in the semiclassical limit  $N \rightarrow \infty$ , all the mass concentrates in the classically allowed region.

Behavior of the expected mass density (as well as the zero distribution in positive dimension) in the forbidden region is subtle, so we pause to introduce the relevant concepts. We show in §4.2.1 that the classically forbidden region decomposes into the disjoint union of the normal ‘flow-outs’ of boundary points of  $P$ ,

$$(12) \quad \text{Flow}(x) := \{e^\tau \cdot z : \tau \in C_x, p\mu(z) = x\}, \quad x \in \partial P \cap p\Sigma^\circ,$$

where  $C_x \subset \mathbb{R}^m$  is the normal cone to  $P$  at  $x$  (see §2.1). Here  $r \cdot z = (r_1 z_1, \dots, r_m z_m)$  denotes the  $\mathbb{R}_+^m$  action on  $(\mathbb{C}^*)^m$ .

We use these flow-outs to divide the forbidden region into subregions  $\mathcal{R}_F$  as follows: Recall that the boundary of a polytope  $P$  decomposes into a disjoint union of its faces (of all dimensions). For each face  $F$  of  $P$  not contained in  $\partial(p\Sigma)$ , we let

$$(13) \quad \mathcal{R}_F = \bigcup_{x \in F} \text{Flow}(x) = \{e^\tau \cdot z : \tau \in C_F, p\mu(z) \in F\},$$

where  $C_F$  is the normal cone to  $P$  along the face  $F$ ; i.e., the normal cone to  $P$  at any point of  $F$  (see §2.1). Note that for the case where  $F$  is the open face  $P^\circ$ ,  $\mathcal{R}_F$  is the classically allowed region  $\mathcal{A}_P$ . We call  $\mathcal{R}_F$  the *flow-out* of the face  $F$ . The regions  $\mathcal{R}_F$  all have nonempty interior. A sample illustration of these regions is given in Figure 4 in §1.2.3.

We can now state our result on mass asymptotics:

**Theorem 1.2.** *Suppose that  $P \subset p\Sigma \subset \mathbb{R}^m$  is a convex integral polytope such that  $P \not\subset \partial(p\Sigma)$ . Then the expected mass density of random  $\mathcal{L}^2$ -normalized polynomials with Newton polytope  $NP$  is given by the asymptotic formulas:*

$$\mathbf{E}_{\nu_{NP}}(|f(z)|_{\text{FS}}^2) \sim c_0 + c_1 N^{-1} + c_2 N^{-2} + \dots, \quad \text{for } z \in \mathcal{A}_P,$$

$$\mathbf{E}_{\nu_{NP}}(|f(z)|_{\text{FS}}^2) = N^{-s/2} e^{-N b_P(z)} [c_0^F(z) + O(N^{-1})], \quad \text{for } z \in \mathcal{R}_F^\circ, F \subset \Sigma^\circ,$$

where  $c_0 = \frac{P^m}{\text{Vol}(P)}$ ,  $s = \text{codim } F$ ,  $c_0^F \in \mathcal{C}^\infty(\mathcal{R}_F^\circ)$ , and  $b_P$  is a positive  $\mathcal{C}^1$  function on  $(\mathbb{C}^*)^m \setminus \overline{\mathcal{A}_P}$ . Furthermore, the remainder estimates are uniform on compact subsets of the open regions  $\mathcal{R}_F^\circ$  and of  $\mathcal{A}_P$ .

We shall give a formula for  $b_P(z)$  below. A more precise asymptotic formula for the expected mass density is given by Theorem 4.1 (see also (74)).

It follows that

$$\mathbf{E}_{\nu_{NP}} (|f(z)|_{\mathbb{F}\mathbb{S}}^2) \longrightarrow \begin{cases} \frac{p^m}{\text{Vol}(P)} & \text{for } z \in \mathcal{A}_P, \\ 0 & \text{for } z \in (\mathbb{C}^*)^m \setminus \overline{\mathcal{A}_P}, \end{cases}$$

as illustrated in Figure 2 (plotted using Maple) for the case where  $P$  is the unit square. (Recall Figure 1 for the depiction of  $\mathcal{A}_P$  for this case.)

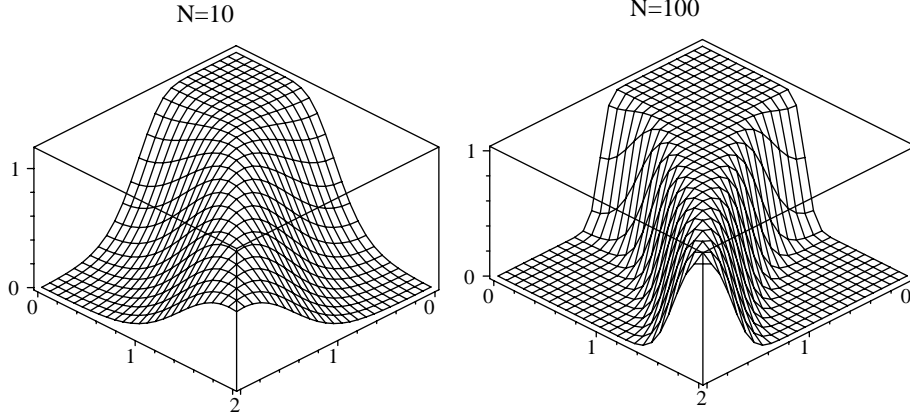


FIGURE 2.  $\frac{1}{4} \mathbf{E}_{\nu_{NP}} (|f(z)|_{\mathbb{F}\mathbb{S}}^2)$  for  $P = [0, 1] \times [0, 1]$

Heuristically, the mass concentration in  $\mathcal{A}_P$  can be understood as follows: the mass of a monomial  $\chi_\alpha \in \text{Poly}(NP)$  concentrates (exponentially) on the torus  $\mu^{-1}(\frac{1}{NP}\alpha)$ . The constraint  $S_f \subset NP$  thus concentrates all the mass in  $\mathcal{A}_P$ , with exponentially small errors, and the mass there is uniformly distributed with small errors. Such mass concentration of monomials is given in formula (76) and in more detail in [STZ2].

We also show that  $\frac{1}{N} \log \mathbf{E}_{\nu_{NP}} (|f(z)|_{\mathbb{F}\mathbb{S}}^2) \rightarrow -b_P(z)$  uniformly on compact subsets of  $(\mathbb{C}^*)^m$  (Proposition 4.2), even at the interfaces of the regions  $\mathcal{R}_F$ . Combined with the Poincaré-Lelong formula, this uniform convergence result leads to the asymptotics of the expected distribution of zeros of a single random polynomial with polytope  $NP$ . Using independence of sections and the Bedford-Taylor Theorem [BT] on continuity of the Monge-Ampère operator, we obtain asymptotics for any number  $1 \leq k \leq m$  of sections.

We shall show that  $b_P$  is  $\mathcal{C}^\infty$  on the regions  $\mathcal{R}_F^\circ$ , but it is not even  $\mathcal{C}^2$  on the interfaces between these regions. To state our formula for  $b_P$ , we let  $q(z)$  denote the unique point in  $\partial P$  such that  $z \in \text{Flow}(q(z))$  and we let  $\tau_z \in C_{q(z)}$  be such that  $z = e^{\tau_z/2} \cdot \xi$  where  $\xi \in \mu^{-1}(\frac{1}{p}q(z))$ . Thus  $q(z)$ ,  $\tau_z$  are given by the conditions:

$$(14) \quad p\mu(e^{-\tau_z/2} \cdot z) = q(z) \in \partial P,$$

$$(15) \quad \tau_z \in C_{q(z)}.$$

The existence and uniqueness of  $q(z)$  and  $\tau_z$  are stated in Lemma 4.3.

We then have

$$(16) \quad b_P(z) = -\langle q(z), \tau_z \rangle + p \log \left( \frac{1 + \|z\|^2}{1 + \|e^{-\tau_z/2} \cdot z\|^2} \right) \quad \text{for } z \in (\mathbb{C}^*)^m \setminus \mathcal{A}_P.$$

To understand formula (16) better, for any  $x \in P$  (not necessarily a lattice point), we consider the ‘monomials’

$$(17) \quad \begin{aligned} |\chi_x(z)| &:= |z|^x = |z_1|^{x_1} \cdots |z_m|^{x_m}, \\ |\widehat{\chi}_x(z)| &:= |\chi_x(z)|_{\text{FS}} = \frac{|\chi_x(z)|}{(1 + \|z\|^2)^{p/2}}, \\ \mathcal{M}_x(z) &:= \frac{|\widehat{\chi}_x(z)|}{\sup |\widehat{\chi}_x|}, \end{aligned}$$

for  $z \in (\mathbb{C}^*)^m$ . Thus the normalized monomial  $\mathcal{M}_x$  has sup-norm 1; in fact it takes its maximum on the torus  $\mu^{-1}(\frac{1}{p}x)$ . One easily checks that

$$(18) \quad b_P(z) = -2 \log \mathcal{M}_{q(z)}(z),$$

and thus Theorem 1.2 says that

$$(19) \quad \mathbf{E}_{\nu_{NP}} (|f(z)|_{\text{FS}}^2) = [c_0^F(z) + O(N^{-1})] N^{-s/2} \mathcal{M}_{q(z)}(z)^{2N}, \quad z \in \mathcal{R}_F^\circ.$$

We will also obtain an integral formula for  $b_P$  (see Proposition 5.3), which allows us to interpret  $b_P$  as an ‘Agmon distance’ (see §1.4), and  $q(z)$  is the closest point of  $P$  to  $p\mu(z)$  in this sense.

We do not assume in Theorem 1.2 that  $P$  has interior. However, if  $P^\circ = \emptyset$  (i.e., if  $\dim P < m$ ), then we must assume that  $P \not\subset \partial(p\Sigma)$  since if  $P$  is contained in a face of  $p\Sigma$ , the decay function  $b_P$  is not defined. Note that if  $P^\circ = \emptyset$ , then  $\mathcal{A}_P$  is vacuous and the first expansion of the theorem does not occur.

1.1.3. *Distribution of zeros: Arbitrary codimension.* Our next results concern the zero set of one or more random polynomials. Since the zero set

$$(20) \quad |Z_{f_1, \dots, f_k}| := \{z \in (\mathbb{C}^*)^m : f_1(z) = \cdots = f_k(z) = 0\}$$

is a submanifold of complex codimension  $k$  (without multiplicity), for almost all polynomials  $f_1, \dots, f_k$ , it defines a *current of integration*  $Z_{f_1, \dots, f_k} \in \mathcal{D}'^{k, k}((\mathbb{C}^*)^m)$ . We recall that this current is given by

$$(21) \quad (Z_{f_1, \dots, f_k}, \varphi) := \int_{|Z_{f_1, \dots, f_k}|} \varphi, \quad \text{for test forms } \varphi \in \mathcal{D}^{m-k, m-k}((\mathbb{C}^*)^m).$$

The zero set  $|Z_{f_1, \dots, f_k}|$  also carries a natural  $(2m - 2k)$ -dimensional Riemannian volume measure (induced from the Fubini-Study metric on  $\mathbb{C}\mathbb{P}^m$ ), denoted  $\|Z_{f_1, \dots, f_k}\|$ , against which one can integrate scalar functions. Since the volume form on any complex  $n$ -dimensional holomorphic submanifold of  $\mathbb{C}\mathbb{P}^m$  is given by the restriction of  $\frac{1}{n!} \omega_{\text{FS}}^n$ , the volume measure on the zero set is given by  $\|Z_{f_1, \dots, f_k}\| = Z_{f_1, \dots, f_k} \wedge \frac{1}{(m-k)!} \omega_{\text{FS}}^{m-k}$ ; i.e.,

$$(22) \quad (\|Z_{f_1, \dots, f_k}\|, \varphi) = \int_{|Z_{f_1, \dots, f_k}|} \varphi d\text{Vol}_{2m-2k} = \frac{1}{(m-k)!} \int_{|Z_{f_1, \dots, f_k}|} \varphi \omega_{\text{FS}}^{m-k},$$

for  $\varphi \in \mathcal{D}((\mathbb{C}^*)^m)$ .

As a measure,  $\|Z_{f_1, \dots, f_k}\|(U) = \text{Vol}(|Z_{f_1, \dots, f_k}| \cap U)$  for open sets  $U \subset (\mathbb{C}^*)^m$ . (For the theory of currents defined by complex algebraic or analytic varieties, see for example, [Sh, I.3].) We shall discuss the expected values of both the current of integration  $Z_{f_1, \dots, f_k}$  and the measure  $\|Z_{f_1, \dots, f_k}\|$ .

We first consider the expected distribution of zeros of one polynomial. We denote by  $\mathbf{E}_{|P}(Z_f) = \mathbf{E}_{\gamma_P|P}(Z_f)$  the conditional expectation of the zero current of a random



polynomial  $f \in \text{Poly}(P)$  with Newton polytope  $P$ . In fact,  $\mathbf{E}|_P(Z_f)$  is actually a smooth  $(1, 1)$ -form on  $(\mathbb{C}^*)^m$  (Proposition 5.1).

Let us recall what happens when  $P = p\Sigma$ . By the uniqueness of the  $\text{SU}(m+1)$ -invariant Kähler form  $\omega_{\text{FS}}$  on  $\mathbb{C}\mathbb{P}^m$ , the expected zero current  $\mathbf{E}(Z_f)$  taken over all polynomials of degree  $p$  is given by  $p\omega_{\text{FS}}$ , where  $\omega_{\text{FS}} = \frac{i}{2\pi}\partial\bar{\partial}\log\|z\|^2$  is the Fubini-Study Kähler form on  $\mathbb{C}\mathbb{P}^m$ . Thus the expected distribution of zeros, as well as the tangents to the zero varieties, is uniform over  $\mathbb{C}\mathbb{P}^m$ . We now describe how the expectation changes if we add the condition that  $P_f = P$ .

**Theorem 1.3.** *Let  $P \subset p\Sigma \subset \mathbb{R}^m$  be a convex integral polytope such that  $P \not\subseteq \partial(p\Sigma)$ . Then there exists a closed semipositive  $(1, 1)$ -form  $\psi_P$  on  $(\mathbb{C}^*)^m$  with piecewise  $C^\infty$  coefficients such that*

- (i)  $N^{-1}\mathbf{E}|_{NP}(Z_f) \rightarrow \psi_P$  in  $\mathcal{L}_{\text{loc}}^1((\mathbb{C}^*)^m)$ .
- (ii)  $\psi_P = p\omega_{\text{FS}}$  on the classically allowed region  $\mu^{-1}(\Sigma \setminus \frac{1}{p}P)$ .
- (iii) On each region  $\mathcal{R}_F^\circ$ , the  $(1, 1)$ -form  $\psi_P$  is  $C^\infty$  and has constant rank equal to  $\dim F$ ; in particular, if  $v \in p\Sigma^\circ$  is a vertex of  $P$ , then  $\psi_P|_{\mathcal{R}_v^\circ} = 0$ .

We see from part (iii) that the zero set  $|Z_f|$  of a polynomial with polytope  $NP$  typically intersects the classically forbidden region  $\mu^{-1}(\Sigma \setminus \frac{1}{p}P)$  in the semiclassical limit  $N \rightarrow \infty$ . However, there are subtler ‘very forbidden regions’ that  $|Z_f|$  avoids in the case where the polytope has vertices in the interior of  $p\Sigma$ , namely the regions  $\mathcal{R}_v^\circ$  comprising the flow-out of these vertices.

As a corollary, we obtain some statistical results on the so-called ‘tentacles’ of amoebas in dimension 2 (see §5.5). Roughly speaking, the (compact) amoeba of a polynomial  $f(z_1, z_2)$  is the image of the Riemann surface  $Z_f$  under the moment map  $\mu$  on  $(\mathbb{C}^*)^2$ , and the tentacles are the ends of the amoeba. Certain tentacles must end at vertices of the triangle  $\Sigma$  while others are ‘free’ to end anywhere along the boundary of  $\Sigma$ . In Corollary 5.6, we will prove that (in the limit  $N \rightarrow \infty$ ) almost all of the free tentacles of typical amoebas tend to end in the classically allowed portion of  $\partial\Sigma$ .

We call the form  $\psi_P$  in Theorem 1.3 the *limit expected zero current*. Our explicit formula for  $\psi_P$  is

$$(23) \quad \psi_P = p\omega_{\text{FS}} - \frac{\sqrt{-1}}{2\pi}\partial\bar{\partial}b_P,$$

where  $b_P$  is given by (16). In fact, (23) holds as an equation of currents, and the current  $\psi_P \in \mathcal{D}'^{1,1}((\mathbb{C}^*)^m)$  is closed and positive. By  $\mathcal{L}_{\text{loc}}^1$  convergence in (i), we mean  $\mathcal{L}^1$  convergence of the coefficients on every compact subset of  $(\mathbb{C}^*)^m$ . (Recall that  $\mathbf{E}|_{NP}(Z_{f_1, \dots, f_k})$  is a  $(k, k)$ -form with smooth coefficients.) If we write  $\psi_P = \sqrt{-1}\sum\psi_{jk}(z)dz_j \wedge d\bar{z}_k$ , then  $(\psi_{jk}(z))$  is a semipositive Hermitian matrix. By the rank of  $\psi_P$  at  $z$ , we mean the rank of the matrix  $(\psi_{jk}(z))$ . Note that if  $\mathcal{R}_F$  and  $\mathcal{R}_{F'}$  are adjoining regions (i.e., have a common codimension 1 interface), then  $F$  and  $F'$  are of different dimensions, so  $\psi_P$  must be discontinuous along the interface.

Boundary points of the  $\mathcal{R}_F$  are called *transition points*. The set of transition points comprises the discontinuities of  $\psi_P$ . Points of  $\partial\mathcal{A}_P$  are always transition points, and there will be others whenever  $P$  has a face  $F \subset p\Sigma^\circ$  of codimension at least 2. For example, consider the case where  $P$  is the unit square. In this case, there are two interior faces whose flow-outs are the connected components of the

classically forbidden region (see Figure 1) and the set of transition points equals  $\partial\mathcal{A}_P$ . We shall also give an example with an interior vertex (see §1.2.3), where the forbidden region is connected but decomposes into flow-outs of 3 faces (see Figure 4) and there are transition points not in  $\partial\mathcal{A}_P$ .

The form  $\psi_P$  not only encodes the expected (normalized) density of the zero set, but also the expected density of tangent directions to the zero set. In the course of the proof of Theorem 1.3, we will show that in the forbidden region, the limit tangent directions are restricted. In particular, as the polytope expands, the tangent spaces to typical zero sets approach tangency to the ‘normal flow’  $\{e^{\tau+i\theta} \cdot z^0 : \tau, \theta \in T_F^\perp \subset \mathbb{R}^m\}$ . A precise formulations of this fact is given in Theorem 5.5. Thus, while the expected distribution of zero densities is absolutely continuous, the expected distribution of zero tangents is singular.

Finally, we consider the general case of  $k \leq m$  independently chosen random polynomials

$$f_j \in \text{Poly}(P_j), \quad P_j \subset p_j\Sigma, \quad P_j \not\subset \partial(p_j\Sigma) \quad (1 \leq j \leq k),$$

and we let  $\mathbf{E}_{|P_1, \dots, P_k}(Z_{f_1, \dots, f_k})$  denote the expected zero current with respect to the probability measure  $\gamma_{P_1|P_1} \times \dots \times \gamma_{P_k|P_k}$  on the product space. If  $P_1 = \dots = P_k = P$ , then we also write  $\mathbf{E}_{|P, \dots, P}(Z_{f_1, \dots, f_k}) = \mathbf{E}_{|P}(Z_{f_1, \dots, f_k})$ .

**Theorem 1.4.** *Let  $P_1, \dots, P_k$  be convex integral polytopes in  $\mathbb{R}_{\geq 0}^m$ . Then*

$$N^{-k} \mathbf{E}_{|NP_1, \dots, NP_k}(Z_{f_1, \dots, f_k}) \rightarrow \psi_{P_1} \wedge \dots \wedge \psi_{P_k} \quad \text{in } \mathcal{L}_{\text{loc}}^1((\mathbb{C}^*)^m), \quad \text{as } N \rightarrow \infty.$$

Theorem 1.1 is a consequence of Theorems 1.3 and 1.4. Furthermore, the expected volume of the zero set of a system of  $k$  polynomials has the following exotic distribution law, as given by the following corollaries:

**Corollary 1.5.** *Let  $P_1, \dots, P_k$  be convex integral polytopes in  $\mathbb{R}_{\geq 0}^m$ . Then for every relatively compact, open set  $U \subset (\mathbb{C}^*)^m$ , we have*

$$\frac{1}{N^k} \mathbf{E}_{|NP_1, \dots, NP_k} \text{Vol}(|Z_{f_1, \dots, f_k}| \cap U) \rightarrow \frac{1}{(m-k)!} \int_U \psi_{P_1} \wedge \dots \wedge \psi_{P_k} \wedge \omega_{\text{FS}}^{m-k}.$$

**Corollary 1.6.** *The expected zero current  $N^{-k} \mathbf{E}_{|NP}(Z_{f_1, \dots, f_k})$  tends to 0 at all points of each forbidden subregion  $\mathcal{R}_F$  with  $\dim F < k$ .*

**1.2. Examples.** We illustrate the results of Theorems 1.2–1.3 in some simple cases in dimension  $m = 2$ . For our examples, we shall describe the forbidden subregions  $\mathcal{R}_F$  given by (12)–(13) and compute the decay function  $b_P(z)$  and the limit expected zero current  $\psi_P$ .

To simplify our computation of  $\psi_P$  from (23), we write

$$u(z) = -b_P(z) + p \log(1 + \|z\|^2) = p \log(1 + \|e^{-\tau z/2} \cdot z\|^2) + \langle q(z), \tau_z \rangle,$$

so that

$$(24) \quad \psi_P = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \left[ p \log(1 + \|e^{-\tau z/2} \cdot z\|^2) + \langle q(z), \tau_z \rangle \right].$$

**1.2.1. Example 1: The square.** For our first example, we let  $P$  be the unit square with vertices  $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$  and we let  $p = 2$ . Recalling that

$$\mu(z_1, z_2) = \left( \frac{|z_1|^2}{1 + |z_1|^2 + |z_2|^2}, \frac{|z_2|^2}{1 + |z_1|^2 + |z_2|^2} \right),$$

we see that the classically allowed region is given by

$$\mathcal{A}_P = \{(z_1, z_2) : |z_1|^2 - 1 < |z_2|^2 < |z_1|^2 + 1\},$$

as illustrated in Figure 1. The forbidden region consists of two subregions:

$$\mathcal{R}_F = \{(z_1, z_2) : |z_2|^2 \geq |z_1|^2 + 1\}, \quad F = \{(x_1, 1) : 0 \leq x_1 \leq 1\},$$

$$\mathcal{R}_{F^*} = \{(z_1, z_2) : |z_2|^2 \leq |z_1|^2 - 1\}, \quad F^* = \{(1, x_2) : 0 \leq x_2 \leq 1\}.$$

Suppose that  $z$  is a point in the upper forbidden region  $\mathcal{R}_F$ . Recalling (14)–(15), we write  $\tau_z = (\tau_1, \tau_2)$ ; then  $\tau_1 = 0$  since  $\tau \perp T_F$ . The boundary point  $q(z)$  of  $P$  whose normal flow contains  $z$  is given by

$$q(z) = 2\mu(e^{-\tau_z/2} \cdot z) = (a, 1) \in F \subset \partial P.$$

Writing

$$|z_1|^2 = s_1, \quad |z_2|^2 = s_2, \quad |e^{-\tau_2/2} z_2|^2 = e^{-\tau_2} s_2 = \tilde{s}_2,$$

we have

$$\frac{s_1}{1 + s_1 + \tilde{s}_2} = \frac{a}{2}, \quad \frac{\tilde{s}_2}{1 + s_1 + \tilde{s}_2} = \frac{1}{2}.$$

Therefore

$$s_1 = \frac{a}{1-a}, \quad \tilde{s}_2 = \frac{1}{1-a} = \frac{s_1}{a}, \quad a = \frac{s_1}{1+s_1} = \frac{|z_1|^2}{1+|z_1|^2},$$

$$e^{-\tau_2} = \tilde{s}_2/s_2 = \frac{s_1}{as_2} = \frac{1+|z_1|^2}{|z_2|^2}.$$

We have

$$\log(1 + \|e^{-\tau_z/2} \cdot z\|^2) = \log\left(1 + |z_1|^2 + \frac{1+|z_1|^2}{|z_2|^2}|z_2|^2\right) = \log(1 + |z_1|^2) + \log 2,$$

$$\langle q(z), \tau_z \rangle = \left\langle \left(\frac{|z_1|^2}{1+|z_1|^2}, 1\right), \left(0, \log \frac{|z_2|^2}{1+|z_1|^2}\right) \right\rangle = \log |z_2|^2 - \log(1 + |z_1|^2).$$

Therefore

$$u = \log |z_2|^2 + \log(1 + |z_1|^2) + \log 4.$$

We conclude that

$$\psi_P = \begin{cases} \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + |z_1|^2) & \text{for } z \in \mathcal{R}_F^\circ \\ & (|z_1|^2 + 1 < |z_2|^2), \\ \frac{\sqrt{-1}}{\pi} \partial \bar{\partial} \log(1 + |z_1|^2 + |z_2|^2) = 2\omega_{\text{FS}} & \text{for } z \in \mathcal{A}_P \\ & (|z_1|^2 - 1 < |z_2|^2 < |z_1|^2 + 1), \\ \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + |z_2|^2) & \text{for } z \in \mathcal{R}_{F^*}^\circ \\ & (|z_2|^2 < |z_1|^2 - 1) \end{cases}$$

(where the third case is by symmetry). Note that  $\psi_P$  has constant rank 1 in both of the forbidden regions  $\mathcal{R}_F$ ,  $\mathcal{R}_{F^*}$ , as indicated in Theorem 1.3(iii).

On  $\mathcal{R}_F$ , we have

$$e^{-b_P(z)} = \frac{4|z_2|^2(1+|z_1|^2)}{(1+|z_1|^2+|z_2|^2)^2} \quad \text{for } |z_1|^2 + 1 < |z_2|^2.$$

On the boundary  $\{|z_1|^2 = |z_2|^2 - 1\}$ , we have  $e^{-b_P(z)} = 1$  as expected. On  $\{|z_1| = c\}$ , we have the growth rate  $e^{-b_P(z)} \sim 1/|z_2|^2$  as  $z_2 \rightarrow \infty$ . To obtain  $e^{-b_P(z)}$  on  $\mathcal{R}_{F^*}$ , we interchange  $z_1$  and  $z_2$  in the above.

1.2.2. *Example 2: A trapezoid.* We next consider the trapezoidal polytope of Figure 3. Comparing with the case of the square, we see that the classically allowed region is given by  $|z_2|^2 < |z_1|^2 + 1$ , and the forbidden region coincides with the upper forbidden region  $\mathcal{R}_F = \{|z_2|^2 \geq |z_1|^2 + 1\}$  from Example 1. Thus, the map  $z \mapsto (\tau_z, q(z))$  is the same as before when  $z$  is in the forbidden region  $\mathcal{R}_F$ . Hence,  $e^{-b_P(z)}$  and  $\psi_P$  are also the same as in Example 1 on  $\mathcal{R}_F$ . On the classically allowed region,  $e^{-b_P(z)} = 1$  and  $\psi_P = 2\omega_{\text{FS}}$ .

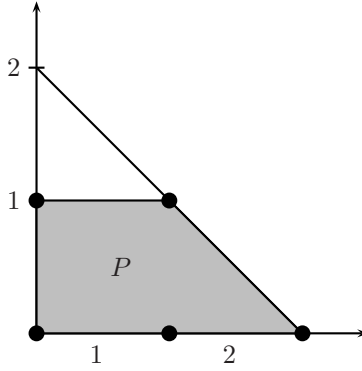


FIGURE 3.  $P \cap \mathbb{Z}^2 = \{(0, 0), (1, 0), (2, 0), (0, 1), (1, 1)\}$

1.2.3. *Example 3: Trapezoids of higher degree.* Now let  $n \geq 2$  and let  $P$  be the trapezoid with vertices  $(0, 0), (n+1, 0), (0, 1), (1, 1)$  given in Figure 4. In this case,  $p = n+1$  and  $P$  has an interior vertex  $v = (1, 1) \in \Sigma^\circ$ .

We see that the classically allowed region  $\mu^{-1}(\frac{1}{p}P^\circ)$  is given by

$$(25) \quad \mathcal{A}_P = \left\{ (z_1, z_2) : |z_2|^2 < \min \left\{ \frac{|z_1|^2 + 1}{n}, \frac{1}{n-1} \right\} \right\}.$$

This time, the forbidden region consists of three subregions:  $\mathcal{R}_F$ ,  $\mathcal{R}_v$ ,  $\mathcal{R}_{F'}$ , where  $F = \{(x_1, 1) : 0 \leq x_1 < 1\}$ ,  $F' = \{(x_1, x_2) : x_2 = \frac{1}{n}(n+1-x_1), 1 < x_1 \leq n+1\}$ .

(See Figure 4.)

Suppose that  $z$  is a point in the region  $\mathcal{R}_F$ . Then  $\tau_z = (0, \tau_2)$ , and

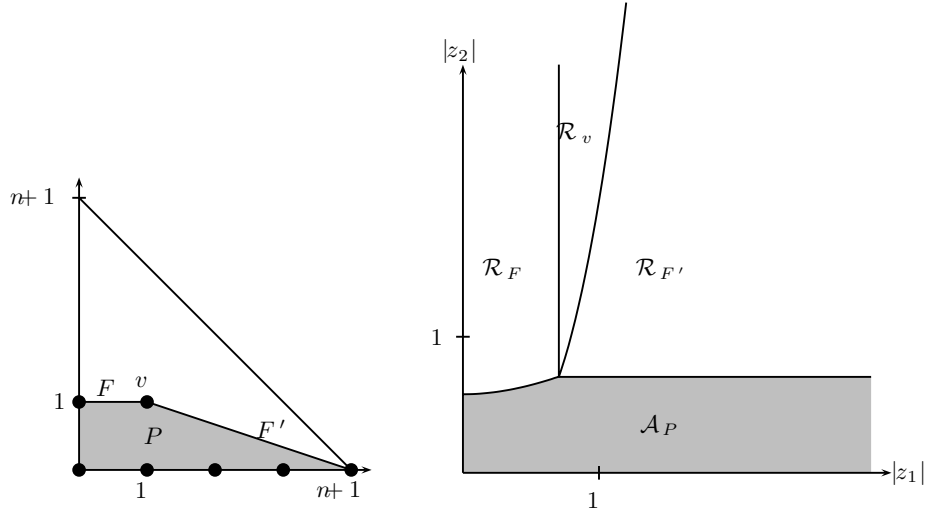
$$q(z) = (n+1)\mu(e^{-\tau_2/2} \cdot z) = (a, 1) \in (n+1)F \subset \partial P \quad (0 < a < 1).$$

Again writing

$$|z_1|^2 = s_1, \quad |z_2|^2 = s_2, \quad |e^{-\tau_2/2} z_2|^2 = e^{-\tau_2} s_2 = \tilde{s}_2,$$

we have

$$\frac{s_1}{1 + s_1 + \tilde{s}_2} = \frac{a}{n+1}, \quad \frac{\tilde{s}_2}{1 + s_1 + \tilde{s}_2} = \frac{1}{n+1}.$$

FIGURE 4.  $P \cap \mathbb{Z}^2 = \{(0, 0), (1, 0), \dots, (n+1, 0), (0, 1), (1, 1)\}$ 

Therefore

$$s_1 = \frac{a}{n-a}, \quad \tilde{s}_2 = \frac{1}{n-a} = \frac{s_1}{a}, \quad a = \frac{ns_1}{1+s_1} = \frac{n|z_1|^2}{1+|z_1|^2},$$

$$e^{-\tau_2} = \tilde{s}_2/s_2 = \frac{s_1}{as_2} = \frac{1+|z_1|^2}{n|z_2|^2}.$$

In particular,

$$a < 1 \Leftrightarrow |z_1|^2 < \frac{1}{n-1}$$

and therefore

$$(26) \quad \mathcal{R}_F = \left\{ (z_1, z_2) : |z_2|^2 \geq \frac{|z_1|^2 + 1}{n}, |z_1|^2 < \frac{1}{n-1} \right\}.$$

We have

$$\log(1 + \|e^{-\tau_z/2} \cdot z\|^2) = \log \left( 1 + |z_1|^2 + \frac{1 + |z_1|^2}{n|z_2|^2} |z_2|^2 \right) = \log(1 + |z_1|^2) + \log \frac{n+1}{n},$$

$$\langle q(z), \tau_z \rangle = \left\langle \left( \frac{n|z_1|^2}{1+|z_1|^2}, 1 \right), \left( 0, \log \frac{n|z_2|^2}{1+|z_1|^2} \right) \right\rangle = \log |z_2|^2 - \log(1 + |z_1|^2) + \log n.$$

Therefore

$$u = \log |z_2|^2 + n \log(1 + |z_1|^2) + (n+1) \log(n+1) - n \log n.$$

Hence,

$$\left. \begin{aligned} \psi_P &= n \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + |z_1|^2) \\ e^{-b_P(z)} &= \frac{(n+1)^{n+1}}{n^n} \frac{|z_2|^2 (1 + |z_1|^2)^n}{(1 + |z_1|^2 + |z_2|^2)^{n+1}} \end{aligned} \right\} \quad \text{for } z \in \mathcal{R}_F.$$

Now suppose that  $z$  is a point in  $\mathcal{R}_{F'}$ . Since  $\tau_z \perp T_{F'}$ , we can write  $\tau_z = (\tau_1, n\tau_1)$ . Let

$$(27) \quad q(z) = (n+1)\mu(e^{-\tau_z/2} \cdot z) = \left(c, \frac{1}{n}(n+1-c)\right) \in F'.$$

As before, we write

$$s_1 = |z_1|^2, \quad s_2 = |z_2|^2, \quad \tilde{s}_1 = |e^{-\tau_1/2} z_1|^2 = e^{-\tau_1} s_1, \quad \tilde{s}_2 = |e^{-n\tau_1/2} z_2|^2 = e^{-n\tau_1} s_2.$$

By (27), we have

$$\frac{\tilde{s}_1}{1 + \tilde{s}_1 + \tilde{s}_2} = \frac{c}{n+1}, \quad \frac{\tilde{s}_2}{1 + \tilde{s}_1 + \tilde{s}_2} = \frac{1}{n} \left(1 - \frac{c}{n+1}\right).$$

Solving for  $\tilde{s}_1, \tilde{s}_2$ , we obtain

$$(28) \quad \tilde{s}_1 = \frac{n}{n-1} \frac{c}{n+1-c}, \quad \tilde{s}_2 = \frac{1}{n-1}.$$

Therefore,  $e^{-n\tau_1} = \tilde{s}_2/s_2 = |z_2|^{-2}/(n-1)$ , so we have

$$\tau_1 = \frac{1}{n} \log(n-1)|z_2|^2, \quad \tilde{s}_1 = \frac{|z_1|^2}{(n-1)^{1/n}|z_2|^{2/n}}.$$

Thus,

$$\log(1 + \|e^{-\tau_z/2} \cdot z\|^2) = \log(1 + \tilde{s}_1 + \tilde{s}_2) = \log\left(\frac{n}{n-1} + \frac{|z_1|^2}{(n-1)^{1/n}|z_2|^{2/n}}\right).$$

By (27),

$$\langle q(z), \tau_z \rangle = \left\langle \left(c, \frac{1}{n}(n+1-c)\right), (\tau_1, n\tau_1) \right\rangle = (n+1)\tau_1 = \frac{n+1}{n} \log(n-1)|z_2|^2.$$

Hence by (24),

$$\psi_P = (n+1) \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log\left(\frac{n}{n-1} + \frac{|z_1|^2}{(n-1)^{1/n}|z_2|^{2/n}}\right) \quad \text{for } z \in \mathcal{R}_{F'}.$$

(Note that  $\psi_P$  has constant rank 1 on  $\mathcal{R}_{F'}$  as indicated by Theorem 1.3(iii), since  $\psi_P = (n+1)g^*\omega_{\mathbb{C}\mathbb{P}^1}$  on  $\mathcal{R}_{F'}$ , where  $g$  is the multi-valued holomorphic map to  $\mathbb{C}\mathbb{P}^1$  given by  $g(z_1, z_2) = (c_n z_1, z_2^{1/n})$ .)

By Theorem 1.3(iii), we know that  $\psi_P = 0$  on  $\mathcal{R}_v$ . To complete the description of  $\psi_P$ , it remains to describe the regions  $\mathcal{R}_{F'}$  and  $\mathcal{R}_v$ . We note that a forbidden point  $z$  lies in  $\mathcal{R}_{F'}$  if and only if  $c > 1$ . By (28), this is equivalent to  $\tilde{s}_1 > \frac{1}{n-1}$ , or

$$|z_2|^2 < (n-1)^{n-1}|z_1|^{2n}.$$

Therefore

$$(29) \quad \mathcal{R}_{F'} = \left\{ (z_1, z_2) : \frac{1}{n-1} \leq |z_2|^2 < (n-1)^{n-1}|z_1|^{2n}, \quad |z_1|^2 > \frac{1}{n-1} \right\}.$$

This leaves us with

$$(30) \quad \mathcal{R}_v = \left\{ (z_1, z_2) : |z_2|^2 \geq (n-1)^{n-1}|z_1|^{2n}, \quad |z_1|^2 \geq \frac{1}{n-1} \right\}.$$

These subregions are illustrated in Figure 4. To summarize:

$$\psi_P = \begin{cases} (n+1) \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + |z_1|^2 + |z_2|^2) = (n+1) \omega_{\text{FS}} & \text{for } z \in \mathcal{A}_P \\ & \text{(see (25)),} \\ n \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + |z_1|^2) & \text{for } z \in \mathcal{R}_F^\circ \\ & \text{(see (26)),} \\ (n+1) \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log\left(\frac{n}{n-1} + \frac{|z_1|^2}{(n-1)^{1/n} |z_2|^{2/n}}\right) & \text{for } z \in \mathcal{R}_{F'}^\circ \\ & \text{(see (29)),} \\ 0 & \text{for } z \in \mathcal{R}_v^\circ \\ & \text{(see (30)).} \end{cases}$$

Thus, as stated in Theorem 1.3,  $\psi_P$  has constant rank 1 on the flow-outs of the one-dimensional faces  $F$ ,  $F'$  and vanishes on the flow-out of the vertex  $v$ .

**1.3. Methods of proof.** Having stated our principal results, we now briefly outline some key ideas in the proofs. The key result is Theorem 1.2 on the mass of polynomials with Newton polytope  $P$ . To prove it, we begin with an easy formula

$$(31) \quad \mathbf{E}_{\nu_{NP}}(|f(z)|_{\text{FS}}^2) = \frac{1}{\#(NP)} \mathbf{E}_{|NP}(|f(z)|_{\text{FS}}^2) = \frac{1}{\#(NP)} \Pi_{|NP}(z, z),$$

(see §4) relating expected mass to the *conditional Szegő kernel*  $\Pi_{|NP}$ , i.e., the orthogonal projection onto  $\text{Poly}(NP)$ . In general, the term ‘Szegő kernel’ of a space  $\mathcal{S}$  of  $\mathcal{L}^2$  functions refers to the kernel for the orthogonal projection to  $\mathcal{S}$  from the space of all  $\mathcal{L}^2$  functions; i.e., it is of the form  $\Pi(x, y) = \sum_j s_j(x) \overline{s_j(y)}$ , where  $\{s_j\}$  is an orthonormal basis of  $\mathcal{S}$ . Precise asymptotics for  $\frac{1}{\#(NP)} \Pi_{|NP}(z, z)$  are given in Theorem 4.1. Note that the normalizing factor  $\frac{1}{\#(NP)}$  is straightforward to evaluate since

$$\#(NP) = \dim \text{Poly}(NP) = \text{Vol}(P) N^m + \dots$$

is the Ehrhart polynomial [Eh].

To analyze the conditional Szegő kernels  $\Pi_{|NP}(z, w)$ , we introduce the *polytope characters*

$$(32) \quad \chi_{NP}(e^{i\varphi}) := \sum_{\alpha \in NP} e^{i\langle \alpha, \varphi \rangle}, \quad e^{i\varphi} = (e^{i\varphi_1}, \dots, e^{i\varphi_m}).$$

We observe that

$$(33) \quad \Pi_{|NP}(z, w) = \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} \Pi_{NP}(z, e^{i\varphi} \cdot w) \chi_{NP}(e^{i\varphi}) d\varphi,$$

where

$$(34) \quad \mathbf{T}^m = \{(\zeta_1, \dots, \zeta_m) \in (\mathbb{C}^*)^m : |\zeta_j| = 1, 1 \leq j \leq m\}$$

is the  $m$ -torus, and where  $\Pi_{NP}(z, w)$  is the Szegő kernel of  $\text{Poly}(NP\Sigma)$ , the space of all homogeneous polynomials of degree  $NP$ . To obtain asymptotics, we need to analyze the behavior of  $\chi_{NP}(e^w)$  as  $N \rightarrow \infty$ .

To do so, we use the Euler-MacLaurin formula of Khovanskii-Pukhlikov [KP], Brion-Vergne [BV1, BV2], and Guillemin [Gu]:

$$(35) \quad \chi_P(e^w) = \text{Todd}(\mathcal{F}_P, \partial/\partial h) \left( \int_{P(h)} e^{\langle w, x \rangle} dx \right) \Big|_{h=0} \quad (w \in \mathbb{C}^m, \|w\| < \varepsilon),$$

where  $P(h)$  is of the form  $\{x : \langle u_j, x \rangle + \lambda_j + h_j \geq 0, 1 \leq j \leq n\}$  ( $P(0) = P$ ) and where  $\text{Todd}(\mathcal{F}_P, \partial/\partial h)$  is a certain infinite order differential *Todd operator*. Upon dilating the polytope, one obtains

$$(36) \quad \chi_{NP}(e^w) = N^m \text{Todd}(\mathcal{F}_P, N^{-1}\partial/\partial h) \left( \int_{P(h)} e^{N\langle w, x \rangle} dx \right) \Big|_{h=0} \quad (\|w\| < \varepsilon).$$

In Proposition 3.1, we show that the family  $\chi_{NP}$  is a complex oscillatory integral of the form

$$(37) \quad \chi_{NP}(e^w) = \int_P e^{N\langle w, x \rangle} [A_0(x, w)N^n + A_1(x, w)N^{n-1} + \cdots + A_n(x, w)] d\text{Vol}_n(x),$$

for  $\Im w$  sufficiently small (but  $\Re w$  arbitrary), where the  $A_l$  are analytic functions that are holomorphic in  $w$  and algebraic in  $x$ .

We then substitute this expression in (33) and use the method of stationary phase for complex oscillatory integrals [Hö, Ch. 7] to obtain the asymptotics of  $\Pi_{|NP}(z, z)$ . For the case where  $z$  is in the classically allowed region, we easily find that the critical point of the phase is given by  $\varphi = 0$  and  $x = p\mu(z)$ . Since the phase vanishes and has nondegenerate Hessian at the critical point, we immediately obtain an asymptotic expansion. The case where  $z$  is in the classically forbidden region is more subtle. Since  $p\mu(z)$  lies outside of  $P$  for this case, the phase has no critical points. To complete the analysis, we must deform the contour to pick up critical points. In particular, we consider the complexification  $(\mathbb{C}^*)^m$  of  $\mathbf{T}^m$  and deform  $\mathbf{T}^m$  to a contour of the form  $(\log|\zeta_1|, \dots, \log|\zeta_m|) = \tau \in \mathbb{R}^m$ , on which the phase has a ‘critical point’ with  $\varphi = 0$  and  $x = x_c \in \partial P$ . To be precise, the derivative tangential to the face of  $P$  containing  $x_c$  vanishes at  $x_c$  (while the normal derivative is nonvanishing), and furthermore the phase takes its maximal real part at  $x_c$ . Indeed,  $\tau$  is the unique vector  $\tau_z$  and  $x_c$  is the unique point  $q(z) \in \partial P$  used in formula (16) for the decay rate  $b_P(z)$ , and the maximal real part of the phase is  $-b_P(z)$ . We then obtain an asymptotic expansion from the method of stationary phase on domains with boundary.

**1.4. Brief outline and remarks.** To assist the reader in navigating this article, we give a brief outline: We begin in §2 by reviewing some notation and terminology involving convex polytopes and Szegő kernels for spaces of polynomials. Section 3 states and proves our oscillatory integral formula for the characters  $\chi_{NP}(e^w)$ . The heart of the paper is §4, where we derive the diagonal asymptotics of the Szegő kernel (Theorem 4.1) for the spaces  $\text{Poly}(NP)$ , using our character formula from §3 as one of the ingredients. (Theorem 1.2 is an immediate consequence of Theorem 4.1.) Then applying our Szegő kernel asymptotics, we prove Theorems 1.1, 1.3 and 1.4 on the distribution of zeros in §5. We also include an appendix (§6), where we describe an alternative geometric approach to our results using toric varieties. An index of notation is included at the end of the paper.



Before embarking on the proofs, we comment on our borrowed terminology from physics and on the role of toric geometry:

1.4.1. *Tunnelling of zeros.* Some of the terminology we use—allowed and forbidden regions, mass density—is taken from the semiclassical analysis of ground states of Schrödinger operators  $H_{\hbar} = -\hbar^2 \Delta + V$  on  $\mathbb{R}^n$ . The well-known ‘Agmon estimates’ of ground states (cf. [Ag]) show that  $\mathcal{L}^2$ -normalized ground states or low-lying eigenfunctions  $H_{\hbar}\varphi = E\varphi$  of  $H$  are concentrated as  $\hbar \rightarrow 0$  in the classically allowed region  $C_E = \{x \in \mathbb{R}^n : V(x) \leq E\}$  and  $|\varphi(x)|^2$  decays exponentially as  $\hbar \rightarrow 0$  for points  $x$  in the complement. In our setting, the Hamiltonian is  $\bar{\partial}^* \bar{\partial}$  on  $\mathcal{L}^2$ -sections of powers of the hyperplane section bundle on  $\mathbb{C}\mathbb{P}^m$ ,  $\hbar = 1/N$ , and the ground states are the holomorphic sections. Replacing  $|\varphi(x)|^2$  is the expected mass density  $|f(z)|_{\text{FS}}^2$  of the polynomials in this subspace. Theorems 1.1–1.2 show that the Newton-polytope constraint on the polynomials creates a ‘tunnelling theory’ for zeros and mass.

The decay function  $b_P$  in Theorem 1.2 is analogous to the Agmon action to the allowed region [Ag]. We will also obtain (Proposition 5.3) the integral formula

$$(38) \quad b_P(z) = \int_0^{\tau_z} \left[ -q(e^{-\sigma/2} \cdot z) + p\mu(e^{-\sigma/2} \cdot z) \right] \cdot d\sigma$$

(where  $\int_0^{\tau_z}$  denotes the integral over any path in  $\mathbb{R}^m$  from 0 to  $\tau_z$ ), which could be interpreted as an action, thereby bringing the results closer to classical Agmon estimates.

1.4.2. *Role of toric varieties.* Since  $\text{Poly}(NP)$  is naturally isomorphic to the space  $H^0(M_P, L_P)$  of holomorphic sections of a natural line bundle  $L_P$  over a toric variety  $M_P$  associated to  $P$  (see §6), the reader may wonder whether toric varieties play any role in this paper. The answer is that neither the statements nor proofs of our results involve toric varieties in any essential way. However, the theory of these varieties does give an alternative approach to the asymptotics of the conditional Szegő kernel, as will be explained in §6. It was also the approach in the original version of this paper [SZ2] and motivated some of the ideas.

The reader may also wonder how the results of this paper would change if, instead of defining the Gaussian measures  $\gamma_{1P}$  to be the conditional measures of one fixed ensemble, we defined the measures on  $\text{Poly}(NP) \simeq H^0(M_P, L_P^N)$  to be the Gaussian measures  $\gamma_N^{M_P}$  induced by the  $\mathcal{L}^2$  inner product induced by a Hermitian metric on  $L_P$  and its curvature form  $\omega_{M_P}$ . As a special case of our results in [SZ1, Prop. 4.4], one obtains (under the added assumption that  $M_P$  is smooth or maybe has orbifold singularities)

$$(39) \quad \frac{1}{N^k} \mathbf{E}_{\gamma_N^{M_P}}(Z_{f_1, \dots, f_k}) = \omega_{M_P}^k + O\left(\frac{1}{N}\right) = \left( \sum_{j=1}^m dI_j \wedge d\theta_j \right)^k + O\left(\frac{1}{N}\right),$$

where  $I_j, \theta_j$  are the action-angle variables of the moment map  $\mu_P : M_P \rightarrow \mathbb{R}^m$  of the  $\mathbf{T}^m$  action on  $M_P$ . Related results in a somewhat different set-up have also been obtained by Malajovich and Rojas [MR]. This quite different law shows that the measures  $\gamma_N^{M_P}$  are singular relative to  $\gamma_{N|NP}$  in the limit as  $N \rightarrow \infty$ .

As mentioned previously, the polytope  $P$  in this article is only used as a constraint on the polynomials in creating conditional measures. The norms of the monomials  $z^\alpha$  are fixed (as their  $\mathbb{C}\mathbb{P}^m$  norms). Thus, the change in the distribution of zeros as  $P$  varies is due solely to the choice of which monomials occur in the

polynomials. In the case of  $\gamma_N^{M_P}$ , the norms of the monomials vary as  $P$  varies since they are  $\mathcal{L}^2$ -normalized on the toric variety  $M_P$ . Hence the variances of the coefficients of a polynomial in the  $\gamma_N^{M_P}$  ensembles depend on the choice of metric on  $L_P$ . This dependence creates complicated biases towards some monomials and away from others as  $P$  varies, making it difficult to understand what a comparison between the  $\gamma_N^{M_P}$  ensembles would be measuring.

In fact, our results using conditional measures also apply to polytopes that are not convex and hence do not correspond to any toric variety. Indeed, suppose that  $P$  is a nonconvex ‘lattice’ polytope. Then Theorems 1.1–1.4 hold with the following modifications: The point  $q(z)$  satisfying (14)–(15) is not always unique; instead we choose  $q(z)$  to minimize  $b_P(z)$ . Then the function  $b_P(z)$  is  $\mathcal{C}^0$ , not  $\mathcal{C}^1$ , and  $\psi_P$  is a positive current, which has singular support. Part (iii) of Theorem 1.3 applies only to the absolutely continuous part of  $\psi_P$ . More significantly, the limit measure in Theorem 1.1 does not vanish on the forbidden region; instead, it is orthogonal to volume measure there. This observation is relevant to the theory of ‘fewnomial’ systems, which we discuss in a forthcoming paper [SZ4].

1.4.3. *Acknowledgements.* Our interest in polynomials with a fixed Newton polytope was in part stimulated by a discussion with A. Varchenko at the outset of this work. We would like to thank M. Brion for many helpful comments regarding polytopes and the Euler-MacLaurin formula. We are also grateful to T. Theobald for giving us permission to use Figure 7 from his paper and to A. Carass for providing the illustrations in Figures 1, 3, 4, 5, and 6.

## 2. BACKGROUND

2.1. **Fans.** By a *convex integral polytope*, we mean the convex hull in  $\mathbb{R}^m$  of a finite set in the lattice  $\mathbb{Z}^m$ . A convex integral polytope  $P$  with nonempty interior  $P^\circ$  can be defined by linear equations

$$(40) \quad \ell_j(x) := \langle x, u_j \rangle + \lambda_j \geq 0 \quad (j = 1, \dots, d),$$

where  $u_j \in \mathbb{Z}^m$  is the primitive inward-pointing normal to the  $j$ -th *facet* (codimension-one face)

$$F_j^{m-1} := \{x \in P : \ell_j(x) = 0, \ell_k(x) > 0 \text{ for } k \neq j\}.$$

For each point  $x \in P$ , we consider the *normal cone to  $P$  at  $x$* ,

$$(41) \quad C_x = C_x^P := \{u \in \mathbb{R}^m : \langle u, x \rangle = \sup_{y \in P} \langle u, y \rangle\},$$

which is a closed convex polyhedral cone. We decompose  $P$  into a finite union of *faces*, each face being an equivalence class under the equivalence relation  $x \sim y \iff C_x = C_y$ . For each face  $F$ , we let  $C_F$  denote the normal cone of the points of  $F$ . **Note that by our convention, the faces are disjoint sets.** We shall use the term *closed face* to refer to the closure of a face of  $P$ .

Each face of dimension  $r$  ( $0 \leq r \leq m$ ) is an open polytope in an  $r$ -plane in  $\mathbb{R}^m$ ; i.e., the 0-dimensional faces are the vertices of  $P$ , the 1-dimensional faces are the edges with their end points removed, and so forth. The facets  $F_j = F_j^{m-1}$  and their normal cones are given by

$$\bar{F}_j = \{x \in P : \ell_j(x) = 0\}, \quad C_{F_j} = \{-tu_j : t \geq 0\}.$$

The  $m$ -dimensional face is the interior  $P^\circ$  of the polytope with normal cone  $C_{P^\circ} = \{0\}$ .

For each  $x \in P$ , we let

$$(42) \quad \mathcal{J}(x) = \{j \in \mathbb{Z} : 1 \leq j \leq d, \ell_j(x) = 0\} .$$

One easily sees that  $C_x = C_y \iff \mathcal{J}(x) = \mathcal{J}(y)$ , and hence we can write  $\mathcal{J}(F) = \mathcal{J}(x)$ , where  $x$  lies in the face  $F$ . In particular,  $\mathcal{J}(F_j^{m-1}) = \{j\}$  and  $\mathcal{J}(P^\circ) = \emptyset$ . The polytope  $P$  is called *simple* if  $\#\mathcal{J}(v) = m$  for each vertex  $v$  of  $P$ . In this case,  $\{u_j : j \in \mathcal{J}(v)\}$  is a basis for  $\mathbb{R}^m$  for each vertex  $v$  (since  $P^\circ \neq \emptyset$ ), and furthermore  $\#\mathcal{J}(F) = \text{codim } F$  for all faces  $F$ . The polytope  $P$  is said to be *Delzant* if  $\{u_j : j \in \mathcal{J}(v)\}$  generates the lattice  $\mathbb{Z}^m$  for each vertex  $v$  of  $P$ .

The convex integral polytope  $P$  determines the fan  $\mathcal{F}_P := \{C_F : F \text{ is a face of } P\}$ . A *fan*  $\mathcal{F}$  in  $\mathbb{R}^m$  is a collection of closed convex rational polyhedral cones such that a closed face of a cone in  $\mathcal{F}$  is an element of  $\mathcal{F}$  and the intersection of two cones in  $\mathcal{F}$  is a closed face of each of them. (Fans are used in the algebraic construction of toric varieties; see [Fu, §1.4].) An example of a convex integral polytope and its fan is given in Figure 5.

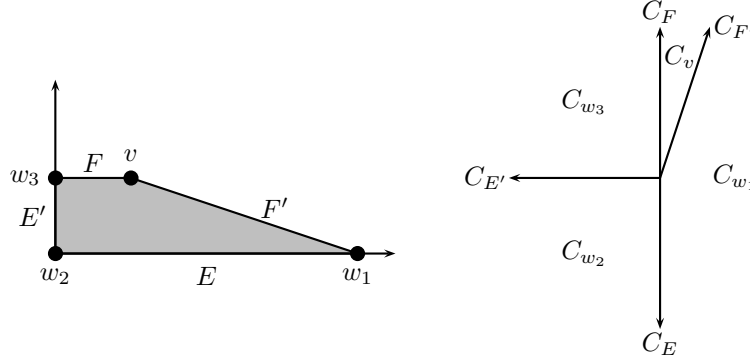


FIGURE 5. A convex polytope and its fan

We shall also consider convex integral polytopes  $P$  with empty interior. In this case, the faces of  $P$  and the normal cones  $C_x^P$  are defined exactly as above. If  $P^\circ = \emptyset$ , then the normal cones of  $P$  all contain the linear subspace of  $\mathbb{R}^m$  orthogonal to  $P$  (contrary to the case  $P^\circ \neq \emptyset$ , where the normal cones are *pointed*, i.e., do not contain any lines in  $\mathbb{R}^m$ ). A convex polytope in  $\mathbb{R}^m$  of dimension  $n < m$  is also said to be *simple* if it is linearly isomorphic to a simple polytope in  $\mathbb{R}^n$  or, equivalently, if every vertex is the intersection of precisely  $n$  closed faces of dimension  $n - 1$ .

**2.2. Szegő kernels.** By homogenizing the polynomials in  $\text{Poly}(P)$ , we obtain a finite dimensional subspace  $\mathcal{S}$  of  $\mathcal{L}^2(S^{2m+1})$ . By the ‘Szegő projector,’ we mean the orthogonal projection  $\Pi_{\mathcal{S}} : \mathcal{L}^2(X) \rightarrow \mathcal{S}$ . We recall that the projector  $\Pi_{\mathcal{S}}$  is given by a kernel of the form

$$(43) \quad \Pi_{\mathcal{S}}(x, y) = \sum_{j=1}^{k_N} s_j^N(z) \overline{s_j^N(y)} ,$$

where  $\{s_j\}$  is an orthonormal basis of  $\mathcal{S}$ .

We now describe two sequences of Szegő kernels, which play a crucial role in our main results:

2.2.1. *The projective Szegő kernels.* As a first example, we recall that  $\text{Poly}(p\Sigma)$  can be identified with the space of degree- $p$  homogeneous polynomials in  $m+1$  variables by identifying the (nonhomogeneous) polynomial

$$f(z_1, \dots, z_m) = \sum_{|\alpha| \leq p} c_\alpha z^\alpha \quad (z^\alpha = z_1^{\alpha_1} \dots z_m^{\alpha_m})$$

with the homogeneous polynomial

$$F(\zeta_0, \dots, \zeta_m) = \sum_{|\alpha| \leq p} c_\alpha \zeta_0^{p-|\alpha|} \zeta_1^{\alpha_1} \dots \zeta_m^{\alpha_m} .$$

We equip the space  $\text{Poly}(p\Sigma)$  with the  $\mathcal{L}^2$  inner product on  $S^{2m+1}$ :

$$(44) \quad \langle f, \bar{g} \rangle = \frac{1}{m!} \int_{S^{2m+1}} F \bar{G} d\nu = \frac{1}{m!} \int_{\mathbb{C}^m} \frac{f(z) \overline{g(z)}}{(1 + \|z\|^2)^p} \omega_{\text{FS}}^m(z), \quad f, g \in \text{Poly}(p\Sigma),$$

where

$$(45) \quad \omega_{\text{FS}} = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + \|z\|^2)$$

is the Fubini-Study Kähler form on  $\mathbb{C}^m \subset \mathbb{C}\mathbb{P}^m$ .

A basis for  $\text{Poly}(p\Sigma)$  consists of the monomials  $\chi_\alpha(z) = z_1^{\alpha_1} \dots z_m^{\alpha_m}$ ,  $|\alpha| \leq p$ . The monomials  $\{\chi_\alpha\}$  are orthogonal but not normalized. Their  $\mathcal{L}^2$  norms given by the inner product (44) are

$$(46) \quad \|\chi_\alpha\| = \left[ \frac{(p - |\alpha|)! \alpha_1! \dots \alpha_m!}{(p + m)!} \right]^{\frac{1}{2}} .$$

Note that the norm  $\|\cdot\|$  depends on  $p$ . Thus we have an orthonormal basis for  $\text{Poly}(p\Sigma)$  given by the monomials

$$(47) \quad \frac{1}{\|\chi_\alpha\|} \chi_\alpha = \left[ \frac{(p + m)!}{(p - |\alpha|)! \alpha_1! \dots \alpha_m!} \right]^{\frac{1}{2}} \chi_\alpha = \sqrt{\frac{(p + m)!}{p!} \binom{p}{\alpha}} \chi_\alpha, \quad |\alpha| \leq p,$$

where

$$(48) \quad \binom{p}{\alpha} = \frac{p!}{(p - |\alpha|)! \alpha_1! \dots \alpha_m!} .$$

We let  $\widehat{\chi}_\alpha^p : S^{2m+1} \rightarrow \mathbb{C}$  denote the homogenization of  $\chi_\alpha$ :

$$(49) \quad \widehat{\chi}_\alpha^p(x) = x_0^{p-|\alpha|} x_1^{\alpha_1} \dots x_m^{\alpha_m} .$$

Hence the Szegő kernel  $\Pi_p$  for the orthogonal projection to  $\text{Poly}(p\Sigma)$  is given by

$$(50) \quad \Pi_p(x, y) = \sum_{|\alpha| \leq p} \frac{1}{\|\chi_\alpha\|^2} \widehat{\chi}_\alpha(x) \overline{\widehat{\chi}_\alpha(y)} = \frac{(p + m)!}{p!} \langle x, \bar{y} \rangle^p ,$$

for  $x, y \in S^{2m+1}$ . (The sum  $\sum_{p=0}^\infty \Pi_p$  is the usual Szegő kernel for the sphere; see [BSZ, §1.3.1].)

We also identify the point  $z \in (\mathbb{C}^*)^m$  with the lift  $x = \frac{1}{(1 + \|z\|^2)^{1/2}} (1, z_1, \dots, z_m) \in S^{2m+1}$ , and we write

$$(51) \quad \widehat{\chi}_\alpha^p(z) = \frac{z^\alpha}{(1 + \|z\|^2)^{p/2}} .$$

The Szegő kernel can then be written explicitly as

$$(52) \quad \begin{aligned} \Pi_p(z, w) &= \frac{(p+m)!}{p!} \frac{\sum_{|\alpha| \leq p} \binom{p}{\alpha} z^\alpha \bar{w}^\alpha}{(1+\|z\|^2)^{p/2} (1+\|w\|^2)^{p/2}} \\ &= \frac{(p+m)!}{p!} \left[ \frac{1 + \langle z, \bar{w} \rangle}{(1+\|z\|^2)^{1/2} (1+\|w\|^2)^{1/2}} \right]^p. \end{aligned}$$

2.2.2. *The conditional Szegő kernels associated to a polytope  $P$ .* In this case, the relevant space of polynomials is the subspace  $\text{Poly}(P) \subset \text{Poly}(p\Sigma)$  of polynomials with Newton polytope  $P$ . Here, we may choose any  $p \geq \deg P := \max\{|\alpha| : \alpha \in P\}$ , but we normally choose  $p = \deg P$ . In the conditional Szegő kernel, we retain the Fubini-Study inner product on this subspace. Hence this example is very similar to the previous one. The main difference is that an orthonormal basis of  $\text{Poly}(P)$  is given by

$$\left\{ \frac{1}{\|\chi_\alpha\|} \chi_\alpha : \alpha \in P \right\}.$$

**Definition.** The conditional Szegő kernel  $\Pi|_P$  is the kernel for the orthogonal projection to  $\text{Poly}(P)$  with respect to the induced Fubini-Study inner product:

$$(53) \quad \Pi|_P(x, y) = \sum_{\alpha \in P} \frac{1}{\|\chi_\alpha\|^2} \widehat{\chi}_\alpha^p(x) \overline{\widehat{\chi}_\alpha^p(y)}.$$

When defining the term ‘random polynomial with fixed Newton polytope  $P$ ’, we wish to use an  $\mathcal{L}^2$ -norm on monomials which is defined independently of  $P$ . This explains why the conditional Szegő kernel is the essential one in our problem. The conditional Szegő kernel can be written explicitly on  $\mathbb{C}^m$  as

$$(54) \quad \Pi|_P(z, w) = \frac{(p+m)!}{p!} \frac{\sum_{\alpha \in P} \binom{p}{\alpha} z^\alpha \bar{w}^\alpha}{(1+\|z\|^2)^{p/2} (1+\|w\|^2)^{p/2}}.$$

### 3. POLYTOPE CHARACTER

This section is devoted to the asymptotic analysis of the polytope characters (32). Our interest is in the asymptotics of the ‘ray’ of characters  $\chi_{NP}(e^w)$  defined on  $(\mathbb{C}^*)^m$  by

$$(55) \quad \chi_{NP}(e^w) = \sum_{\alpha \in NP} e^{\langle w, \alpha \rangle}, \quad w \in \mathbb{C}^m.$$

We shall derive a formula expressing the character  $\chi_{NP}(e^w)$  as an oscillatory integral over the original polytope  $P$  with the same phase  $\langle w, x \rangle$  as in formula (35) of [KP, BV1, Gu], which is the starting point of our derivation. Although (35) holds only for small  $w$ , our formula holds by analytic extension for all points  $w$  in a ‘strip’ of the form

$$(56) \quad S(\varepsilon) := \{w \in \mathbb{C}^m : |\Im w_j| < \varepsilon \text{ for } 1 \leq j \leq m\}.$$

Our integral formula for  $\chi_{NP}(e^w)$  is given by the following proposition, which is also of independent interest.

**Proposition 3.1.** *Let  $P$  be a simple integral polytope in  $\mathbb{R}^m$  of dimension  $n$  ( $1 \leq n \leq m$ ). Then there exists  $\varepsilon > 0$  such that the characters  $\chi_{NP}$  can be given as integrals over the polytope  $P$  of the form*

$$\chi_{NP}(e^w) = \int_P e^{N\langle w, x \rangle} [A_0(x, w)N^n + A_1(x, w)N^{n-1} + \cdots + A_n(x, w)] d\text{Vol}_n(x),$$

for all  $w \in S(\varepsilon)$ ,

where the  $A_l$  are analytic functions on  $P \times S(\varepsilon)$  that are holomorphic in  $w$  and algebraic in  $x$ , and

- (i)  $A_l(x, w) \in \mathbb{R}$  whenever  $w \in \mathbb{R}^m$ ,
- (ii)  $A_0(x, 0) = 1$ ,
- (iii)  $A_0(x, w) \in \mathbb{R}^+$  whenever  $w$  is in the normal cone to  $P$  at  $x$ , i.e., whenever  $w \in C_x \subset \mathbb{R}^m$ .

As discussed in the introduction, Proposition 3.1 will later be used to obtain asymptotics of the conditional Szegő kernel through formula (33). When  $w = 0$ , the character (55) equals the number of lattice points of  $NP$  and Proposition 3.1 reduces to the Ehrhart formula [Eh]:

$$(57) \quad \#(NP) = \sum_{j=0}^n a_j N^{n-j}, \quad n = \dim P, \quad a_0 = \text{Vol}_n(P).$$

We note that one can easily extend (57) to an arbitrary convex integral polytope  $P$  (not necessarily simple) by triangulating  $P$  into  $n$ -simplices  $\{\Delta_1, \dots, \Delta_k\}$  whose vertices are vertices of  $P$  (see [GKZ, pp. 214–216] for an elementary proof of this fact), and then applying the inclusion-exclusion principle.

*Remark.* The characters  $\chi_{NP}(e^w)$  have a natural interpretation in toric geometry, and  $\chi_{NP}(1) = \#(NP)$  can also be given by the Riemann-Roch formula on the toric variety associated to the polytope  $P$ . See §6 for further details and references.

**3.1. Proof of Proposition 3.1.** We first reduce to the case where  $P$  has nonempty interior: If  $\dim P = n < m$ , then by a translation we may assume that  $0 \in P$ . Next by an orthogonal projection to the  $n$ -plane  $\langle P \rangle$  spanned by  $P$ , we may assume that  $w \in \langle P \rangle$ . Since  $\langle P \rangle \cap \mathbb{Z}^m \approx \mathbb{Z}^n$ , we may assume that  $\langle P \rangle = \mathbb{R}^n$ . Hence we may assume that  $n = m$ .

Recalling (40), we express  $P$  in the form

$$(58) \quad P = \{x : \langle x, u_j \rangle + \lambda_j \geq 0, \quad j = 1, \dots, d\},$$

where the vectors  $u_j$  are the inward-pointing normal vectors to the facets of  $P$ . The normals  $u_j$  are normalized so that  $u_j$  is a primitive element of  $\mathbb{Z}^m$  (i.e.,  $ru_j \in \mathbb{Z}^m \iff r \in \mathbb{Z}$ ).

We shall use the formula given independently by Brion and Vergne [BV1, Theorem 3.12] and by Guillemin [Gu, Theorem 4.1]:

$$(59) \quad \chi_P(e^w) = \text{Todd}(\mathcal{F}, \partial/\partial h) \left( \int_{P(h)} e^{\langle w, x \rangle} dx \right) \Big|_{h=0}, \quad w \in \mathbb{C}^m, \quad \|w\| < \varepsilon_{\mathcal{F}}.$$

Here,

$$(60) \quad P(h) = \{x : \langle u_j, x \rangle + \lambda_j + h_j \geq 0, \quad 1 \leq j \leq d\},$$

$\mathcal{F} = \mathcal{F}_P$  is the fan associated to  $P$ , and  $\varepsilon_{\mathcal{F}}$  is a positive constant. (E.g., if  $P$  is Delzant, then  $\varepsilon_{\mathcal{F}} = 2\pi$ .) Also,  $\text{Todd}(\mathcal{F}, \partial/\partial h)$  is the (generalized) Todd operator constructed by Brion and Vergne [BV1, pp. 374–376] as follows:

For each cone  $C_F \in \mathcal{F}$ , we consider the open  $k$ -parallelogram

$$U_F := \left\{ \sum_{j \in \mathcal{J}(F)} t_j u_j : 0 < t_j < 1 \right\},$$

where  $k = \text{codim } F = \dim C_F$ . (Recall (42).) We let  $U$  denote the (disjoint) union of the  $U_F$ , where  $F$  runs over the faces of  $P$  (including the open face  $P^\circ$ , where  $U_{P^\circ} = \{0\}$ ). Equivalently,  $U$  is a ‘bouquet of  $m$ -parallelograms’,

$$U = \bigcup_{\text{vertices } v} \left\{ \sum_{j \in \mathcal{J}(v)} t_j u_j : 0 \leq t_j < 1 \right\},$$

and is an open neighborhood of 0. We consider the finite set

$$\Gamma_{\mathcal{F}} := U \cap \mathbb{Z}^m.$$

(It is easy to see that  $\Gamma_{\mathcal{F}} = \{0\}$  if and only if the polytope  $P$  is Delzant.) For each lattice point  $\gamma \in \Gamma_{\mathcal{F}}$ , we define a unique vector  $g(\gamma) = (g_1(\gamma), \dots, g_m(\gamma)) \in ([0, 1) \cap \mathbb{Q})^m$  as follows: Let  $F$  be the face such that  $\gamma \in U_F$ ; then

$$\gamma = \sum_{j=1}^m g_j(\gamma) u_j, \quad g_j(\gamma) \neq 0 \iff j \in \mathcal{J}(F).$$

We note that the map  $\gamma \mapsto g(\gamma)$  is injective. The Todd operator is then given by

$$(61) \quad \text{Todd}(\mathcal{F}, \partial/\partial h) = \sum_{\gamma \in \Gamma_{\mathcal{F}}} \left[ \prod_{j=1}^d \text{Todd}(e^{i2\pi g_j(\gamma)}, \partial/\partial h_j) \right], \quad \text{Todd}(a, z) = \frac{z}{1 - ae^{-z}}.$$

As noted in [BV1] (Remark 2.16), if  $P$  is a Delzant polytope, then  $\text{Todd}(\mathcal{F}, \partial/\partial h)$  is the usual Todd operator

$$(62) \quad \text{Todd}(\partial/\partial h) = \prod_{j=1}^d \text{Todd}(\partial/\partial h_j), \quad \text{Todd}(z) = \text{Todd}(1, z) = z(1 - e^{-z})^{-1},$$

and formula (59) is due to Khovanskii and Pukhlikov [KP] for that case. (Formula (59) has a generalization to arbitrary convex rational polytopes; see [BV2].) Since the term in (61) corresponding to  $\gamma = 0$  is the usual Todd operator and the other terms vanish at the origin, the constant term in the series expansion of  $\text{Todd}(\mathcal{F}, \partial/\partial h)$  equals 1 for all simple polytopes. Furthermore, the coefficients of the expansion of  $\text{Todd}(\mathcal{F}, \partial/\partial h)$  are real since if we regard the  $g(\gamma)$  as elements of  $\mathbb{Q}^m/\mathbb{Z}^m$ , then the set  $\{g(\gamma)\}$  is a union of subgroups and hence is invariant under multiplication by  $-1$ .

Let us check how formula (59) dilates. The fan does not change when the polytope is dilated, so the Todd operator is the same for  $P$  and for  $NP$ . Noting that

$$(63) \quad (NP)(Nh) = \{x : \langle u_j, x \rangle + N\lambda_j + Nh_j \geq 0\} = N(P(h)),$$

we change variables  $h \mapsto Nh$ ,  $x \mapsto Nx$  in (59) to obtain

$$(64) \quad \chi_{NP}(e^w) = N^m \text{Todd}(\mathcal{F}, N^{-1}\partial/\partial h) \left( \int_{P(h)} e^{N\langle w, x \rangle} dx \right) \Big|_{h=0}, \quad \|w\| < \varepsilon_{\mathcal{F}}.$$

To prove Proposition 3.1, we use the following lemma to change variables so that the integral is over the fixed polytope  $P$  while the integrand depends on  $h$ .

We say that a map  $f : P \rightarrow \mathbb{R}^m$  is a *regular rational map* if it can be written in the form  $f = (p_1/q_1, \dots, p_m/q_m)$ , where the  $p_j$  and  $q_j$  are polynomials on  $\mathbb{R}^m$  and the  $q_j$  are nonvanishing on  $P$ .

**Lemma 3.2.** *Let  $P$  be a simple polytope in  $\mathbb{R}^m$  with nonempty interior. Then there exist regular rational maps  $f_j : P \rightarrow \mathbb{R}^m$ ,  $1 \leq j \leq d$ , such that for  $h \in \mathbb{R}^d$  sufficiently small, the map  $\gamma_h : P \rightarrow \mathbb{R}^m$  given by*

$$\gamma_h(x) = x + \sum_{j=1}^d h_j f_j(x)$$

*is a diffeomorphism from  $P$  onto  $P(h)$  mapping each face of  $P$  onto the corresponding face of  $P(h)$ .*

*Proof.* Recall that the assumption that  $P$  is simple means that there are precisely  $m$  hyperplanes  $\{x : \langle x, u_j \rangle + \lambda_j = 0\}$  through each vertex, and hence for small  $h$ , the faces of  $P(h)$  correspond to the faces of  $P$ .

We first construct an ‘algebraic partition of unity’  $\{\psi_k\}$  over  $P$  as follows. Let  $\{v^1, \dots, v^s\}$  denote the set of vertices of  $P$ . Recalling (40), we let  $\tilde{\psi}_k : P \rightarrow [0, +\infty)$  be given by

$$\tilde{\psi}_k = \prod \{\ell_j : j \notin \mathcal{J}(v^k)\}, \quad k = 1, \dots, s.$$

We then let

$$\psi_k = \frac{\tilde{\psi}_k}{\sum_{j=1}^s \tilde{\psi}_j}.$$

We easily see that  $0 \leq \psi_k \leq 1$ ,  $\sum_{k=1}^s \psi_k = 1$  and that  $\psi_k^{-1}(0)$  is the union of the closed facets of  $P$  that do not contain  $v^k$ .

The vertices  $v^k(h)$  of  $P(h)$  are of the form

$$v^k(h) = v^k + L(k)h, \quad 1 \leq k \leq s, \quad \|h\| < \varepsilon,$$

where  $L(k) : \mathbb{R}^d \rightarrow \mathbb{R}^m$  is a linear transformation. (In fact, it is easy to see that  $L(k)$  is given by first projecting onto the  $m$  coordinates corresponding to the facets incident to  $v^k$  and then applying an element of  $\text{GL}(m, \mathbb{R})$ .) Let

$$w^{kj} = L(k)E^j \in \mathbb{R}^m, \quad 1 \leq k \leq s, \quad 1 \leq j \leq d,$$

where  $\{E^j\}$  is the standard basis of  $\mathbb{R}^d$ , so that

$$(65) \quad v^k(h) = v^k + \sum_{j=1}^d h_j w^{kj}, \quad \|h\| < \varepsilon.$$

For each  $k$ , let  $v^{\nu(k,1)}, \dots, v^{\nu(k,m)}$  be the vertices connected to  $v^k$  by an edge (1-dimensional face) of  $P$ . For  $1 \leq k \leq s$ ,  $1 \leq j \leq d$ , let  $f_{kj} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  be the unique affine map such that

$$f_{kj}(v^l) = w^{lj} \quad \text{for } l = k, \nu(k,1), \dots, \nu(k,m).$$

So by (65), if  $l = k$  or if  $v^l$  is a vertex connected to  $v^k$  by an edge of  $P$ , then

$$(66) \quad v^l(h) = v^l + \sum_{j=1}^d h_j f_{kj}(v^l), \quad \text{for } l = k, \nu(k,1), \dots, \nu(k,m), \quad \|h\| < \varepsilon.$$



We claim that the maps

$$f_j := \sum_{k=1}^s \psi_k f_{kj}$$

satisfy the conclusions of the lemma. To verify the claim, we let

$$\gamma_h^k(x) = x + \sum_{j=1}^d h_j f_{kj}(x), \quad 1 \leq k \leq s,$$

so that

$$\gamma_h = \sum \psi_k \gamma_h^k.$$

It follows from (66) that when  $x$  is a vertex we have

$$\gamma_h(v^k) = \gamma_h^k(v^k) = v^k(h),$$

since  $\psi_l(v^k) = 0$  when  $k \neq l$ .

Next we consider points on the 1-skeleton of  $P$ . Suppose that  $x = tv^1 + (1-t)v^2$ , where  $0 < t < 1$  and  $v^1, v^2$  are joined by an edge of  $P$ . Then  $\psi_k(x) = 0$  for  $k > 2$  and we again conclude from (66) that  $\gamma_h^k(v^1) = v^1(h)$  and  $\gamma_h^k(v^2) = v^2(h)$  for  $k = 1, 2$ . Since the  $\gamma_h^k$  are affine maps, it follows that

$$\gamma_h(x) = tv^1(h) + (1-t)v^2(h)$$

so that  $\gamma_h$  takes each edge of  $P$  linearly onto the corresponding edge of  $P(h)$ .

In general we do not have linearity on faces of dimension  $> 1$ , so we need to modify the argument for the higher skeletons. Suppose now that  $x$  lies in a 2-dimensional face  $F$  with vertices  $v^1, \dots, v^r$  ( $r \geq 3$ ) and  $x = t_1v^1 + t_2v^2 + t_3v^3$ , where  $\min t_j \geq 0$  and  $\sum t_j = 1$ . Let  $1 \leq l \leq r$ . Since the points  $v^1(h), \dots, v^r(h)$  lie on a plane  $S_h$  and the  $\gamma_h^k$  are linear, it follows from the above argument that  $\gamma_h^k(v^l) \in S_h$  for  $1 \leq k \leq r$ . (For example, if  $v^2, v^3$  are each connected to  $v^1$  by an edge, then  $\gamma_h^1(v^l) = v^l(h)$  for  $l = 1, 2, 3$ , but for  $3 < l \leq r$ , we can conclude only that  $\gamma_h^1(v^l)$  lies on the plane through  $v^1(h), v^2(h), v^3(h)$ .) For  $k > r$ , there exists a closed facet containing  $v_1, v_2, v_3$  but not containing  $v_k$ , and hence  $\psi_k(x) = 0$ . Therefore,

$$\gamma_h(x) = \sum_{1 \leq j, k \leq r} t_j \psi_k(x) \gamma_h^k(v^j) \in S_h.$$

Furthermore, for  $h$  small,  $D\gamma_h \approx D\gamma_0 \equiv I$ , so  $\gamma_h|_F : F \rightarrow S_h$  is a local homeomorphism. Therefore,  $\partial_{S_h}(\gamma_h(F)) = \gamma_h(\partial F)$ , which is the boundary of the convex polygon  $F(h)$  with vertices  $v^1(h), \dots, v^r(h)$ . It follows that  $\gamma_h(F) = F(h)$  and hence  $\gamma_h|_{\overline{F}} : \overline{F} \rightarrow \overline{F(h)}$  is a (global) homeomorphism for sufficiently small  $h$ .

Continuing by induction on  $\dim F$ , we conclude that  $\gamma_h : \overline{F} \approx \overline{F(h)}$  for all faces  $F$  of  $P$ , and in particular,  $\gamma_h$  maps  $P$  diffeomorphically onto  $P(h)$ , for  $h$  small.  $\square$

*Completion of the proof of Proposition 3.1.* Let  $\gamma_h : P \approx P(h)$  be as in Lemma 3.2. Making the change of variables  $x \mapsto \gamma_h(x)$ , we have

$$\begin{aligned} \int_{P(h)} e^{N\langle w, x \rangle} dx &= \int_P e^{N\langle w, \gamma_h(x) \rangle} \det D\gamma_h(x) dx \\ (67) \qquad &= \int_P e^{N[\langle w, x \rangle + \sum_j \langle w, f_j(x) \rangle h_j]} G(x, h) dx, \end{aligned}$$

where  $G$  is a regular rational function on  $P \times \mathbb{R}$  of the form

$$G(x, h) = \sum_{\alpha \in \mathbb{N}^d, |\alpha| \leq m} G_\alpha(x) h^\alpha, \quad G_0 \equiv 1.$$

Hence by (64) and (67), we have  
(68)

$$\chi_{NP}(e^w) = N^m \int_P \text{Todd}(\mathcal{F}, N^{-1} \partial / \partial h) \left( e^{N[\langle w, x \rangle + \sum_j \langle w, f_j(x) \rangle h_j]} G(x, h) \right) \Big|_{h=0} dx, \quad \|w\| < \varepsilon.$$

To justify interchanging integration and Todd differentiation in the above and to simplify (68), we use the following identity from [KP, Lemma 1]): Suppose that  $G(h) \in \mathbb{C}[h_1, \dots, h_k]$  is a polynomial and  $F$  is a convergent power series about  $0 \in \mathbb{C}^k$  with domain of convergence  $\Omega$ . Then

$$(69) \quad F(\partial / \partial h) \left[ G(h) e^{\sum_{j=1}^k q_j h_j} \right] = G(\partial / \partial q) \left[ F(q) e^{\sum_{j=1}^k q_j h_j} \right], \quad q \in \Omega, \quad h \in \mathbb{C}^k.$$

(We verify (69) by noting that both sides equal  $F(\partial / \partial h) G(\partial / \partial q) \exp(\sum_{j=1}^k q_j h_j)$ .)

We apply (69) with  $F(q) = \text{Todd}(\mathcal{F}, N^{-1} q)$ ,  $G(q) = G(x, q)$  to the integrand of (68):

$$\begin{aligned} & \text{Todd}(\mathcal{F}, N^{-1} \partial / \partial h) \left[ e^{N[\langle w, x \rangle + \sum_j \langle w, f_j(x) \rangle h_j]} G(x, h) \right] \Big|_{h=0} \\ &= e^{N\langle w, x \rangle} \text{Todd}(\mathcal{F}, N^{-1} \partial / \partial h) \left[ G(x, h) e^{\sum_{j=1}^d q_j h_j} \right] \Big|_{h=0, q_j = N\langle w, f_j(x) \rangle} \\ &= e^{N\langle w, x \rangle} G(x, \partial / \partial q) \left[ \text{Todd}(\mathcal{F}, N^{-1} q) e^{\sum_{j=1}^d q_j h_j} \right] \Big|_{h=0, q_j = N\langle w, f_j(x) \rangle} \\ &= e^{N\langle w, x \rangle} G(x, \partial / \partial q) \text{Todd}(\mathcal{F}, N^{-1} q) \Big|_{q_j = N\langle w, f_j(x) \rangle} \\ (70) \quad &= e^{N\langle w, x \rangle} G(x, N^{-1} \partial / \partial \tilde{q}) \text{Todd}(\mathcal{F}, \tilde{q}) \Big|_{\tilde{q}_j = \langle w, f_j(x) \rangle}. \end{aligned}$$

In particular, if we let  $\text{Todd}_k$  denote the sum of the terms of the Todd series of degree  $\leq k$ , then

$$\begin{aligned} & \text{Todd}_k(\mathcal{F}, N^{-1} \partial / \partial h) \left[ e^{N[\langle w, x \rangle + \sum_j \langle w, f_j(x) \rangle h_j]} G(x, h) \right] \Big|_{h=0} \\ &= e^{N\langle w, x \rangle} G(x, N^{-1} \partial / \partial \tilde{q}) \text{Todd}_k(\mathcal{F}, \tilde{q}) \Big|_{\tilde{q}_j = \langle w, f_j(x) \rangle} \\ &\rightarrow \text{Todd}(\mathcal{F}, N^{-1} \partial / \partial h) \left[ e^{N[\langle w, x \rangle + \sum_j \langle w, f_j(x) \rangle h_j]} G(x, h) \right] \Big|_{h=0} \end{aligned}$$

uniformly for  $x \in P$  and  $w$  sufficiently small, which justifies the interchange of integration and Todd differentiation in (68).

Let  $\Lambda(x)$  denote the  $d \times m$  matrix whose rows are the  $f_j$ ; i.e.,  $\Lambda(x) : \mathbb{R}^m \rightarrow \mathbb{R}^d$  is the linear map given by  $(\Lambda(x)y)_j = \langle y, f_j(x) \rangle$ . By (70)

$$(71) \quad \begin{aligned} & \text{Todd}(\mathcal{F}, N^{-1} \partial / \partial h) \left[ e^{N[\langle w, x \rangle + \sum_j \langle w, f_j(x) \rangle h_j]} G(x, h) \right] \Big|_{h=0} \\ &= e^{N\langle w, x \rangle} \sum_{l=0}^m N^{-l} T_l(x, \Lambda(x)w), \end{aligned}$$

where  $T_l(x, q)$  is a convergent power series in  $q$  with coefficients that are regular rational functions on  $P$ . More specifically,

$$T_0(x, q) = \text{Todd}(\mathcal{F}, q), \quad T_l(x, q) = G_l(x, \partial / \partial q) \text{Todd}(\mathcal{F}, q) \quad (1 \leq l \leq m),$$

where  $G_l(x, \partial/\partial q)$  is a homogeneous differential operator of order  $l$  whose coefficients are regular rational functions of  $x \in P$ .

It follows from (68) and (71) that the integral formula of Proposition 3.1 holds for small  $w$ , with

$$(72) \quad A_l(x, w) = T_l(x, \Lambda(x)w) .$$

Moreover, from the construction of the  $T_l$ , we see that the integrand has no poles if  $\|\Lambda(x)\Im w\| < \varepsilon_{\mathcal{F}}$  for all  $x \in P$ . Thus by analytic continuation, the identity holds on  $S(\varepsilon)$ , where  $\varepsilon = \varepsilon_{\mathcal{F}} / \sup_{x \in P} \|\Lambda(x)\|$ .

Statements (i) and (ii) follow from (72) and the fact that the coefficients of the Todd function are real and the constant term is 1.

To verify (iii), we fix  $x \in P$  and  $w \in C_x$ , and we let  $y_j = \langle w, f_j(x) \rangle$ , for  $j = 1, \dots, d$ .

**Claim 1.** (i)  $y_j \geq 0$ , for  $j = 1, \dots, d$ .  
(ii)  $y_j = 0$ , for  $j \notin \mathcal{J}(x)$ .

*Proof of Claim 1.* (i) Let  $F$  be the face containing  $x$ . Then  $x - tf_j(x)$  is on the face  $F(h)$  corresponding to  $F$  of the polytope  $P(h)$ , where  $h_j = -t$  (for  $t > 0$  sufficiently small) and  $h_k = 0$  for  $k \neq j$ . In particular,  $x - tf_j(x) \in P$ . Since  $w$  is in the normal cone to  $F$ , we have  $\langle w, x - tf_j(x) \rangle \leq \langle w, x \rangle$ ; i.e.,  $y_j = \langle w, f_j(x) \rangle \geq 0$ .

(ii) If  $j \notin \mathcal{J}(x)$ , then  $F(h) \subset F$  for  $h$  as above, and hence  $f_j(x)$  is tangent to  $F$ , so that  $y_j = \langle w, f_j(x) \rangle = 0$ .  $\square$

**Claim 2.** Let  $G_F = \Gamma_{\mathcal{F}} \cap \overline{U_F}$ , where  $F$  is the face containing  $x$ . Then

$$A_0(x, w) = \sum_{\gamma \in G_F} \prod_{j \in \mathcal{J}(F)} \text{Todd}(e^{i2\pi g_j(\gamma)}, y_j) .$$

*Proof of Claim 2.* We have

$$A_0(x, w) = \text{Todd}(\mathcal{F}, y) , \quad y = (y_1, \dots, y_d) , \quad y_j = \langle w, f_j(x) \rangle .$$

We first note that

$$G_F = \left\{ \gamma \in \mathbb{Z}^m : \gamma = \sum_{j \in \mathcal{J}(F)} t_j u_j, \quad 0 \leq t_j < 1 \right\} .$$

Thus, if  $\gamma \in \Gamma_{\mathcal{F}} \setminus G_F$ , then there exists  $j \notin \mathcal{J}(F)$  such that  $g_j(\gamma) \neq 0$ . But by Claim 1, the corresponding  $y_j = 0$ . Hence the term corresponding to  $\gamma$  in the sum (61) for  $\text{Todd}(\mathcal{F}, y)$  vanishes.

Now suppose that  $\gamma \in \Gamma_{\mathcal{F}} \cap G_F$ . Then by definition,  $g_j(\gamma) = 0$ , for all  $j \notin \mathcal{J}(F)$ . Therefore

$$\text{Todd}(e^{i2\pi g_j(\gamma)}, y_j) = \text{Todd}(1, 0) = 1 \quad \text{for } j \notin \mathcal{J}(F) .$$

Hence the sum (61) for  $\text{Todd}(\mathcal{F}, y)$  reduces to the sum in Claim 2.  $\square$

We easily see that the set

$$Q_F := \{(g_j(\gamma))_{j \in \mathcal{J}(F)} \in \mathbb{Q}^n / \mathbb{Z}^n : \gamma \in G_F\}$$

is a subgroup of  $\mathbb{Q}^n / \mathbb{Z}^n$ , where  $n = \#\mathcal{J}(F) = \text{codim } F$ . The final conclusion (iii) of the proposition is then a consequence of Claims 1–2 and the following elementary lemma.

**Lemma 3.3.** *Let  $\Gamma$  be a finite subgroup of  $\mathbf{T}^n$  and let  $y_j \geq 0$ , for  $1 \leq j \leq n$ . Then*

$$\sum_{a \in \Gamma} \prod_{j=1}^n \text{Todd}(a_j, y_j) > 0, \quad a = (a_1, \dots, a_n).$$

*Proof.* We can assume without loss of generality that the  $y_j$  are positive. For suppose that  $y_n = 0$ , for example. Then  $\text{Todd}(a_n, y_n) = 1$  if  $a_n = 1$ , and  $\text{Todd}(a_n, y_n) = 0$  if  $a_n \neq 1$ . Then the sum reduces to the corresponding sum over the group  $\Gamma' := \{a' \in \mathbf{T}^{n-1} : (a', 1) \in \Gamma\}$ .

Let  $r_j = e^{-y_j} < 1$ . Then

$$\begin{aligned} S &:= \frac{1}{y_1 \cdots y_n} \sum_{a \in \Gamma} \prod_{j=1}^n \text{Todd}(a_j, y_j) = \sum_{a \in \Gamma} \prod_{j=1}^n (1 - a_j r_j)^{-1} \\ &= \sum_{a \in \Gamma} \sum_{\beta \in \mathbb{N}^n} a^\beta r^\beta = \sum_{\beta \in \mathbb{N}^n} c_\beta r^\beta, \end{aligned}$$

where

$$c_\beta = \sum_{a \in \Gamma} a^\beta.$$

We claim that  $c_\beta$  is either 0 or  $o(\Gamma)$ . Indeed, if  $a^\beta = 1$  for all  $a \in \Gamma$ , then  $c_\beta = o(\Gamma)$ . On the other hand, if there exists  $b \in \Gamma$  with  $b^\beta \neq 1$ , then  $c_\beta = \sum_{a \in \Gamma} a^\beta = \sum_{a \in \Gamma} (b \cdot a)^\beta = b^\beta c_\beta$ , and hence  $c_\beta = 0$ . Since  $c_0 = o(\Gamma)$ , it then follows that  $S > 0$ .  $\square$

This completes the proof of Proposition 3.1.  $\square$

#### 4. MASS ASYMPTOTICS: PROOF OF THEOREM 1.2

In this section, we prove a precise asymptotic formula for the conditional Szegő kernel on the diagonal (Theorem 4.1), which yields the mass asymptotics of Theorem 1.2.

First, we verify formula (31) for the expected mass density. Let  $P$  denote a convex integral polytope in  $\mathbb{R}^m$ . Recalling the definition (7) of the conditional probability measure  $\gamma_{p|P}$  on the space  $\text{Poly}(P)$  of polynomials with Newton polytope  $P$ , we see that the expected value of the mass density with respect to  $\gamma_{p|P}$  is given by

$$\mathbf{E}_{|P} (|f(z)|_{\text{FS}}^2) = \sum_{\alpha, \beta \in P} \frac{\mathbf{E}(\lambda_\alpha \bar{\lambda}_\beta) \chi_\alpha(z) \overline{\chi_\beta(z)}}{\|\chi_\alpha\| \|\chi_\beta\| (1 + \|z\|^2)^{p/2}}.$$

Since the  $\lambda_\alpha$  are independent complex random variables with variance 1 (i.e.,  $\mathbf{E}_{|P}(\lambda_\alpha \bar{\lambda}_\beta) = \delta_\alpha^\beta$ ), we have by (53)

$$(73) \quad \mathbf{E}_{|P} (|f(z)|_{\text{FS}}^2) = \sum_{\alpha \in P} \frac{|\chi_\alpha(z)|_{\text{FS}}^2}{\|\chi_\alpha\|^2} = \Pi_{|P}(z, z).$$

It then follows by expressing the Gaussian in spherical coordinates that

$$\mathbf{E}_{\nu_P} (|f(z)|_{\text{FS}}^2) = \frac{1}{\#P} \mathbf{E}_{|P} (|f(z)|_{\text{FS}}^2) = \frac{1}{\#P} \Pi_{|P}(z, z).$$

Replacing  $P$  with  $NP$ , we obtain

$$(74) \quad \mathbf{E}_{\nu_{NP}} (|f(z)|_{\text{FS}}^2) = \frac{1}{\#(NP)} \Pi_{|NP}(z, z).$$

Recall that by (57), the number  $\#(NP)$  of lattice points in the polytope  $NP$  is a polynomial in  $N$  of degree equal to the dimension of  $P$ . The mass asymptotics of Theorem 1.2 is an immediate consequence of (74) and the asymptotic expansion of the conditional Szegő kernel on the diagonal given in Theorem 4.1 below.

**Theorem 4.1.** *Suppose that  $P$  is a convex integral polytope in  $\mathbb{R}^m$  such that  $P \subset p\Sigma$  and  $P \not\subset \partial(p\Sigma)$ . Then*

(i) *If  $P^\circ \neq \emptyset$ , then for  $z$  in the classically allowed region  $\mathcal{A}_P$ , we have*

$$\Pi_{|NP}(z, z) = \prod_{j=1}^m (Np + j) + R_N(z), \quad \|R_N\|_{\mathcal{C}^k(K)} = O(e^{-\lambda_K N})$$

*for all compact  $K \subset \mathcal{A}_P$  and for all  $k \in \mathbb{N}$ , where  $\lambda_K > 0$ .*

(ii) *On each open forbidden region  $\mathcal{R}_F^\circ$ ,*

$$\Pi_{|NP}(z, z) = N^{\frac{m+r}{2}} e^{-Nb_P(z)} [c_0^F(z) + c_1^F(z)N^{-1} + \cdots + c_k^F(z)N^{-k} + R_k^F(z)],$$

*where  $r = \dim F$  and*

- (a)  $c_j^F \in \mathcal{C}^\infty(\mathcal{R}_F^\circ)$  and  $c_0^F > 0$  on  $\mathcal{R}_F^\circ$ ;
- (b)  $\|R_k^F\|_{\mathcal{C}^j(K)} = O(N^{-k-1})$ , for all compact  $K \subset \mathcal{R}_F^\circ$  and for all  $j, k$ ;
- (c)  $b_P > 0$  on  $(\mathbb{C}^*)^m \setminus \overline{\mathcal{A}_P}$ ;
- (d)  $b_P$  is given by formula (16);
- (e)  $b_P \in \mathcal{C}_{\mathbb{R}}^1((\mathbb{C}^*)^m)$  (with  $b_P = 0$  on  $\mathcal{A}_P$ ), and  $b_P$  is  $\mathcal{C}^\infty$  on each closed region  $\mathcal{R}_F$ .

Note that  $b_P$  fails to be  $\mathcal{C}^2$  at transition points, since the limit expected zero current  $\psi_P$  of Theorem 1.3 is discontinuous at the transition points and is given in terms of the second derivatives of  $b_P$  by formula (23).

*Remark.* For the case where  $P = \{\beta\}$  is a single lattice point in  $p\Sigma^\circ$ , the allowed region  $\mathcal{A}_{\{\beta\}}$  is all of  $(\mathbb{C}^*)^m$ . Since the point  $q(z)$  given by (14) lies in  $\partial P$ , in this case  $q(z) = \beta$  for all  $z \in (\mathbb{C}^*)^m$ . Recalling (17)–(18), we then have

$$(75) \quad b_{\{\beta\}}(z) = -2 \log \mathcal{M}_\beta(z) = \log \left| \widehat{\chi}_\beta^p(\mu^{-1}(\frac{1}{p}\beta)) \right|^2 - \log \left| \widehat{\chi}_\beta^p(z) \right|^2.$$

Theorem 4.1 then says that

$$(76) \quad \Pi_{|N\{\beta\}}(z, z) = \frac{1}{\|\chi_\beta\|^2} |\chi_\beta(z)|_{\text{FS}}^2 = N^{\frac{m}{2}} (c_0 + c_1 N^{-1} + c_2 N^{-2} + \cdots) \mathcal{M}_\beta(z)^{2N}.$$

(The asymptotic formula (76) is also an immediate consequence of formula (46) for the  $\mathcal{L}^2$  norms of monomials and Stirling's formula. A more precise formula and general results on the decay of monomials normalized over toric varieties are given in [STZ2].)

Replacing  $\beta$  with an arbitrary (nonlattice) point  $x \in P^\circ$  in (75), we define

$$(77) \quad b_{\{x\}} := -2 \log \mathcal{M}_x(z).$$

Recalling (18), we see that for a convex lattice polytope  $P$ , we have

$$b_P(z) = b_{q(z)}(z), \quad \Pi_{|NP}(z, z) = N^{\frac{m+r}{2}} e^{-Nb_{q(z)}(z)} [c_0^F(z) + c_1^F(z)N^{-1} + \cdots].$$

To compute  $b_{\{x\}}$ , we use (16) with  $q(z) = x$  and  $\mu(e^{-\tau/2} \cdot z) = \frac{1}{p}x$ , which yields

$$\tau_j = \log |z_j|^2 - \log x_j + \log(p - \sum_{j=1}^m x_j),$$

and hence

$$(78) \quad b_{\{x\}}(z) = \sum_{j=0}^m x_j \log \frac{x_j}{p} - \log \frac{|z|^{2x}}{(1 + \|z\|^2)^p} \quad (x_0 = p - \sum_{j=1}^m x_j).$$

The asymptotics of Theorem 4.1 are uniform away from the transition points only. To take care of the transition points, we shall also prove the following local uniform convergence result on all of  $(\mathbb{C}^*)^m$ :

**Proposition 4.2.** *Let  $P$  be as in Theorem 4.1. Then*

$$\frac{1}{N} \log \Pi_{|NP}(z, z) \rightarrow -b_P(z)$$

*uniformly on all compact subsets of  $(\mathbb{C}^*)^m$ .*

Proposition 4.2 is the main ingredient in our proofs of Theorems 1.3–1.4 on the asymptotics of zeros in §5.

The remainder of this section is devoted to the proofs of Theorem 4.1 and Proposition 4.2. To motivate the argument, we begin by using our integral formula for the polytope character (Proposition 3.1) to show that the integral is rapidly decaying when  $z$  is in the classically forbidden region, while for  $z$  in the allowed region, the phase is of positive type with a nondegenerate critical point and hence  $\Pi_{|NP}(z, z)$  has an asymptotic expansion in powers of  $\frac{1}{N}$ . To obtain our expansion in the forbidden region, we use Proposition 3.1 to give an oscillatory integral formula (89)–(92) for the conditional Szegö kernel  $\Pi_{|NP}(z, z)$ . We then seek deformations of the contour of integration so that the (analytic continuation of) the phase picks up a critical point. In §4.2, we formulate necessary and sufficient geometric conditions for the existence of a critical point where the phase has maximal real part, and in §4.2.1 we show that there is a contour for which these conditions are satisfied. Assuming that  $P$  is simple, we obtain in §4.2.2 the asymptotic expansion (ii) by performing a stationary phase analysis over the polytope when the critical points of the phase lie on the boundary. In §4.2.3, we show that the decay function satisfies properties (c) and (e). In §4.2.4, we verify the precise asymptotic formula (i), and in §4.2.5 we complete the proof of part (ii) of Theorem 4.1 by extending our results to nonsimple polytopes. We prove Proposition 4.2 in §4.2.6.

We start our proof of Theorem 4.1 with a simple integral formula for the conditional Szegö kernel:

$$(79) \quad \Pi_{|NP}(z, z) = \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} \Pi_{NP}(z, e^{i\varphi} \cdot z) \chi_{NP}(e^{i\varphi}) d\varphi,$$

where we recall that

$$\mathbf{T}^m = \{(\zeta_1, \dots, \zeta_m) \in (\mathbb{C}^*)^m : |\zeta_j| = 1, 1 \leq j \leq m\}.$$

The identity (79) follows immediately from the explicit formulas (52) and (54) for the Szegö kernels and the definition (55) of the character  $\chi_{NP}$ .

Substituting formula (52) for  $\Pi_{NP}(z, e^{i\varphi} \cdot z)$  in (79), we obtain

$$(80) \quad \Pi_{|NP}(z, z) = \frac{(Np + m)!}{(Np)!(2\pi)^m} \int_{\mathbf{T}^m} \left[ \frac{1 + \sum_{j=1}^m e^{-i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right]^{Np} \chi_{NP}(e^{i\varphi}) d\varphi.$$

**We now assume that  $P$  is a simple polytope.** In §4.2.5, we shall derive the general case as a corollary of our results for simple polytopes.

**4.1. The classically allowed region.** In order to motivate our argument, we give a short proof of part (i) of Theorem 4.1 with the exponentially small error term replaced with an  $O(N^{-\infty})$  error. The actual proof of part (i) will be given in §4.2.4 and does not depend on the results of this section.

Assume for simplicity that  $P^\circ \neq \emptyset$ . We first choose a cut-off function  $\xi \in \mathcal{C}^\infty(\mathbb{R}^m)$  vanishing outside the  $\varepsilon_{\mathcal{F}}$ -ball and with  $\xi \equiv 1$  on the  $(\varepsilon_{\mathcal{F}}/2)$ -ball. Since  $\Pi_{NP}(z, e^{i\varphi} \cdot z)$  decays exponentially (in  $N$ ) when  $\|\varphi\| \geq \varepsilon_{\mathcal{F}}$ , it follows from Proposition 3.1 and (80) that

$$(81) \quad \Pi_{|NP}(z, z) \sim N^{2m} \frac{\sigma(N)}{(2\pi)^m} \int_{\mathbf{T}^m} \int_P \xi(\varphi) e^{N\Psi_{\mathcal{A}}(\varphi, x; z)} [A_0(x, \varphi) + A_1(x, \varphi)N^{-1} + \dots] dx d\varphi,$$

where  $\sigma(N) = (p + \frac{1}{N}) \cdots (p + \frac{m}{N}) = p^m + \sigma_1 N^{-1} + \dots + \sigma_M N^{-m}$ ,  $A_0(x, 0) = 1$  and the phase  $\Psi_{\mathcal{A}}$  is given by

$$(82) \quad \Psi_{\mathcal{A}}(\varphi, x; z) = i\langle \varphi, x \rangle + p \log \left( \frac{1 + \sum_{j=1}^m e^{-i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right).$$

Here, ‘ $\sim$ ’ means modulo  $O(N^{-\infty})$ .

The integral in (81) is an ‘oscillatory integral’ with ‘complex phase’  $\Psi_{\mathcal{A}}$ ; to evaluate it by the method of stationary phase [Hö, Ch. 7], we must find the critical points of  $\Psi_{\mathcal{A}}$ . The (interior) critical point equations

$$d_x \Psi_{\mathcal{A}} = i\varphi = 0, \quad d_\varphi \Psi_{\mathcal{A}}|_{\varphi=0} = ix - \frac{ip}{1 + \|z\|^2} (|z_1|^2, \dots, |z_m|^2) = ix - ip\mu_\Sigma(z) = 0$$

yield

$$(83) \quad \varphi = 0, \quad x = p\mu(z).$$

We note that  $\Re \Psi_{\mathcal{A}} \leq 0$  and  $\Psi_{\mathcal{A}} = 0$  at the critical point.

The Hessian  $\mathcal{H}\Psi_{\mathcal{A}}$  of  $\Psi_{\mathcal{A}}$  (with respect to the variables  $\varphi, x$ ) is of the form

$$(84) \quad \mathcal{H}\Psi_{\mathcal{A}}|_{(0,x)} = \begin{pmatrix} \mathbf{C} & i\mathbf{I} \\ i\mathbf{I} & 0 \end{pmatrix}$$

where  $\mathbf{I}$  denotes the  $m \times m$  identity matrix. The matrix  $\mathbf{C}$  is given by

$$C_{jk} = \frac{\partial^2 \Psi_{\mathcal{A}}}{\partial \varphi_j \partial \varphi_k} \Big|_{(0,x)} = p(-I_j \delta_j^k + I_j I_k),$$

where

$$I_j = \frac{|z_j|^2}{1 + \|z\|^2}; \quad \text{i.e., } \mu(z) = (I_1, \dots, I_m).$$

Hence

$$(85) \quad \det \mathcal{H}\Psi_{\mathcal{A}}|_{(0,x)} = 1$$

and the inverse Hessian is given by

$$\mathcal{H}\Psi_{\mathcal{A}}|_{(0,x)}^{-1} = \begin{pmatrix} 0 & -i\mathbf{I} \\ -i\mathbf{I} & \mathbf{C} \end{pmatrix}.$$

Hence, the Hessian operator is given by

$$H = -i \sum_j \partial^2 / \partial \varphi_j \partial x_j + \sum_{jk} C_{jk} \partial^2 / \partial x_j \partial x_k.$$

By (81) and [Hö, Theorem 7.7.5], we have a stationary phase expansion of the form

$$\Pi_{|NP}(z, z) \sim p^m \sigma(N) \sum_{2k \geq 3j \geq 0} \frac{1}{j!k!2^k} N^{m+j-k} H^k g_3^j \Big|_{\varphi=0, x=p\mu(z)},$$

where  $g_3$  is the third order Taylor remainder of the phase  $\Psi_A$ . We note that  $g_3$  is a function of  $\varphi$  alone, but that every  $\varphi$  derivative is accompanied by an  $x$ -derivative in  $H$ . It follows that only the very first term of the summation is nonzero. Thus, the stationary phase expansion is simply  $p^m \sigma(N) N^m$ .

**4.2. The classically forbidden region.** By (83),  $\Psi$  has a critical point if and only if  $z$  is in the classically allowed region. Thus, in the classically forbidden region, there are no critical points and the integral (81) is rapidly decaying.

We now let  $z$  be a point in the classically forbidden region. To evaluate the integral by the method of stationary phase, we need to deform the contour of integration to pick up a critical point with maximal real part along the contour. We complexify the real torus  $\mathbf{T}^m$  to  $(\mathbb{C}^*)^m$  with variables  $\zeta_j = \rho_j e^{i\varphi_j}$  so that (80) may be written as

$$(86) \quad \Pi_{|NP}(z, z) = \frac{(Np+m)!}{(Np)!(2\pi i)^m} \int_{\mathbf{T}^m} \left( \frac{1 + \sum_{j=1}^m \zeta_j^{-1} |z_j|^2}{1 + \|z\|^2} \right)^{Np} \chi_{NP}(\zeta) \prod_{j=1}^m \frac{d\zeta_j}{\zeta_j}.$$

Since the integrand is holomorphic in  $\zeta \in (\mathbb{C}^*)^m$ , we can deform the contours in (86) and instead integrate over a torus of the form

$$\mathbf{T}_\tau^m := \{ \zeta \in (\mathbb{C}^*)^m : |\zeta_1| = e^{\tau_1}, \dots, |\zeta_m| = e^{\tau_m} \}.$$

We shall show below how to choose the parameter  $\tau = (\tau_1, \dots, \tau_m) \in \mathbb{R}^m$  to obtain our required critical point.

Deforming (86) to  $\mathbf{T}_\tau^m$ , we have

$$(87) \quad \Pi_{|NP}(z, z) = \frac{(Np+m)!}{(Np)!(2\pi)^m} \int_{\mathbf{T}^m} e^{N\Psi_0(\varphi; \tau, z)} \chi_{NP}(e^{\tau+i\varphi}) d\varphi,$$

where

$$(88) \quad \Psi_0(\varphi; \tau, z) = p \log \left( \frac{1 + \sum_{j=1}^m e^{-\tau_j - i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right).$$

Using a  $\mathcal{C}^\infty$  partition of unity  $\{\xi_1, \xi_2\}$  of  $\mathbf{T}^m$ , where  $\xi_1(\varphi)$  has support in a small neighborhood of 0 and is identically 1 near 0, we decompose (87) into two integrals:

$$(89) \quad \begin{aligned} \Pi_{|NP}(z, z) &= \frac{(Np+m)!}{(Np)!} (\mathcal{I}_N^1 + \mathcal{I}_N^2), \\ \mathcal{I}_N^j &:= \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} \xi_j(\varphi) e^{N\Psi_0(\varphi; \tau, z)} \chi_{NP}(e^{\tau+i\varphi}) d\varphi. \end{aligned}$$

We shall show that  $\mathcal{I}_N^1$  has the desired asymptotic expansion (when  $z$  is not a transition point) and that  $\mathcal{I}_N^2 = O(e^{-\delta} \mathcal{I}_N^1)$  for some  $\delta > 0$ .

Let  $\dim P = n$ . Under our assumption that  $P$  is simple, Proposition 3.1 lets us write the primary part  $\mathcal{I}_N^1$  of the Szegö kernel given in (89) as an oscillatory integral over  $\mathbf{T}^m \times P$ :

$$(90) \quad \mathcal{I}_N^1 = N^n \int_{\mathbf{T}^m} \int_P e^{N\Psi(\varphi, x; \tau, z)} A(\varphi, x, N; \tau) dx d\varphi,$$



where the phase is given by

$$(91) \quad \Psi(\varphi, x; \tau, z) = \langle \tau + i\varphi, x \rangle + p \log \left( \frac{1 + \sum_{j=1}^m e^{-\tau_j - i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right),$$

and the amplitude is given by

$$(92) \quad A(\varphi, x, N; \tau) = \xi_1(\varphi) [A_0(x, \tau + i\varphi) + A_1(x, \tau + i\varphi)N^{-1} + \cdots + A_N(x, \tau + i\varphi)N^{-n}].$$

We therefore look for complex critical points of  $\Psi$ . We note that as before, by the triangle inequality,

$$(93) \quad \Re\Psi(\varphi, x; \tau, z) < \Re\Psi(0, x; \tau, z), \quad \text{for } \varphi \neq 0.$$

Thus for each fixed  $\tau$  and  $z$ , the maximal value of  $\Re\Psi$  occurs where  $\varphi = 0$ . From (91), we obtain

$$(94) \quad d_\varphi \Psi = ix - \frac{ip}{1 + \sum_j e^{-\tau_j - i\varphi_j} |z_j|^2} (e^{-\tau_1 - i\varphi_1} |z_1|^2, \dots, e^{-\tau_m - i\varphi_m} |z_m|^2).$$

Hence the critical-point equation  $d_\varphi \Psi(0, x; \tau, z) = 0$  is equivalent to

$$(95) \quad \frac{1}{p} x = \mu(e^{-\tau/2} \cdot z).$$

From the second critical point equation

$$(96) \quad d_x \Psi = \tau + i\varphi = 0,$$

it follows that there are no critical points on the interior of  $P$ , since by (95) we must choose  $\tau \neq 0$  to obtain a critical point when  $z$  is in the classically forbidden region.

Suppose  $\tau \in \mathbb{R}^m$  and  $x^0 \in P$ . Recall that  $\tau$  is in the normal cone to  $P$  at  $x^0$  if  $\langle \tau, x^0 \rangle \geq \langle \tau, y \rangle$  for all  $y \in P$ . Since  $\Psi(\tau, 0, x, z) = \langle \tau, x \rangle + c_{\tau, z}$ , where  $c_{\tau, z}$  is independent of  $x$ , it follows that  $\Re\Psi$  takes its maximum along the contour  $\mathbf{T}_\tau^m$  at the point  $\{\varphi = 0, x = x^0\}$  if (and only if)  $\tau$  is in the normal cone  $C_{x^0}^P$  to  $P$  at  $x^0$ . Recall that if  $\tau$  is in the normal cone at  $x^0$ , then  $\tau$  must be orthogonal to the face of  $P$  containing  $x^0$ .

Therefore, to obtain a ‘critical point’ that determines the asymptotics of the integral (90), we must find  $\tau \in \mathbb{R}^m, x \in \partial P$  such that (95) holds and  $\tau \in C_x$ . We show the existence of such  $\tau, x$  in the next section.

4.2.1. *Existence of critical points.* Recall that the normal cone  $C_F$  of a face  $F$  of a convex polytope  $Q \subset \mathbb{R}^m$  is given by

$$(97) \quad C_F = \{u \in \mathbb{R}^m : \langle u, x \rangle = \sup_{y \in Q} \langle u, y \rangle, \forall x \in F\},$$

so that  $C_F = C_x^Q$  for all points  $x \in F$ . The purpose of this section is to prove the following general existence result:

**Lemma 4.3.** *Let  $Q$  be a convex polytope contained in  $\Sigma$  and suppose that  $Q \not\subset \partial\Sigma$ . Then for each  $z \in (\mathbb{C}^*)^m$ , there exist unique  $\tau_z \in \mathbb{R}^m$  and  $x_z \in Q$  so that*

- $\mu(e^{-\tau_z/2} \cdot z) = x_z,$
- $\tau_z \in C_{x_z}^Q.$

Here we do not assume that  $Q$  is simple or even that  $Q$  is integral. We note that even though  $\tau_z$  is orthogonal to  $\partial Q$  at  $x_z$ , the orbit  $\{\mu(e^{r\tau_z} \cdot z) : r \in \mathbb{R}\}$  is in general not orthogonal to  $\partial Q$ , even when  $x_z$  lies in a facet. (See also Figure 6 and the associated remark.) Our proof of Lemma 4.3 uses the following elementary, but not so well known, fact about the invertibility of Lipschitz maps.

**Lemma 4.4** ([Fan]). *Let  $f : U \rightarrow \mathbb{R}^m$  be a Lipschitz map, where  $U$  is open in  $\mathbb{R}^m$ . Then  $f$  has a local orientation-preserving Lipschitz inverse with Lipschitz constant  $L$  at a point  $x_0 \in U$  if and only if there exists  $\varepsilon > 0$  such that*

- (i)  $\liminf_{v \rightarrow 0} |f(x+v) - f(x)|/|v| \geq 1/L$  for all  $x \in B_\varepsilon(x_0)$ ,
- (ii)  $\det f'(x) > 0$  for all  $x \in B_\varepsilon(x_0)$  such that  $f$  is differentiable at  $x$ ,
- (iii)  $f(x_0) \notin f(\partial B_\varepsilon(x_0))$  and  $\deg [f : \partial B_\varepsilon(x_0) \rightarrow \mathbb{R}^m \setminus \{f(x_0)\}] = 1$ .

Here, we let  $B_\varepsilon(x_0) = \{x \in \mathbb{R}^m : |x - x_0| < \varepsilon\}$  denote the  $\varepsilon$ -ball about  $x_0 \in \mathbb{R}^m$ . By the degree in (iii), we mean the degree of  $f_*(\partial B_\varepsilon(x_0)) \in H_{m-1}(\mathbb{R}^m \setminus \{f(x_0)\}, \mathbb{Z})$ . The hypotheses (i) and (iii) in [Fan] differ slightly from those above, but Fan's proof of sufficiency uses only (i)–(iii) above. (Necessity is obvious.) Hypothesis (i) is given as a lemma in [Fan]. Hypothesis (iii) above is replaced in [Fan] by the condition that the index of  $f$  at  $x_0$  is 1, which is easily seen to be equivalent to (iii) under the assumptions (i) and (ii).

Our proof of Lemma 4.3 also makes use of the ‘normal bundle’ of a polytope, which we define below as the collection of normal cones at the points of the polytope.

**Definition.** The *normal bundle*  $\mathcal{N}(Q)$  of a convex polytope  $Q \subset \mathbb{R}^m$  is the subset of  $T_{\mathbb{R}^m} = \mathbb{R}^m \times \mathbb{R}^m$  consisting of pairs  $(x, v)$ , where  $x \in Q$  and  $v \in C_x^Q$ . (Note that  $\mathcal{N}(Q)$  is not a fiber bundle over  $Q$ .)

The normal bundle  $\mathcal{N}(Q)$  is a piecewise smooth submanifold of  $T_{\mathbb{R}^m}$ ; it is homeomorphic to  $\mathbb{R}^m$  via the ‘exponential map’

$$(98) \quad \mathcal{E}_Q : \mathcal{N}(Q) \rightarrow \mathbb{R}^m, \quad \mathcal{E}_Q(x, v) = x + v, \quad x \in F, \quad v \in C_F \quad (F \text{ a face of } Q).$$

It is easily seen that  $\mathcal{E}_Q$  is a homeomorphism and is a  $C^\infty$  (in fact, linear) diffeomorphism on each of the ‘pieces’  $\bar{F} \times C_F \subset \mathcal{N}(Q)$ .

*Proof of Lemma 4.3.* We note that the  $\mathbb{R}_+^m$  action on  $(\mathbb{C}^*)^m$  descends via the moment map  $\mu$  to an  $\mathbb{R}_+^m$  action on  $\Sigma^\circ$  given by  $r \odot \mu(z) = \mu(r \cdot z)$ . We let

$$\mathcal{N}^\circ = \mathcal{N}(Q) \cap (\Sigma^\circ \times \mathbb{R}^m),$$

and we consider the map

$$\Phi : \mathcal{N}^\circ \rightarrow \Sigma^\circ, \quad \Phi(x, \tau) = e^{\tau/2} \odot x.$$

We observe that the conclusion of the lemma is equivalent to the bijectivity of  $\Phi$ , since if  $z \in (\mathbb{C}^*)^m$ , we have

$$\Phi(x, \tau) \stackrel{\text{def}}{=} e^{\tau/2} \odot x = \mu(z) \iff x = e^{-\tau/2} \odot \mu(z) \stackrel{\text{def}}{=} \mu(e^{-\tau/2} \cdot z),$$

for  $(x, \tau) \in \mathcal{N}^\circ$ , i.e., for  $x \in \Sigma^\circ$  and  $\tau \in C_x^Q$ .

To show that  $\Phi$  is a bijection, we let  $\mathcal{L} : \Sigma^\circ \rightarrow \mathbb{R}^m$  be the diffeomorphism given by

$$(99) \quad \mathcal{L}(x) = \left( \log \frac{x_1}{1 - \sum x_j}, \dots, \log \frac{x_n}{1 - \sum x_j} \right),$$

so that

$$\mathcal{L} \circ \mu(w) = (\log |w_1|^2, \dots, \log |w_m|^2).$$

Thus, writing  $x = \mu(w)$ , we have

$$\begin{aligned} \mathcal{L} \circ \Phi(x, \tau) &= e^{\tau/2} \odot u_{\Sigma}(w) = \mathcal{L} \circ \mu(e^{\tau/2} \cdot w) \\ &= (\tau_1 + \log |w_1|^2, \dots, \tau_m + \log |w_m|^2) \\ (100) \quad &= \tau + \mathcal{L} \circ \mu(w) = \tau + \mathcal{L}(x). \end{aligned}$$

We first show that  $\mathcal{L} \circ \Phi : \mathcal{N}^\circ \rightarrow \mathbb{R}^m$  is proper: suppose on the contrary that the sequence  $\{(x^n, \tau^n)\}$  is unbounded in  $\mathcal{N}^\circ$ , but  $\mathcal{L} \circ \Phi(x^n, \tau^n) = \mathcal{L}(x^n) + \tau^n \rightarrow a \in \mathbb{R}^m$ . By passing to a subsequence, we can assume that  $x^n \rightarrow x^0 \in Q$ . Then  $x^0 \in \partial\Sigma$ , since otherwise  $\tau^n \rightarrow a - \mathcal{L}(x^0)$ . Write  $\tau^n = r^n u^n$ , where  $r^n > 0$ ,  $|u^n| = 1$ . We can assume without loss of generality that  $u^n \rightarrow u^0$ . We first consider the case where  $\sum x_j^0 < 1$ . If  $x^0 = 0$ , then  $u_j^0 \geq 0$  ( $1 \leq j \leq m$ ), since otherwise  $\mathcal{L} \circ \Phi(x^n, \tau^n)$  would diverge. But  $u^0 \in C_{\{0\}}$ ; hence  $Q \subset \{x : \langle u^0, x \rangle \leq 0\} \subset \mathbb{R}^m \setminus \Sigma^\circ$ , a contradiction. Now suppose that  $x_1^0 = \dots = x_k^0 = 0$ ,  $x_l^0 > 0$  for  $k < l \leq m$ . Then we conclude as before that  $u_j^0 \geq 0$  for  $1 \leq j \leq k$ . Furthermore, in this case,  $\tau_1^n \rightarrow +\infty$  and  $\tau_l^n \rightarrow \tau_l^\infty \in \mathbb{R}$  and thus  $u_l^0 = 0$  for  $k < l \leq m$ , and we again obtain a contradiction. Finally, if  $\sum x_j^0 = 1$ , we can suppose that  $x_1^0 \neq 0$  and make the coordinate change  $\tilde{x}_1 = 1 - \sum x_j$ ,  $\tilde{x}_2 = x_2, \dots, \tilde{x}_m = x_m$  (which corresponds to permuting the homogeneous coordinates in  $\mathbb{C}\mathbb{P}^m$ ) to reduce to the previous case. Therefore,  $\mathcal{L} \circ \Phi : \mathcal{N}^\circ \rightarrow \mathbb{R}^m$  is a proper map.

Let  $\mathcal{E} = \mathcal{E}_Q : \mathcal{N}(Q) \xrightarrow{\approx} \mathbb{R}^m$ , and let  $U = \mathcal{E}(\mathcal{N}^\circ) \subset \mathbb{R}^m$ . We consider the map  $f : U \rightarrow \mathbb{R}^m$  given by  $f \circ \mathcal{E}|_{\mathcal{N}^\circ} = \mathcal{L} \circ \Phi$ , i.e., by the commutative diagram

$$\begin{array}{ccc} \mathcal{N}^\circ & \xrightarrow{\Phi} & \Sigma^\circ \\ \approx \downarrow \mathcal{E} & & \approx \downarrow \mathcal{L} \\ U & \xrightarrow{f} & \mathbb{R}^m \end{array}$$

Since  $\mathcal{L} \circ \Phi$  is a proper map,  $f$  is also proper. Hence to show that  $\Phi$  is a bijection, it suffices to show that  $f$  is a local homeomorphism and is therefore a (global) homeomorphism.

To describe the map  $f$ , for each  $x \in \Sigma^\circ$ , we let  $q_x$  denote the point in  $Q$  that is closest to  $x$ . We note that  $q_x \in \Sigma^\circ$ ; if  $x \notin Q$ , then  $q_x \in \partial Q$ ; if  $x \in Q$ , then  $q_x = x$ . Furthermore,  $\mathcal{E}^{-1}(x) = (q_x, x - q_x)$  and hence

$$f(x) = x - q_x + \mathcal{L}(q_x).$$

We shall show that  $f$  satisfies the hypotheses of Lemma 4.4. Let  $F$  be an arbitrary face of  $Q \cap \Sigma^\circ$ . To verify (i) and (ii), it suffices to show that  $\det Df > 0$  on the (noncompact) polyhedron

$$U_F := \mathcal{E}(\bar{F} \times C_F) \cap U = \mathcal{E}[(\bar{F} \cap \Sigma^\circ) \times C_F].$$

To compute the determinant, we let  $T_F \subset \mathbb{R}^m$ ,  $N_F = T_F^\perp \subset \mathbb{R}^m$  denote the tangent and normal spaces, respectively, of  $F$ . Let  $x^0 \in F$  be fixed. For  $y \in U_F$ , we have

$$y = x + v \xrightarrow{f} \mathcal{L}(x) + v, \quad x - x^0 \in T_F, v \in N_F.$$

Choose orthonormal bases  $\{Y_1, \dots, Y_r\}$ ,  $\{Y_{r+1}, \dots, Y_m\}$  of  $T_F$ ,  $N_F$ , respectively. We let  $Df$  denote the matrix of the derivative  $(f|_{U_F})'$  with respect to the basis

$\{Y_1, \dots, Y_m\}$  of  $\mathbb{R}^m$ . We have

$$(101) \quad Df = \begin{pmatrix} T^t \mathcal{L}'(x) T & 0 \\ * & I \end{pmatrix},$$

where  $T$  is the  $m \times r$  matrix  $[Y_1 \cdots Y_r]$ . We have by (99)

$$(102) \quad (\mathcal{L}'(x))_{jk} = \frac{1}{x_0} + \delta_j^k \frac{1}{x_j}, \quad x_0 = 1 - \sum_{j=1}^m x_j > 0,$$

for  $x \in \Sigma^\circ$ . Hence  $\mathcal{L}'(x)$  is a positive definite symmetric matrix, it being the sum of a semipositive matrix (all of whose entries are  $\frac{1}{x_0}$ ) and a positive definite diagonal matrix. Therefore,  $T^t \mathcal{L}'(x) T$  is positive definite, and hence  $\det Df(x) = \det(T^t \mathcal{L}'(x) T) > 0$ , completing the proof that hypotheses (i) and (ii) of Lemma 4.4 are satisfied.

Note that (102) implies that the eigenvalues of  $\mathcal{L}'(x)$  are  $> 1$  for  $x \in \Sigma^\circ$  and hence by (101),  $Df(x)$  is a diagonalizable matrix whose eigenvalues are real and  $\geq 1$ .

We verify (iii) by a homotopy argument: Choose a point  $q^0 \in Q^\circ$ . We contract  $Q$  to  $q^0$ ; i.e., for  $0 \leq t \leq 1$ , we let

$$Q_t = (1-t)\{q^0\} + tQ,$$

so that  $Q_0 = \{q^0\}$ ,  $Q_1 = Q$ . For  $0 \leq t < 1$ ,  $Q_t \subset \Sigma^\circ$ , and hence we have a map

$$\Phi_t : \mathcal{N}(Q_t) \rightarrow \Sigma^\circ.$$

For  $0 \leq t < 1$ , we define  $f_t : \mathbb{R}^m \rightarrow \mathbb{R}^m$  by the commutative diagram

$$\begin{array}{ccc} \mathcal{N}(Q_t) & \xrightarrow{\Phi_t} & \Sigma^\circ \\ \approx \downarrow \varepsilon_{Q_t} & & \approx \downarrow \mathcal{L} \\ \mathbb{R}^m & \xrightarrow{f_t} & \mathbb{R}^m \end{array}$$

As before we have

$$(103) \quad f_t(x) = x - q_x^t + \mathcal{L}(q_x^t),$$

where  $q_x^t$  is the (unique) point of  $Q_t$  closest to  $x$ . The above argument shows that the maps  $f_t$  also satisfy (i) and (ii) of Lemma 4.4.

We write

$$F : (\mathbb{R}^m \times [0, 1]) \cup (U \times \{1\}) \rightarrow \mathbb{R}^m, \quad (x, t) \mapsto f_t(x),$$

where  $f_1 = f : U \rightarrow \mathbb{R}^m$ . One easily sees that

$$\left| q_x^t - q_{x'}^{t'} \right| \leq |x - x'| + |t - t'|,$$

and hence  $F$  is continuous. Furthermore,  $F$  is uniformly continuous on  $\mathbb{R}^m \times [0, t]$ , for each  $t < 1$ .

Let  $H$  denote the set of  $t \in [0, 1)$  such that  $f_t : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is an orientation preserving homeomorphism. We shall show that  $H = [0, 1)$ . We first observe that  $0 \in H$  since

$$f_0(x) = x - q + \mathcal{L}(q) = x + \text{constant}.$$

Note that (101) also holds for the maps  $f_t$  and hence as above, the eigenvalues of  $Df_t$  are real and  $\geq 1$ . Hence  $f_t$  satisfies conditions (i) and (ii) of Lemma 4.4 with  $L = 1$ , for all  $t \in [0, 1]$ .

We now show that  $H$  is open in  $[0, 1)$ . Suppose  $t_0 \in H$ . Then by Lemma 4.4,  $\text{Lip}(f_{t_0}^{-1}) \leq 1$  and hence

$$|x - x_0| = 1 \implies |f_{t_0}(x) - f_{t_0}(x_0)| \geq 1 .$$

By the uniform continuity of  $F$  on  $\mathbb{R}^m \times [0, \frac{t_0+1}{2}]$ , we can choose  $\varepsilon > 0$  such that for all  $x_0 \in \mathbb{R}^m$ , we have

$$(104) \quad |f_t(x) - f_t(x_0)| \geq \frac{1}{2}, \quad \text{for } x \in \partial B_1(x_0), |t - t_0| < \varepsilon .$$

To simplify notation, we shall write

$$\deg(f, x_0, r) := \deg [f : \partial B_r(x_0) \rightarrow \mathbb{R}^m \setminus \{f(x_0)\}] .$$

We conclude from (104) that for  $|t - t_0| < \varepsilon$ ,

$$\deg(f_t, x_0, 1) = \deg(f_{t_0}, x_0, 1) = 1 ,$$

and hence  $f_t$  is a local homeomorphism at  $x_0$ , by Lemma 4.4. By the very first part of the argument that  $f : U \rightarrow \mathbb{R}^m$  is proper, we easily see that  $f_t$  is proper. Since  $x_0 \in \mathbb{R}^m$  is arbitrary, it follows that  $f_t : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is a covering map and therefore is a homeomorphism.

Next we show that  $H$  is closed in  $[0, 1)$  and hence  $H = [0, 1)$ . Let  $t_n \in H$  such that  $t_n \rightarrow t_0 \in [0, 1)$ . Since  $f_{t_0}$  satisfies condition (i) of Lemma 4.4, we can choose  $\varepsilon > 0$  so that  $f_{t_0}(x_0) \notin f_{t_0}(\partial B_\varepsilon(x_0))$ . Then for  $n$  sufficiently large,

$$\deg(f_{t_0}, x_0, \varepsilon) = \deg(f_{t_n}, x_0, \varepsilon) = 1 .$$

It follows as above that  $f_{t_0} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is a homeomorphism.

We have shown that  $f_t$  is a homeomorphism for  $0 \leq t < 1$ . To complete the proof of the lemma, we must show that  $f = f_1$  is a local homeomorphism. So we let  $x_0 \in U$  be arbitrary, and we choose  $\varepsilon > 0$  such that  $\overline{B_\varepsilon(x_0)} \subset U$  and  $f(x_0) \notin f(\partial B_\varepsilon(x_0))$ . Then for  $t < 1$  sufficiently close to 1, we have as before

$$\deg(f, x_0, \varepsilon) = \deg(f_t, x_0, \varepsilon) = 1 .$$

Thus by Lemma 4.4,  $f$  is a local homeomorphism. This completes the proof of Lemma 4.3.  $\square$

We have thus shown (as a consequence of (93) and Lemma 4.3 with  $Q = \frac{1}{p}P$ ) that for each  $z \in (\mathbb{C}^*)^m$ , there exist (unique)  $\tau_z \in \mathbb{R}^m$  and  $q(z) = px_z \in P$  such that the phase function  $\Psi(\cdot, \cdot; \tau_z, z) : \mathbf{T}^m \times P \rightarrow \mathbb{C}$  given by (88) takes its maximum real part at the point  $(0, q(z))$ , and furthermore  $(0, q(z))$  is a critical point of  $\Psi|_{\mathbf{T}^m \times F}$ , where  $F$  is the face of  $P$  containing  $q(z)$ . The latter statement follows from (95), (96) and the fact that vectors in the normal cone  $C_F$  are orthogonal to  $F$ ; i.e.,  $C_F \subset F^\perp$ .

By the *boundary* of the normal cone  $C_F$ , we mean the boundary  $\partial C_F$  of  $C_F$  in  $F^\perp$ ; the complement  $C_F \setminus \partial C_F$  is called the *interior* of  $C_F$ . We note the following consequence of the proof of Lemma 4.3:

**Lemma 4.5.** *A point  $z \in (\mathbb{C}^*)^m$  is a transition point if and only if  $\tau_z \in \partial C_{q(z)}^P$ .*

*Proof.* It follows from (100) that  $2 \text{Log}(\mathcal{R}_F) = \mathcal{L}(F) + C_F$ , where  $\text{Log}(z) = (\log |z_1|, \dots, \log |z_m|)$  and that

$$(105) \quad \partial[2 \text{Log}(\mathcal{R}_F)] = 2 \text{Log}(\partial \mathcal{R}_F) = \mathcal{L}(F) + \partial C_F .$$

Thus

$$z \in \partial\mathcal{R}_F \iff 2 \operatorname{Log}(z) \in \mathcal{L}(F) + \partial C_F \iff 2 \operatorname{Log}(z) = \mathcal{L}(x_z) + \tau_z, \tau_z \in \partial C_F.$$

□

*Remark.* The identity (105) tells us that we can decompose  $\mathbb{R}^m$  as the disjoint union of the sets  $\mathcal{L}(F) + C_F$ . If  $z$  is a transition point, then  $2 \operatorname{Log}(z)$  must lie in the common boundary of at least two of the sets  $\mathcal{L}(F) + C_F$ . Figure 6 shows the transition points in log coordinates (as solid lines) for the case where  $P$  is the polytope of Figure 5. Note that although  $C_{F'}$  is orthogonal to  $F'$  in Figure 5,  $C_{F'}$  is not orthogonal to  $\mathcal{L}(F')$ . Similarly, in Lemma 4.3, if  $q(z) \in F'$ , then the orbit  $\{\mu(e^{r\tau_z} \cdot z) : r \in \mathbb{R}\}$  is not orthogonal to  $F'$ .

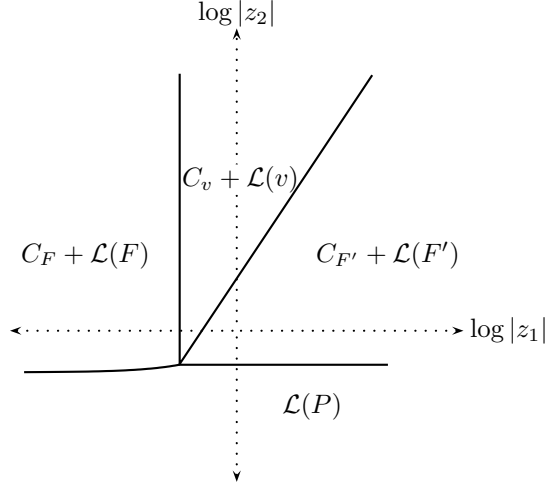


FIGURE 6. Transition points in log coordinates for the polytope of Figure 5

4.2.2. *Continuation of the proof of Theorem 4.1.* Since the critical point of  $\Psi$  on the contour  $\mathbf{T}_r^m$  always occurs on the boundary of  $P$  when  $z$  is in the forbidden region, we shall use the method of stationary phase for complex oscillatory integrals on quadrants. Since this general method is not so well known, we state here the basic result we need:

**Lemma 4.6.** *Let  $\Psi(\xi, y), A(\xi, y) \in C^\infty(\mathbb{R}_{\geq 0}^k \times \mathbb{R}^s)$  such that  $d_y \Psi(0, 0) = 0$ ,  $A$  has compact support and*

- (1)  $\frac{\partial \Psi}{\partial \xi_j} \neq 0$  on  $\operatorname{Supp}(A)$  for  $1 \leq j \leq k$ ,
- (2)  $d_y \Psi(0, y) \neq 0$  for  $(0, y) \in \operatorname{Supp}(A) \setminus \{0\}$ ,
- (3)  $\det \mathcal{H}_y \Psi(0, 0) \neq 0$  (where  $\mathcal{H}_y$  denotes the Hessian with respect to  $y$ ),
- (4)  $\Re \Psi \leq \Re \Psi(0, 0)$  on  $\operatorname{Supp}(A)$ .

Then

$$\begin{aligned} & \int_{\mathbb{R}^s} \int_{\mathbb{R}_{\geq 0}^k} e^{N\Psi(\xi, y)} A(\xi, y) d\xi dy \\ &= N^{-k-s/2} e^{N\Psi(0, 0)} [c_0 + c_1 N^{-1} + c_2 N^{-2} + \cdots + c_l N^{-l} + O(N^{-l-1})] \end{aligned}$$

for  $l = 1, 2, 3, \dots$ , where

$$c_0 = \frac{(2\pi)^{s/2} A}{\sqrt{\det(-\mathcal{H}_y \Psi)} \prod_{j=1}^k \partial\psi/\partial\xi_j} \Big|_{\xi=y=0}.$$

(If  $s = 0$ , then hypotheses (2) and (3) are vacuous, and the determinant in the formula for  $c_0$  is 1.)

*Proof.* (See also [AGV].) Let us first consider the case  $k = 1$ . Integrating by parts,

$$(106) \quad \int_{\mathbb{R}^s} \int_0^{+\infty} e^{N\Psi} A d\xi_1 dy = \frac{1}{N} \int_{\mathbb{R}^s} e^{N\Psi} \frac{A}{\partial\Psi/\partial\xi_1} \Big|_{\xi_1=0} dy \\ - \frac{1}{N} \int_{\mathbb{R}^s} \int_0^{+\infty} e^{N\Psi} \frac{\partial}{\partial\xi_1} \left[ \frac{A}{\partial\Psi/\partial\xi_1} \right] d\xi_1 dy.$$

Applying the stationary phase expansion [Hö, Th. 7.7.5] to the first term of (106) and iterating, we obtain the desired expansion.

The case  $k > 1$  follows by induction on  $k$ , integrating by parts as above.  $\square$

We consider a point  $z$  in the classically forbidden region. Let  $\tau_z, x_z$  be as in Lemma 4.3, and let  $q(z) = px_z$ . Hence  $\tau = \tau_z, x = q(z)$  satisfies the critical point equation (95), and as we observed above,  $q(z) \in \partial P$ . The decay function  $b_P(z)$  from (16) is given by

$$(107) \quad b_P(z) = -\Psi(0, q(z)).$$

Recalling (90), we further decompose the integral  $\mathcal{I}_N^1$  into two parts:

$$(108) \quad \mathcal{I}_N^1 = \mathcal{I}'_N + \mathcal{I}''_N, \quad \mathcal{I}'_N = N^n \int_{\mathbf{T}^m} \int_P e^{N\Psi(\varphi, x; \tau_z, z)} \rho(x) A(\varphi, x, N; \tau_z) dx d\varphi,$$

where  $A(\varphi, x, N; \tau_z)$  is given by (92) and  $\rho$  is supported in a small neighborhood of  $q(z)$  and is  $\equiv 1$  near  $q(z)$ . Note that it follows from (94)–(95) that  $d_\varphi \Psi(0, x; \tau_z, z) \neq 0$  for  $x \neq q(z)$ . Since  $d_\varphi \Psi$  does not vanish on the support of  $1 - \rho(x)$  and  $\sup_{\mathbf{T}^m \times P} \Psi = \Psi(0, q(z)) = -b_P(z)$ , we conclude by performing the  $\varphi$  integration first that

$$(109) \quad \mathcal{I}''_N = e^{-Nb_P(z)} \cdot O(N^{-\infty})$$

(see [Hö]). Hence, we need consider only  $\mathcal{I}'_N$ .

Since we are not assuming that  $P$  has interior, we let  $n = \dim P$  and we let  $\langle P \rangle$  denote the  $n$ -plane containing  $P$ . The polytope  $P$  can be given by

$$P = \{x \in \langle P \rangle : \ell_j(x) \geq 0 \text{ for } 1 \leq j \leq d\}, \quad \ell_j(x) = \langle u_j, x \rangle + \lambda_j,$$

where the  $u_j$  are primitive vectors tangent to  $P$ . Suppose that  $q(z)$  is in a face  $F \subset P$  of dimension  $r$ . After permuting the indices of the normal vectors  $\{u_j\}$ , we may assume that

$$\bar{F} = \{x \in P : \ell_j(x) = 0, \text{ for } 1 \leq j \leq n - r\}.$$

We consider the translated cone

$$\Gamma := \{x \in \langle P \rangle : \ell_j(x) \geq 0\} \supset P$$

with vertex  $q(z)$ , which coincides with  $P$  in a sufficiently small neighborhood of  $q(z)$ . We then choose (nonhomogeneous) linear functions  $t_1, \dots, t_r$  on  $\mathbb{R}^m$  such that the map

$$L = (\ell_1, \dots, \ell_{n-r}, t_1, \dots, t_r) : \langle P \rangle \rightarrow \mathbb{R}^n$$

is bijective and  $L(q(z)) = 0$ . We note that  $L : \Gamma \xrightarrow{\cong} \mathbb{R}_{\geq 0}^{n-r} \times \mathbb{R}^r$  and  $L : F \hookrightarrow \{0\} \times \mathbb{R}^r$ .

We shall evaluate  $\mathcal{I}'_N$  by applying Lemma 4.6 with  $\xi_j = \ell_j(x)$  for  $1 \leq j \leq n-r$  and  $y = (t_1, \dots, t_r, \varphi_1, \dots, \varphi_m)$  and with the phase

$$\tilde{\Psi}(\xi, y) := \Psi(\varphi, L^{-1}(\xi; t_1, \dots, t_r); \tau, z)$$

and amplitude  $\rho(x)A(\varphi, x, N; \tau)$ . Since  $d_\varphi \Psi(0, q(z); \tau_z, z) = 0$  and  $\tau_z$  is perpendicular to  $F$ , it follows (using (91)) that  $d_y \tilde{\Psi}(0, 0) = 0$ . Since  $\tau_z$  is in the normal cone at  $q(z)$ , the real part of the phase  $\Psi$  takes its maximum at  $q(z)$ , and hence hypothesis (4) of the lemma holds. Also, hypothesis (2) is satisfied since  $d_\varphi \Psi(0, x; \tau_z, z) \neq 0$  for  $x \neq q(z)$ , and hence  $d_y \tilde{\Psi}(0, y) \neq 0$  for  $y \neq 0$ .

We now verify (3): Recalling (84), we note that the Hessian matrix (with respect to  $\varphi, x$ ) of  $\Psi$  is given by

$$(110) \quad \mathcal{H}\Psi(\varphi, x; \tau, z) = \mathcal{H}\Psi_{\mathcal{A}}(\varphi, x; e^{-\tau/2} \cdot z),$$

and hence

$$(111) \quad \mathcal{H}\Psi|_{(0,x)} = \begin{pmatrix} \mathbf{C} & i\mathbf{I} \\ i\mathbf{I} & 0 \end{pmatrix}$$

where

$$(112) \quad C_{jk} = \frac{\partial^2 \Psi}{\partial \varphi_j \partial \varphi_k} \Big|_{(0,x)} = p(-I_j \delta_j^k + I_j I_k), \quad (I_1, \dots, I_m) = \mu(e^{-\tau/2} \cdot z).$$

**Lemma 4.7.** *The matrix  $\mathbf{C}$  given by (112) is strictly negative definite.*

*Proof.* Suppose that  $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m \setminus \{0\}$  is arbitrary. Let  $v = (v_1, \dots, v_m)$ ,  $w = (w_1, \dots, w_m)$  be given by  $v_j = \sqrt{I_j}$ ,  $w_j = \sqrt{I_j} \lambda_j$ . Since  $e^{-\tau/2} \cdot z \in (\mathbb{C}^*)^m$ ,  $I_j \neq 0$  for  $1 \leq j \leq m$ , and hence  $w \neq 0$ . Furthermore  $\|v\|^2 = \sum I_j < 1$ . Therefore

$$\frac{1}{p} \sum_{j,k=1}^m C_{jk} \lambda_j \lambda_k = -\|w\|^2 + (v \cdot w)^2 < -\|v\|^2 \|w\|^2 + (v \cdot w)^2 \leq 0.$$

□

Changing variables, we see that  $\mathcal{H}_y \tilde{\Psi}(0, 0)$  is of the form

$$\begin{pmatrix} 0 & i\mathbf{B} \\ i\mathbf{B}^t & \mathbf{C} \end{pmatrix}$$

where  $\mathbf{B}$  is an  $r \times m$  matrix of maximal rank  $r$  and  $\mathbf{C}$  is nonsingular; hence  $\mathcal{H}_y \tilde{\Psi}(0, 0)$  is nonsingular, verifying hypothesis (3).

To verify (1), we must assume that  $z$  is **not a transition point**. By Lemma 4.5,  $\tau_z$  is not in the boundary of the normal cone to  $F$ , and hence

$$\langle \tau_z, x - q(z) \rangle < 0 \quad \text{for all } x \in P \setminus \bar{F}.$$

Let  $q^j = L^{-1}(E^j) \in \Gamma$ , for  $1 \leq j \leq n-r$  (where  $E^j = (\delta_1^j, \delta_2^j, \dots, \delta_n^j) \in \mathbb{R}^n$ ); i.e.,  $q^j$  lies on the  $j$ -th edge of  $\Gamma$ . Let  $\eta_j = q^j - q(z)$ . Then

$$\partial \tilde{\Psi} / \partial \xi_j = \nabla_{\eta_j} \Psi \equiv \langle \tau_z, v_j \rangle < 0 \quad (1 \leq j \leq n-r),$$

and hence hypothesis (1) holds.



Therefore, by (90), Lemma 4.6 (with  $k = n - r$ ,  $s = m + r$ ) and (109),

$$(113) \quad \mathcal{I}_N^1(z) = N^{\frac{r-m}{2}} e^{N\Psi(0, q(z); \tau_z, z)} [c_0 + c_1 N^{-1} + \dots + c_l N^{-l} + O(N^{-l-1})].$$

We let

$$b_P(z) = -\Psi(0, q(z); \tau_z, z) = -\langle q(z), \tau_z \rangle + p \log \left( \frac{1 + \|z\|^2}{1 + \|e^{-\tau_z/2} \cdot z\|^2} \right)$$

as in (16).

We now show that  $\mathcal{I}_N^2$  decays faster than  $\mathcal{I}_N^1$ : Recall that

$$\mathcal{I}_N^2 = \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} \xi_2(\varphi) \left( \frac{1 + \sum_{j=1}^m e^{-\tau_j - i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right)^{pN} \chi_{NP}(e^{\tau_z + i\varphi}) d\varphi,$$

$$\xi_2(\varphi) = 0 \text{ for } \text{dist}_{\mathbf{T}^m}(\varphi, 0) \geq \varepsilon,$$

where we write  $\tau_z = (\tau_1, \dots, \tau_m)$ . We choose  $\lambda = \lambda(\varepsilon, z) > 0$  such that

$$\left| \frac{1 + \sum_{j=1}^m e^{-\tau_j - i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right| \leq e^{-\lambda} \frac{1 + \sum_{j=1}^m e^{-\tau_j} |z_j|^2}{1 + \|z\|^2} \text{ for } \text{dist}_{\mathbf{T}^m}(\varphi, 0) \geq \varepsilon.$$

Therefore,

$$(114) \quad |\mathcal{I}_N^2| \leq e^{-Np\lambda} \left( \frac{1 + \sum_{j=1}^m e^{-\tau_j} |z_j|^2}{1 + \|z\|^2} \right)^{pN} \chi_{NP}(e^{\tau_z}).$$

Furthermore,

$$(115) \quad \chi_{NP}(e^{\tau_z}) = \sum_{\alpha \in NP} e^{N\langle \tau_z, \alpha/N \rangle} \leq \#(NP) e^{N\langle \tau_z, q(z) \rangle} \leq (Np)^n e^{N\langle \tau_z, q(z) \rangle},$$

since  $\alpha/N \in P$  and  $\tau_z$  is in the normal cone to  $P$  at  $q(z)$ . Combining (114) and (115), we obtain

$$(116) \quad |\mathcal{I}_N^2| \leq (Np)^n e^{-Np\lambda} e^{-Nb_P(z)}.$$

The asymptotic expansion of part (ii) of Theorem 4.1 (for  $P$  simple) with  $b_P(z)$  given by (16) now follows from (89), (113) and (116).

**4.2.3. The decay function.** In this section we verify that the function  $b_P$  given by (16) satisfies parts (c) and (e) of Theorem 4.1(ii) for any convex polytope  $P$ .

We first note that when  $z \in \mathcal{A}_P$ , we have  $\tau_z = 0$  and hence if we extend (16) to all of  $(\mathbb{C}^*)^m$ , we then have  $b_P = 0$  on  $\mathcal{A}_P$ . Furthermore, since  $z \mapsto \tau_z$  is easily seen to be continuous on  $(\mathbb{C}^*)^m$  and  $\mathcal{C}^\infty$  on each  $\overline{\mathcal{R}_F}$ , the same holds for  $b_P(z)$ . We now show that  $b_P$  is  $\mathcal{C}^1$  by computing its derivative. Recalling that  $b_P(z) = -\Psi(0, q(z); \tau_z, z)$ , we have

$$-db_P = [d_x \Psi(0, x; \tau, z) \cdot Dq(z) + d_\tau \Psi(0, x; \tau, z) \cdot D\tau_z + d_z \Psi(0, x; \tau, z)]_{x=q(z), \tau=\tau_z}.$$

Here,  $D\tau_z$  and  $Dq(z)$  are only piecewise continuous, being discontinuous at transition points. However, from (91) and the fact that  $(0, q(z))$  is a critical point of  $\Psi$ , we have

$$d_\tau \Psi(0, x; \tau, z)|_{\{x=q(z), \tau=\tau_z\}} = -i d_\varphi \Psi(\varphi, x; \tau, z)|_{\{\varphi=0, x=q(z), \tau=\tau_z\}} = 0.$$

Furthermore, since  $d_x \Psi|_{\varphi=0, \tau=\tau_z} = \tau_z$  and  $\tau_z$  is perpendicular to the face containing  $q(z)$ , we also have

$$d_x \Psi(0, x; \tau, z)|_{\{x=q(z), \tau=\tau_z\}} \cdot Dq(z) = \tau_z \cdot Dq(z) = 0.$$

Therefore

$$(117) \quad db_P = -d_z \Psi(0, x; \tau, z)|_{\{x=q(z), \tau=\tau_z\}} \in \mathcal{C}^0((\mathbb{C}^*)^m),$$

completing the proof of (e).

Property (c) is a special case of the following lemma on the dependence of the decay function  $b_P(z)$  on the polytope.

**Lemma 4.8.** *Let  $P' \subset P \subset p\Sigma$  be convex polytopes in  $\mathbb{R}^m$ . Suppose that  $z \in (\mathbb{C}^*)^m$  and let  $q(z) \in P$  be given as in (16). Then*

$$b_{P'}(z) \geq b_P(z),$$

with equality if and only if  $q(z) \in P'$ .

*Proof.* Fix  $z \in (\mathbb{C}^*)^m$  and consider the function

$$(118) \quad b(x; z) := -\Psi(0, x; \tau(x), z) = -\langle x, \tau(x) \rangle + p \log \left( \frac{1 + \|z\|^2}{1 + \|e^{-\tau(x)/2} \cdot z\|^2} \right), \quad x \in \Sigma^\circ,$$

where  $\tau(x) \in \mathbb{R}^m$  is given by

$$(119) \quad p\mu(e^{-\tau(x)/2} \cdot z) = x.$$

Note that  $\tau(q(z)) = \tau_z$  and  $b(q(z); z) = b_P(z)$ . Hence it suffices to show that  $b(x; z) > b(q(z); z)$  for all  $x \in P \setminus \{q(z)\}$ .

Equation (119) yields  $e^{-\tau_j(x)} |z_j|^2 = \frac{x_j}{1 - \sum_k x_k}$ , and therefore

$$(120) \quad \tau_j(x) = \log |z_j|^2 + \log(1 - \sum_k x_k) - \log x_j.$$

Since

$$d_\tau \Psi|_{\{\varphi=0, \tau=\tau(x)\}} = \frac{1}{i} d_\varphi \Psi|_{\{\varphi=0, \tau=\tau(x)\}} = x - p\mu(e^{-\tau(x)/2} \cdot z) = 0,$$

we conclude from (118) that

$$(121) \quad d_x b = -d_x \Psi|_{\{\varphi=0, \tau=\tau(x)\}} = -\tau(x),$$

and hence

$$(122) \quad \frac{\partial^2 b}{\partial x_j \partial x_k} = -\frac{\partial \tau_k}{\partial x_j} = \frac{1}{1 - \sum_{l=1}^m x_l} + \delta_j^k \frac{1}{x_j}.$$

Therefore, the Hessian  $(\frac{\partial^2 b}{\partial x_j \partial x_k})$  is positive definite. (It coincides with the derivative  $\mathcal{L}'(x)$  from §4.2.1.)

Now suppose that  $x^0 \in P$ ,  $x^0 \neq q(z)$ . Let

$$f(t) = b(q(z) + tv; z), \quad v = x^0 - q(z).$$

Then by (121),  $f'(0) = -\langle v, \tau_z \rangle \geq 0$ , and by the positivity of the Hessian,  $f''(t) > 0$ . Therefore  $b(x^0; z) - b(q(z); z) = f(1) - f(0) > 0$ .  $\square$

Now let  $z \in (\mathbb{C}^*)^m \setminus \overline{\mathcal{A}_P}$ . We apply Lemma 4.8 with  $P', P$  replaced with  $P, p\Sigma$ , respectively. Since  $q_{p\Sigma}(z) = p\mu(z) \notin P$  by assumption, we conclude by Lemma 4.8 that

$$b_P(z) > b_{p\Sigma}(z) = 0.$$

Thus we have shown that the function  $b_P$  satisfies (c) and (e) for all convex polytopes  $P$ .

4.2.4. *Precise asymptotics on the classically allowed region.* In this section, we prove part (i) of Theorem 4.1. We let  $P$  be a convex integral polytope (not necessarily simple). By (50),

$$\Pi_{Np\Sigma}(z, z) = \sum_{\alpha \in Np\Sigma} \frac{1}{\|\chi_\alpha\|^2} |\widehat{\chi}_\alpha^{Np}(z)|^2 = \frac{(Np+m)!}{(Np)!} = \prod_{j=1}^m (Np+j),$$

and we then have by (53),

$$\Pi_{|NP}(z, z) = \prod_{j=1}^m (Np+j) - R_N(z),$$

where

$$(123) \quad R_N(z) = \sum_{\alpha \in Np\Sigma \setminus NP} \frac{1}{\|\chi_\alpha\|^2} |\widehat{\chi}_\alpha^{Np}(z)|^2.$$

To simplify our computations, we introduce the renormalized monomials

$$(124) \quad \begin{aligned} \widehat{m}_\alpha^{Np}(z) &:= \left[ \frac{(Np)!}{(Np+m)!} \right]^{\frac{1}{2}} \frac{\widehat{\chi}_\alpha(z)}{\|\chi_\alpha\|} = \binom{Np}{\alpha}^{\frac{1}{2}} \widehat{\chi}_\alpha^{Np}(z) \\ &= \binom{Np}{\alpha}^{\frac{1}{2}} \frac{z^\alpha}{(1+\|z\|^2)^{Np/2}}, \quad |\alpha| \leq Np. \end{aligned}$$

Let  $\alpha \in Np\Sigma \setminus NP$  be fixed. We easily check that

$$(125) \quad |\widehat{m}_\alpha^{Np}(z)|^2 = \frac{(Np)!}{(Np+m)!} \Pi_{\{\alpha\}}(z, z) = \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} e^{\Psi_N(\varphi, \alpha; \tau, z)} d\varphi,$$

where the phase  $\Psi_N$  is given by

$$(126) \quad \Psi_N(\varphi, \alpha; \tau, z) = \langle \tau + i\varphi, \alpha \rangle + Np \log \left( \frac{1 + \sum_{j=1}^m e^{-\tau_j - i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right).$$

(Equations (125)–(126) also follow from (53) and (87) with  $P = \{\alpha\}$  and  $p$  replaced by  $Np$ ;  $\Psi_N$  is the phase (91) with  $p$  replaced by  $Np$ .) Since  $\Re \Psi_N(\varphi, x; \tau, z) \leq \Psi_N(0, x; \tau, z)$ , we have

$$(127) \quad |\widehat{m}_\alpha^{Np}(z)|^2 \leq e^{\Psi_N(0, \alpha; \tau, z)}.$$

Choosing  $\tau$  such that  $Np\mu(e^{-\tau/2} \cdot z) = \alpha$  and recalling (118), we have

$$(128) \quad -\Psi_N(0, \alpha; \tau, z) = Nb \left( \frac{\alpha}{N}; z \right) = Nb_{\{\frac{\alpha}{N}\}}(z).$$

By Lemma 4.8,  $b_{\{x\}}(z) > b_P(z) = 0$  for  $x \in p\Sigma \setminus P^\circ$  and  $z \in \mathcal{A}_P$  (i.e.,  $p\mu(z) \in P^\circ$ ).

Now let  $K$  be a compact subset of  $\mathcal{A}_P$ , and let

$$\lambda_K = \frac{1}{2} \inf \{ b_{\{x\}}(z) : x \in p\Sigma \setminus P^\circ, z \in K \} > 0.$$

Therefore,

$$(129) \quad |\widehat{m}_\alpha^{Np}(z)|^2 \leq e^{Nb_{\{\frac{\alpha}{N}\}}(z)} \leq e^{-2\lambda_K N}, \quad \text{for } z \in K, \alpha \in Np\Sigma \setminus NP, N \geq 1.$$

Since  $\#(Np\Sigma \setminus NP) \leq (Np)^m$ , we conclude from (123) and (129) that  $\|R_N\|_{C^0(K)} = O(e^{-\lambda_K N})$ .

To verify the  $\mathcal{C}^1$  estimate, we recall from (51) that

$$|\widehat{m}_\alpha^{Np}(z)|^2 = \binom{Np}{\alpha} h(z)^N |z^\alpha|^2, \quad h(z) = (1 + \|z\|^2)^{-p}.$$

Differentiating with respect to the variables  $\rho_j = \log |z_j|$ , we obtain

$$d_\rho(|\widehat{m}_\alpha^{Np}(z)|^2) = \binom{Np}{\alpha} h(z)^N |z^\alpha|^2 [2\alpha + Nd_\rho \log h] = |\widehat{m}_\alpha^{Np}(z)|^2 [2\alpha + Nd_\rho \log h],$$

and hence the  $\mathcal{C}^1$  estimate follows from the  $\mathcal{C}^0$  estimate. Differentiating repeatedly, we obtain all the  $\mathcal{C}^k$  estimates, completing the proof of part (i) of Theorem 4.1.

4.2.5. *Nonsimple polytopes.* In this section, we prove the general case of part (ii) of Theorem 4.1 by a reduction to the simple case.

Let  $P$  be a nonsimple convex integral polytope of dimension  $n$ . Consider a face  $F$  of  $P$  of dimension  $r$  and let  $z \in \mathcal{R}_F^\circ$ . We say that  $P$  is simple at  $F$  if  $\#\mathcal{J}(F) = n - r$ . (Recall the notation from §2.1. Note that we always have  $\#\mathcal{J}(F) \geq n - r$ . A polytope is simple  $\iff$  it is simple at each of its vertices  $\iff$  it is simple at each of its faces.)

We first consider the case where  $P$  is simple at  $F$ . (This is always the case when  $\text{codim } F \leq 2$ .) We construct a simple integral polytope  $Q \supset P$  coinciding with  $P$  near  $q(z)$  as follows: Recall that  $q(z) \in F$  (by the definition of  $\mathcal{R}_F$ ), and consider the ‘barrier cone’  $B_F(P)$  generated by elements of the form  $x - q(z)$ , where  $x \in P$ . Choose  $L \in \text{GL}(m, \mathbb{Q})$  such that

$$L : B_F(P) \approx \mathbb{R}_{\geq 0}^{n-r} \times \mathbb{R}^r \times \{0\}^{m-n}.$$

(E.g., let  $L$  map the interior normals  $\{u_j : j \in \mathcal{J}_F\}$  at  $F$  to the first  $n - r$  standard basis vectors in  $\mathbb{R}^m$  and map a basis for the tangent space of  $F$  to the next  $r$  standard basis vectors.) We then let

$$Q = q(z) + L^{-1}([0, M]^{n-r} \times [-M, M]^r \times \{0\}^{m-n}),$$

where  $M \in \mathbb{Z}^+$  is chosen so that the vertices of  $Q$  lie in  $\mathbb{Z}^m$ . By the construction,  $B_F(Q) = B_F(P)$  and hence  $Q$  coincides with  $P$  near  $q(z)$ . We then replace  $M$  with a sufficiently high multiple if necessary, so that  $Q \supset P$ .

Since the integral polytope  $Q$  may not be contained in  $p\Sigma$ , we consider the convex rational polytope  $Q' := Q \cap p\Sigma$ . Recalling (124), we have

$$\Pi_{|NP}(z, z) = \Pi_{|NQ'}(z, z) - \frac{(Np + m)!}{(Np)!} \sum_{\alpha \in NQ' \setminus NP} |\widehat{m}_\alpha^{Np}|^2,$$

where by the Szegő kernel  $\Pi_{|NQ'}$ , we mean  $\Pi_{|K_N}$ , where  $K_N$  is the convex hull of  $\mathbb{Z}^N \cap NQ'$ . By (127)–(128), for  $\alpha \in (NQ' \setminus NP) \cap \mathbb{Z}^m$ , we have

$$|\widehat{m}_\alpha^{Np}|^2 \leq \exp(-Nb_{\{\alpha/N\}}(z)) \leq e^{-NC},$$

where

$$C = \inf\{b_{\{x\}}(z) : x \in Q' \setminus P\}.$$

Since  $Q'$  coincides with  $P$  at  $q(z)$ ,  $\tau_z$  is in the normal cone to  $Q'$  at  $q(z)$  and hence  $q(z) = q_{Q'}(z)$ . Therefore by Lemma 4.8,  $b_{\{x\}}(z) > b_{Q'}(z) = b_P(z)$  for all  $x \in Q'$ ,  $x \neq q(z)$ , and hence  $C > b_P(z)$ . Since  $\#(NQ') \leq \frac{1}{m!}(Np)^m$ , it follows that

$$\|\Pi_{|NP}(z, z) - \Pi_{|NQ'}(z, z)\|_{\mathcal{C}^0(K)} = O(e^{-NC'}), \quad b_P(z) < C' < C, \quad K \subset\subset \mathcal{R}_F^\circ.$$

The similar  $\mathcal{C}^l$  estimates follow as in the proof of Theorem 4.1(i) in §4.2.4.

Thus, to obtain the asymptotic expansion of Theorem 4.1(ii) for  $P$ , it suffices to obtain the same expansion for  $Q'$ . To do this, we make the simple observation that

$$(130) \quad \begin{aligned} \Pi_{|NQ'}(z, z) &= \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} \Pi_{NP}(z, e^{i\varphi} \cdot z) \chi_{NQ'}(e^{i\varphi}) d\varphi \\ &= \frac{1}{(2\pi)^m} \int_{\mathbf{T}^m} \Pi_{NP}(z, e^{i\varphi} \cdot z) \chi_{NQ}(e^{i\varphi}), \end{aligned}$$

since lattice points in  $NQ \setminus NP\Sigma = NQ \setminus NQ'$  do not contribute to the integral. By repeating word for word the argument at the beginning of §4.2 and in §4.2.2, with (130) in place of (79), we obtain the expansion

$$\Pi_{|NQ'}(z, z) = N^{\frac{m+r}{2}} e^{-Nb_P(z)} [c_0^F(z) + c_1^F(z)N^{-1} + \cdots + c_k^F(z)N^{-k} + R_k^F(z)],$$

with  $c_0^F > 0$ , and hence part (ii) of Theorem 4.1 holds for  $P$ .

Now suppose that  $P$  is not simple at  $F$ . We subdivide the barrier cone  $B_F(Q)$  into simple rational cones  $\{B^k : k = 1, \dots, s\}$  so that  $B^k \supset B_F(F)$  for all  $k$  and intersections of any number of the  $B^k$  are faces of each of them. (Note that  $B_F(F) = T_F$ .) For a nonempty subset  $\mathcal{E}$  of  $\{1, \dots, s\}$ , let  $P_{\mathcal{E}} = P \cap \bigcap_{k \in \mathcal{E}} (q(z) + B^k)$ , which is simple at  $F$ . Since  $\tau_z$  is in the interior of the normal cone of  $P$  at  $q(z)$ ,  $\tau_z$  is also in the interior of the normal cone at  $q(z)$  of each of the  $P_{\mathcal{E}}$ . (In particular,  $z$  is not a transition point for any of the  $P_{\mathcal{E}}$ .) Thus by the case proven above (which does not use the fact that the vertices of  $P$  are lattice points, only that the vertices of  $Q'$  lie in  $\mathbb{Z}^m$ ), we have the expansion

$$(131) \quad \Pi_{|NP_{\mathcal{E}}}(z, z) = N^{\frac{m+r}{2}} e^{-Nb_P(z)} [c_0^{\mathcal{E}}(z) + c_1^{\mathcal{E}}(z)N^{-1} + \cdots + c_k^{\mathcal{E}}(z)N^{-k} + R_k^{\mathcal{E}}(z)], \quad z \in \mathcal{R}_F^{\circ},$$

where  $c_0^{\mathcal{E}} > 0$ . Note that  $b_P(z)$  is the same for all of the  $P_{\mathcal{E}}$ . Then by the inclusion-exclusion principle

$$(132) \quad \Pi_{|NP}(z, z) = \sum_{\mathcal{E}} (-1)^{\#\mathcal{E}-1} \Pi_{|NP_{\mathcal{E}}}(z, z) \quad (\emptyset \neq \mathcal{E} \subset \{1, \dots, s\}).$$

The desired asymptotic expansion for  $\Pi_{|NP}(z, z)$  now follows from (131)–(132), by noting that  $\Pi_{|NP}(z, z) \geq \Pi_{|NP_{\{1\}}}(z, z)$  and hence  $c_0^F(z) \geq c_0^{\{1\}}(z) > 0$ .

The proof of Theorem 4.1 is now complete.

4.2.6. *Proof of Proposition 4.2.* By Theorem 4.1(ii),

$$(133) \quad \frac{1}{N} \log \Pi_{|NP}(z, z) \rightarrow -b_P(z),$$

for all nontransition points  $z$ . We use the log coordinates  $\rho_j + i\theta_j = \log z_j$ , so that

$$(134) \quad u_N(z) = \frac{1}{N} \log \Pi_{|NP}(z, z) + p \log(1 + \|z\|^2) = \frac{1}{N} \log \sum_{\alpha \in NP} \binom{Np}{\alpha} e^{2\langle \alpha, \rho \rangle}, \quad z \in (\mathbb{C}^*)^m.$$

Thus for all nontransition points  $z$ , we have

$$(135) \quad u_N(z) \rightarrow u_{\infty}(z) := p \log(1 + \|z\|^2) - b_P(z).$$

We must show that convergence of (135) also holds at the transition points and is uniform on compact subsets of  $(\mathbb{C}^*)^m$ . From (134), we obtain

$$(136) \quad d_{\rho} u_N = \frac{1}{N} d_{\rho} \log \sum_{\alpha \in NP} \binom{Np}{\alpha} e^{2\langle \alpha, \rho \rangle} = \frac{2 \sum_{\alpha \in NP} \binom{Np}{\alpha} e^{2\langle \alpha, \rho \rangle} \alpha}{N \sum_{\alpha \in NP} \binom{Np}{\alpha} e^{2\langle \alpha, \rho \rangle}}.$$

Since  $\|\alpha\| \leq Np$  for all  $\alpha \in NP$ , we therefore have the uniform upper bound

$$\|d_\rho u_N\|_{\mathcal{L}^\infty((\mathbb{C}^*)^m)} \leq \frac{\sum_{\alpha \in NP} \binom{Np}{\alpha} e^{2\langle \alpha, \rho \rangle} 2\|\alpha\|}{N \sum_{\alpha \in NP} \binom{Np}{\alpha} e^{2\langle \alpha, \rho \rangle}} \leq 2p.$$

Since  $\{u_N(z^0)\}$  converges for any nontransition point  $z^0$ , it follows that  $\{u_N\}$  is uniformly bounded and uniformly equicontinuous on compact sets. Therefore it converges uniformly on compact sets in  $(\mathbb{C}^*)^m$ .  $\square$

## 5. DISTRIBUTION OF ZEROS

**5.1. Expected zero current and the Szegő kernel.** In order to deduce Theorems 1.3–1.4 from Proposition 4.2, we need to relate the expected zero current to the conditional Szegő kernel  $\Pi|_P$ . This relationship is given by the following result:

**Proposition 5.1.** *The expected zero current of  $k$  independent random polynomials  $f_j \in \text{Poly}(P_j)$ ,  $1 \leq j \leq k$ , is given by*

$$\mathbf{E}_{|P_1, \dots, P_k}(Z_{f_1, \dots, f_k}) = \bigwedge_{j=1}^k \mathbf{E}_{|P_j}(Z_{f_j}) = \bigwedge_{j=1}^k \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log \Pi_{|P_j}(z, z) + p_j \omega_{\text{FS}} \right) \quad \text{on } (\mathbb{C}^*)^m.$$

Recall that  $\omega_{\text{FS}} = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + \|z\|^2)$ . The case  $k = 1$  of the proposition is essentially the same as in [SZ1, Prop. 3.1], but neither the statement nor the proof there covers the application we need. We remark that Proposition 5.1 is a special case of a more general statement for the zeros of general linear systems on compact Kähler manifolds, which we discuss in [SZ3].

We recall that the space  $\mathcal{D}'^{k,l}(Y)$  of  $(k, l)$ -currents on an  $n$ -dimensional complex manifold  $Y$  consists of the continuous linear functionals on the space  $\mathcal{D}^{n-k, n-l}(Y)$  of  $\mathcal{C}^\infty$  compactly-supported  $(n-k, n-l)$ -forms on  $Y$ . We say that a sequence  $\Psi_j \in \mathcal{D}'^{k,l}(Y)$  converges weakly to  $\Psi_0$  if  $(\Psi_j, \varphi) \rightarrow (\Psi_0, \varphi)$  for all test forms  $\varphi \in \mathcal{D}^{n-k, n-l}(Y)$ .

We begin the proof with the following lemma, which covers the  $k = 1$  case.

**Lemma 5.2.** *Let  $P$  be a convex polytope in  $p\Sigma$  (or more generally, an arbitrary subset of  $p\Sigma \cap \mathbb{Z}^m$ ). Let  $Y$  be an algebraic submanifold of  $(\mathbb{C}^*)^m$ , and let  $h \in \mathcal{C}^\infty(Y)$  be given by*

$$h(z) = \log \Pi|_P(z, z) + p \log(1 + \|z\|^2), \quad z \in Y.$$

Then

$$\mathbf{E}_{|P}(Z_{f|Y}) = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} h \in \mathcal{D}'^{1,1}(Y).$$

*Proof.* The proof is essentially the same as the proof of Proposition 3.1 in [SZ1]. Write  $f_Y = f|Y$  for  $f \in \text{Poly}(P)$ . If  $f_Y \neq 0$ , then the current of integration over the zeros of  $f$  is given by the Poincaré-Lelong formula

$$(137) \quad Z_{f_Y} = \frac{\sqrt{-1}}{\pi} \partial \bar{\partial} \log |f_Y| \in \mathcal{D}'^{1,1}(Y).$$

Recalling (6), we can write

$$f_Y = \sum_{\alpha \in P} \frac{c_\alpha}{\|\chi_\alpha\|} \chi_\alpha = \langle c, G \rangle, \quad c = (c_\alpha)_{\alpha \in P}, \quad G = \left( \frac{1}{\|\chi_\alpha\|} \chi_\alpha|_Y \right)_{\alpha \in P}.$$

(The  $\mathcal{L}^2$  norms  $\|\chi_\alpha\|$  are computed over all of  $\mathbb{C}^m$  as before; see (44)–(46).) We then write  $G(z) = |G(z)|u(z)$  so that  $|u| \equiv 1$  and

$$(138) \quad \log |\langle c, G \rangle| = \log |G| + \log |\langle c, u \rangle| .$$

By (137)–(138), we have for any test form  $\varphi \in \mathcal{D}^{n-1, n-1}(Y)$  (where  $n = \dim Y$ ),

$$\begin{aligned} (\mathbf{E}_{|P}(Z_{f_Y}), \varphi) &= \frac{\sqrt{-1}}{\pi} \left( \int_{\mathbb{C}^k} \log |\langle c, G \rangle| d\gamma(c), \partial\bar{\partial}\varphi \right) \\ &= \frac{\sqrt{-1}}{\pi} \left( \partial\bar{\partial} \int_{\mathbb{C}^k} \log |G| d\gamma(c), \varphi \right) + \frac{\sqrt{-1}}{\pi} \left( \partial\bar{\partial} \int_{\mathbb{C}^k} \log |\langle c, u \rangle| d\gamma(c), \varphi \right) , \end{aligned}$$

where  $k = \#P$  and  $d\gamma(c) = \frac{1}{\pi^k} e^{-|c|^2} dc$ . Upon integration in  $c$ , the second term becomes constant in  $z$  and  $\partial\bar{\partial}$  kills it. The first term is independent of  $c$  so we may remove the Gaussian integral. Thus

$$(139) \quad \mathbf{E}_{|P}(Z_{f_Y}) = \frac{\sqrt{-1}}{2\pi} \partial\bar{\partial} \log |G|^2 .$$

Recalling (47) and (54), we have

$$\Pi_{|P}(z, z) = (1 + \|z\|^2)^{-p} |G(z)|^2 , \quad z \in Y ,$$

and the formula of the lemma follows.  $\square$

*Proof of Proposition 5.1.* We first note that the expected current is well defined, since for almost all choices of the  $f_j$ ,

$$\int_{|Z_{f_1, \dots, f_k}|} \omega_{\text{FS}}^{m-k} \leq p_1 \cdots p_k ,$$

and hence for each test form  $\varphi \in \mathcal{D}^{m-k, m-k}((\mathbb{C}^*)^m)$ , the function  $(f_1, \dots, f_k) \mapsto (Z_{f_1, \dots, f_k}, \varphi)$  is in  $\mathcal{L}^\infty(\text{Poly}(P_1) \times \cdots \times \text{Poly}(P_k))$ .

We shall verify the current identity by induction on  $k$ . Lemma 5.2 with  $Y = (\mathbb{C}^*)^m$  gives the case  $k = 1$ , so assume that  $k > 1$  and the proposition has been verified for  $k-1$  polynomials. By the inductive assumption, it suffices to show that

$$(140) \quad \mathbf{E}_{|P_1, \dots, P_k}(Z_{f_1, \dots, f_k}) = \mathbf{E}_{|P_1, \dots, P_{k-1}}(Z_{f_1, \dots, f_{k-1}}) \wedge \mathbf{E}_{|P_k}(Z_{f_k}) .$$

Write  $Y = |Z_{f_1, \dots, f_{k-1}}|$ . By Bertini's Theorem, we know that the  $Z_{f_j}$  are smooth and intersect transversally in  $(\mathbb{C}^*)^m$  and hence  $Y$  is smooth (and of codimension  $k-1$ ) in  $(\mathbb{C}^*)^m$  for almost all  $f_1, \dots, f_k$ . Therefore for  $\varphi \in \mathcal{D}^{m-k, m-k}((\mathbb{C}^*)^m)$ , we have

$$(141) \quad (Z_{f_1, \dots, f_k}, \varphi) = \int_{Z_{f_1, \dots, f_k}} \varphi = \int_{Y \cap Z_{f_k}} \varphi = (Z_{f_k}|_Y, \varphi|_Y) ,$$

for almost all  $f_1, \dots, f_k$ . Averaging (141) over  $f_k$  and applying Lemma 5.2, we obtain

$$(142) \quad \begin{aligned} \int_{\text{Poly}(P_k)} (Z_{f_1, \dots, f_k}, \varphi) d\gamma_{p_k|P_k}(f_k) &= \mathbf{E}_{|P_k}(Z_{f_k}|_Y, \varphi|_Y) = \int_Y \mathbf{E}_{|P_k}(Z_{f_k}) \wedge \varphi \\ &= (Z_{f_1, \dots, f_{k-1}}, \mathbf{E}_{|P_k}(Z_{f_k}) \wedge \varphi) , \end{aligned}$$

where

$$\mathbf{E}_{|P_k}(Z_{f_k}) = \frac{\sqrt{-1}}{2\pi} \log \Pi_{|P_k}(z, z) + p\omega_{\text{FS}} \in \mathcal{D}^{1,1}((\mathbb{C}^*)^m) .$$

Averaging (141) over  $f_1, \dots, f_{k-1}$  and applying (142), we obtain

$$\begin{aligned} \mathbf{E}_{|P_1, \dots, P_k}(Z_{f_1, \dots, f_k}, \varphi) &= \mathbf{E}_{|P_1, \dots, P_{k-1}}(Z_{f_1, \dots, f_{k-1}}, \mathbf{E}_{|P_k}(Z_{f_k}) \wedge \varphi) \\ &= (\mathbf{E}_{|P_1, \dots, P_{k-1}}(Z_{f_1, \dots, f_{k-1}}) \wedge \mathbf{E}_{|P_k}(Z_{f_k}), \varphi). \end{aligned}$$

□

**5.2. Asymptotics of zeros: Proofs of Theorems 1.3 and 1.4.** To prove Theorem 1.3, we let  $u_N, u_\infty$  be as in the proof of Proposition 4.2. By Proposition 5.1 (recalling that  $\omega_{\text{FS}} = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log(1 + \|z\|^2)$ ), we have

$$(143) \quad \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_N = \frac{1}{N} \mathbf{E}_{|NP}(Z_f) \quad \text{on } (\mathbb{C}^*)^m,$$

and hence  $u_N$  is plurisubharmonic. By (135) and (143), we conclude that

$$(144) \quad \frac{1}{N} \mathbf{E}_{|NP}(Z_f) \rightarrow \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty = p\omega_{\text{FS}} - \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} b_P,$$

where differentiation is in the sense of currents (see, e.g., [Sh, Chapter I]).

We now show that the current  $\partial \bar{\partial} b_P \in \mathcal{D}'^{1,1}((\mathbb{C}^*)^m)$  is given by a  $(1,1)$ -form with piecewise smooth coefficients; in fact  $\partial \bar{\partial} b_P$  is  $\mathcal{C}^\infty$  on each of the regions  $\mathcal{R}_F$  given by (13). (Equivalently, the Radon measure  $\partial \bar{\partial} b_P \wedge \omega_{\text{FS}}$  does not charge the set of transition points, and hence formula (144) can be interpreted as differentiation in the ordinary sense on the regions  $\mathcal{R}_F$ .) By Theorem 4.1,  $b_P \in \mathcal{C}^1((\mathbb{C}^*)^m)$ ; i.e., if  $z^0 \in \partial \mathcal{R}_F \cap \partial \mathcal{R}_{F'}$ , then the values of  $db_P(z^0)$  computed in the two regions  $\overline{\mathcal{R}_F}$  and  $\overline{\mathcal{R}_{F'}}$  agree. Then for a test form  $\varphi \in \mathcal{D}^{m-1, m-1}((\mathbb{C}^*)^m)$ , we have

$$(145) \quad (\partial \bar{\partial} b_P, \varphi) = \sum_F \int_{\mathcal{R}_F} b_P \partial \bar{\partial} \varphi = \sum_F \int_{\mathcal{R}_F^\circ} \partial \bar{\partial} b_P \wedge \varphi - \sum_F \int_{\partial \mathcal{R}_F} (\bar{\partial} b_P \wedge \varphi + b_P \wedge \partial \varphi).$$

Note that  $\partial \mathcal{R}_F$  consists of  $\mathcal{C}^\infty$  real hypersurfaces (consisting of those points of  $\partial \mathcal{R}_F$  that are contained in the boundary of only one other region  $\mathcal{R}_{F'}$ ) together with submanifolds of real codimension  $\geq 2$ , and hence Stokes' Theorem applies (see, e.g., [Fe, 4.2.14]). Since  $b_P$  and  $\bar{\partial} b_P$  are continuous on  $(\mathbb{C}^*)^m$ , the boundary integral terms in (145) cancel out and we obtain

$$(146) \quad (\partial \bar{\partial} b_P, \varphi) = \sum_F \int_{\mathcal{R}_F^\circ} \partial \bar{\partial} b_P \wedge \varphi = \int_{(\mathbb{C}^*)^m \setminus E} \partial \bar{\partial} b_P \wedge \varphi,$$

where  $E := \bigcup_F \partial \mathcal{R}_F$  is the set of transition points. Formula (146) says that the current  $\partial \bar{\partial} b_P$  is a  $(1,1)$ -form with piecewise smooth coefficients obtained by differentiating  $b_P$  on the regions  $\mathcal{R}_F^\circ$ .

We now let

$$(147) \quad \psi_P = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty = p\omega_{\text{FS}} - \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} b_P$$

on each of the regions  $\mathcal{R}_F$ . By (146), the current  $\psi_P$  is also a piecewise smooth  $(1,1)$ -form; by (144)

$$(148) \quad N^{-1} \mathbf{E}_{|NP}(Z_f) \rightarrow \psi_P \quad (\text{weakly}).$$

We shall show  $\mathcal{L}_{\text{loc}}^1$  convergence of (148) when we prove Theorem 1.4 below. Continuing with the proof of Theorem 1.3, we observe that (ii) is an immediate consequence of (147), since  $b_P = 0$  on the classically allowed region.



To verify (iii), we again use the log coordinates  $\zeta_j = \rho_j + i\theta_j = \log z_j$ , so that

$$(149) \quad u_\infty = p \log \left( 1 + \sum e^{2\rho_j} \right) - b_P(e^\rho).$$

Since  $u_\infty$  is independent of the angle variables  $\theta_j$ , we have

$$(150) \quad \psi_P = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty = \frac{\sqrt{-1}}{8\pi} \sum_{jk} \frac{\partial^2 u_\infty}{\partial \rho_j \partial \rho_k} d\zeta_j \wedge d\bar{\zeta}_k \geq 0.$$

Thus we must show that the Hessian of  $u_\infty(\rho)$  has rank  $r$  at points  $z^0 \in \mathcal{R}_F^\circ$ , where  $r = \dim F$ . From (117) and (91), we have

$$(151) \quad d_\rho b_P = -p d_\rho \log \left( 1 + \sum e^{-\tau_j + 2\rho_j} \right) \Big|_{\tau=\tau_z} + p d_\rho \log \left( 1 + \sum e^{2\rho_j} \right).$$

To simplify (151), we note that for  $\sigma \in \mathbb{R}^m$  we have

$$(152) \quad d_\rho \log \left( 1 + \sum e^{-\sigma_j + 2\rho_j} \right) = \left( \frac{2e^{-\sigma_1 + 2\rho_1}}{1 + \sum e^{-\sigma_j + 2\rho_j}}, \dots, \frac{2e^{-\sigma_1 + 2\rho_1}}{1 + \sum e^{-\sigma_j + 2\rho_j}} \right) \\ = 2\mu(e^{-\sigma/2} \cdot z).$$

Hence by (152) with  $\sigma = \tau_z$  and 0, we have

$$(153) \quad d_\rho b_P = 2p(\mu - \mu \circ \xi),$$

where  $\xi = (\xi_1, \dots, \xi_m) : (\mathbb{C}^*)^m \rightarrow (\mathbb{C}^*)^m$  denotes the map given by

$$(154) \quad \xi(z) = e^{-\tau_z/2} \cdot z = e^{-\tau_z/2 + \rho + i\theta}.$$

By (149), we then obtain

$$(155) \quad d_\rho u_\infty = 2p \mu \circ \xi.$$

Hence the Hessian  $\mathcal{H}_\rho u_\infty$  equals the Jacobian of  $2p\mu \circ \xi$ . Since  $\mu \circ \xi(z) = \frac{1}{p}q(z)$ , we easily see that

$$(156) \quad \mu \circ \xi : \mathcal{R}_F^\circ \rightarrow \frac{1}{p}F,$$

so  $\text{rank } D(\mu \circ \xi) \leq r$ . In fact, (156) is a submersion, so that the rank equals  $r$ . To see that it is a submersion, we recall from the proof of Lemma 4.3 that the ‘lipeomorphism’  $\Phi^{-1} : \Sigma^\circ \rightarrow \mathcal{N}^\circ$  restricts to a diffeomorphism

$$\Sigma^\circ \supset \mu(\mathcal{R}_F^\circ) \xrightarrow{\cong} \frac{1}{p}F \times C_F \subset \mathcal{N}^\circ, \quad \mu(z) \mapsto \left( \frac{1}{p}q(z), \tau_z \right) = (\mu \circ \xi(z), \tau_z),$$

and therefore (156) is a submersion, completing the proof of (iii).

We now prove Theorem 1.4; the case  $k = 1$  of the theorem will then yield part (i) of Theorem 1.3. Let  $P_1, \dots, P_k$  be convex integral polytopes, as in Theorem 1.4. We first show that

$$(157) \quad N^{-k} \mathbf{E}_{|NP}(Z_{f_1, \dots, f_k}) \rightarrow \psi_k := \psi_{P_1} \wedge \dots \wedge \psi_{P_k} \quad \text{weakly.}$$

For  $1 \leq j \leq k$ , we let

$$u_N^j = \frac{1}{N} \log \Pi_{|NP_j}(z, z) + p_j \log(1 + \|z\|^2),$$

so that, recalling (135),

$$u_N^j(z) \rightarrow u_\infty^j(z) := p_j \log(1 + \|z\|^2) - b_{P_j}(z).$$

By (147),

$$\frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty^j = \psi_{P_j} .$$

Recalling Proposition 5.1, we introduce the  $(k, k)$ -forms

$$(158) \quad \kappa_N := \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_N^1 \right) \wedge \cdots \wedge \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_N^k \right) = N^{-k} \mathbf{E}_{|NP_1, \dots, NP_k}(Z_{f_1, \dots, f_k}) .$$

Since  $u_N^j \rightarrow u_\infty^j$  locally uniformly, it follows from the Bedford-Taylor Theorem [BT, Kl] that

$$(159) \quad \kappa_N \rightarrow \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty^1 \right) \wedge \cdots \wedge \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty^k \right) \quad \text{weakly.}$$

In fact, the limit current in (159) is absolutely continuous, and hence it is equal to the locally bounded (piecewise smooth)  $(k, k)$ -form  $\psi_k$ . To see this, we note that by the Bedford-Taylor Theorem,

$$\bigwedge_{j=1}^k (\psi_{P_j} * \varphi_\epsilon) = \bigwedge_{j=1}^k \left[ \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} (u_\infty^j * \varphi_\epsilon) \right] \rightarrow \bigwedge_{j=1}^k \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} u_\infty^j \right) ,$$

where  $\varphi_\epsilon$  denotes an approximate identity. Since the currents  $\psi_{P_j}$  have coefficients in  $\mathcal{L}_{\text{loc}}^\infty$ , the forms  $\bigwedge_{j=1}^k (\psi_{P_j} * \varphi_\epsilon)$  have locally uniformly bounded coefficients; absolute continuity of the limit current follows. The weak limit (157) now follows from (158)–(159).

To complete the proof of Theorem 1.4, we must show that  $\kappa_N \rightarrow \psi_k$  in  $\mathcal{L}^1(K)$  for all compact  $K \subset (\mathbb{C}^*)^m$ . Let  $\epsilon > 0$  be arbitrary, and choose a nonnegative function  $\eta \in \mathcal{D}((\mathbb{C}^*)^m)$  such that  $\eta \equiv 1$  on a neighborhood of  $(E_1 \cup \cdots \cup E_k) \cap K$ , where  $E_j$  is the set of transition points for  $P_j$ , and

$$\int_{(\mathbb{C}^*)^m} \psi_k \wedge \eta \omega_{\text{FS}}^{m-k} < \epsilon .$$

By (157)–(158),

$$\int_{(\mathbb{C}^*)^m} \kappa_N \wedge \eta \omega_{\text{FS}}^{m-k} < 2\epsilon \quad \text{for } N \gg 0 .$$

(Note that the above integrands are nonnegative.) Therefore,

$$\|\kappa_N - \psi_k\|_{\mathcal{L}^1(K)} \leq \|(1 - \eta)(\kappa_N - \psi_k)\|_{\mathcal{L}^1(K)} + \|\eta \kappa_N\|_{\mathcal{L}^1(K)} + \|\eta \psi_k\|_{\mathcal{L}^1(K)} .$$

Since  $\kappa_N$  is a positive  $(k, k)$ -form,

$$\|\eta \kappa_N\|_{\mathcal{L}^1(K)} = \int_K \kappa_N \wedge \eta \omega_{\text{FS}}^{m-k} < 2\epsilon .$$

Similarly,  $\|\eta \psi_k\|_{\mathcal{L}^1(K)} = \int_K \psi_k \wedge \eta \omega_{\text{FS}}^{m-k} < \epsilon$ . Since

$$(1 - \eta)(\kappa_N - \psi_k) \rightarrow 0 \quad \text{uniformly on } K ,$$

it follows that

$$\|\kappa_N - \psi_k\|_{\mathcal{L}^1(K)} \rightarrow 0 ,$$

completing the proof of Theorems 1.3 and 1.4.  $\square$

We now derive our integral formula (38) from (153):

**Proposition 5.3.** *The decay function  $b_P$  of Theorem 1.2 is given by*

$$b_P(z) = \int_0^{\tau_z} \left[ -q(e^{-\sigma/2} \cdot z) + p\mu(e^{-\sigma/2} \cdot z) \right] \cdot d\sigma .$$

*Proof.* Fix a point  $z$  in the classically forbidden region. Then by (153) with the change of variables  $\rho_j = -\sigma_j/2 + \log |z_j|$ , we have

$$(160) \quad \begin{aligned} b_P(e^{-\tau_z/2} \cdot z) - b_P(z) &= \int_0^{\tau_z} d_\sigma b_P(e^{-\sigma/2} \cdot z) \cdot d\sigma \\ &= p \int_0^{\tau_z} [\mu \circ \xi(e^{-\sigma/2} \cdot z) - \mu(e^{-\sigma/2} \cdot z)] \cdot d\sigma , \end{aligned}$$

where the map  $\xi$  is given by (154). By definition,  $q(z) = p\mu \circ \xi(z)$ , and hence  $p\mu \circ \xi(e^{-\sigma/2} \cdot z) = q(e^{-\sigma/2} \cdot z)$ . Furthermore  $b_P(e^{-\tau_z/2} \cdot z) = 0$  since  $e^{-\tau_z/2} \cdot z = \xi(z) \in \partial\mathcal{A}_P$ , and formula (38) then follows from (160).  $\square$

We note that the proof of Theorem 1.4 also gives an asymptotic expansion away from transition points:

**Theorem 5.4.** *Let  $P_1, \dots, P_k$  be convex integral polytopes. Let  $U$  be a relatively compact domain in  $(\mathbb{C}^*)^m$  such that  $\bar{U}$  does not contain transition points for any of the  $P_j$ . Then we have a complete asymptotic expansion of the form*

$$\frac{1}{N^k} \mathbf{E}_{|NP_1, \dots, NP_k}(Z_{f_1, \dots, f_k}) \sim \psi_{P_1} \wedge \dots \wedge \psi_{P_k} + \frac{\varphi_1}{N} + \dots + \frac{\varphi_n}{N^n} + \dots \quad \text{on } U ,$$

with uniform  $\mathcal{C}^\infty$  remainder estimates, where the  $\varphi_j$  are smooth  $(k, k)$ -forms on  $U$ .

*Proof.* By (134) and (158) we have

$$\frac{1}{N^k} \mathbf{E}_{|NP_1, \dots, NP_k}(Z_{f_1, \dots, f_k}) = \bigwedge_{j=1}^k \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \left[ \frac{1}{N} \log \Pi_{|NP_j}(z, z) + p_j \log(1 + \|z\|^2) \right] .$$

The conclusion now follows from the asymptotic expansion of Theorem 4.1.  $\square$

**5.3. Proof of Theorem 1.1.** Theorem 1.1 is an easy consequence of our prior results. Indeed, by Proposition 5.1, we can write  $N^{-m} \mathbf{E}_{|NP}(Z_{f_1, \dots, f_m}) = G_N \omega^m$ , where the  $G_N$  are positive  $\mathcal{C}^\infty$  functions on  $(\mathbb{C}^*)^m$ . We must show the convergence of the sequence of measures

$$d\lambda_N := N^{-m} \mathbf{E}_{|NP}(Z_{f_1, \dots, f_m}) = G_N \omega^m .$$

Now let  $B$  be a Borel subset of  $(\mathbb{C}^*)^m$ . Then

$$\lambda_N(B) = N^{-m} \mathbf{E}_{|NP}(\#\{z \in B : f_1(z) = \dots = f_m(z) = 0\}) .$$

We first consider the case where  $B$  is contained in a compact set  $K \subset (\mathbb{C}^*)^m$ . Then by Theorem 1.4 with  $k = m$  and  $P_1 = \dots = P_k = P$ , it follows that

$$G_N \omega^m \rightarrow \psi_P^m \quad \text{in } \mathcal{L}^1(K) .$$

By Theorem 1.3,  $\psi_P^m = p^m \omega_{\text{FS}}^m$  on  $\mathcal{A}_P$ , and  $\psi_P^m = 0$  on the complement of  $\mathcal{A}_P$  since its rank is less than  $m$  there. Therefore,

$$(161) \quad \lambda_N(B) = \int_K \chi_B G_N \omega^m \rightarrow \int_K \chi_{B \cap \mathcal{A}_P} p^m \omega_{\text{FS}}^m = m! p^m \text{Vol}_{\mathbb{C}\mathbb{P}^m}(B \cap \mathcal{A}_P) .$$

For the general case, let  $\varepsilon > 0$  be arbitrary and choose  $B' \subset B$  such that  $B'$  is relatively compact in  $(\mathbb{C}^*)^m$  and  $\text{Vol}_{\mathbb{C}\mathbb{P}^m}(B \setminus B') < \varepsilon$ . Then by the above

$$\begin{aligned} \liminf_{N \rightarrow \infty} \lambda_N(B) &\geq \liminf_{N \rightarrow \infty} \lambda_N(B') = m!p^m \text{Vol}_{\mathbb{C}\mathbb{P}^m}(B' \cap \mathcal{A}_P) \\ &\geq m!p^m \text{Vol}_{\mathbb{C}\mathbb{P}^m}(B \cap \mathcal{A}_P) - m!p^m \varepsilon. \end{aligned}$$

Since  $\varepsilon > 0$  is arbitrary, we conclude that

$$(162) \quad \liminf_{N \rightarrow \infty} \lambda_N(B) \geq m!p^m \text{Vol}_{\mathbb{C}\mathbb{P}^m}(B \cap \mathcal{A}_P).$$

To obtain the reverse inequality, we recall from the Bernstein-Kouchnirenko Theorem [Be, Ko1, Ko2] that the number of common zeros of  $\{f_1, \dots, f_m\}$  equals  $m!N^m \text{Vol}(P)$  for almost all  $(f_1, \dots, f_m) \in \text{Poly}(P)^m$ , and hence

$$(163) \quad \lambda_N((\mathbb{C}^*)^m) = m! \text{Vol}(P) \quad \text{for all } N \geq 1.$$

It is well known and easy to verify that  $\text{Vol}_{\mathbb{C}\mathbb{P}^m}(\mu^{-1}(\Omega)) = \text{Vol}(\Omega)$  for any  $\Omega \subset \Sigma$ . (This is a special case of a general fact about the moment map in symplectic geometry; see (169) in the Appendix.) Recalling that  $\mathcal{A}_P = \mu^{-1}(\frac{1}{p}P^\circ)$ , we then have

$$(164) \quad \text{Vol}(\mathcal{A}_P) = \frac{1}{p^m} \text{Vol}(P).$$

It follows from (162) with  $B$  replaced by its complement  $B^c = (\mathbb{C}^*)^m \setminus B$  and (163)–(164) that

$$(165) \quad \limsup_{N \rightarrow \infty} \lambda_N(B) = m! \text{Vol}(P) - \liminf_{N \rightarrow \infty} \lambda_N(B^c) \leq m!p^m \text{Vol}_{\mathbb{C}\mathbb{P}^m}(B \cap \mathcal{A}_P).$$

Therefore, (161) holds for a general Borel set  $B \subset (\mathbb{C}^*)^m$ .  $\square$

**5.4. Vanishing of  $\psi_P$  along the normal flow.** The proof of Theorem 1.3(iii) gives us some more information about the expected zero current in the classically forbidden region. Roughly speaking, let  $f \in \text{Poly}(NP)$  be a random polynomial with Newton polytope  $NP$ , with  $N$  large, and let  $z^0 \in |Z_f| \setminus \mathcal{A}_P$ ; then  $|Z_f|$  is highly likely to be close to being tangent to the orbit of the normal flow at  $z^0$ . In particular, if  $z^0$  is in the flow-out of an edge (dimension 1 face) of  $P$ , then  $T_{z^0}(|Z_f|)$  is likely to be a good approximation to the tangent space of the normal flow through  $z^0$ .

To make this statement precise, we define the *complexified normal cone* of a face  $F$ ,

$$\tilde{C}_F := \{\tau + i\theta : \tau \in C_F, \theta \in T_F^\perp\}.$$

(Recall that  $C_F \subset T_F^\perp$ .) We note that  $\tilde{C}_F$  is a semigroup, which acts on  $\mathcal{R}_F$  by the rule  $\eta(z) = e^\eta \cdot z$ ; we call this action the ‘(joint) normal flow.’ The (maximal) orbits of the normal flow are of the form  $\tilde{C}_F \cdot z^0 = \{e^\eta \cdot z^0 : \eta \in \tilde{C}_F\}$ , where  $z^0 \in \mu^{-1}(\frac{1}{p}F)$ . We note that the orbit  $\tilde{C}_F \cdot z^0$  is a complex  $(m-r)$ -dimensional submanifold (with boundary) of  $(\mathbb{C}^*)^m$ . (Indeed,  $(\tilde{C}_F \cdot z^0) \cap \mathcal{R}_F^\circ$  is a submanifold without boundary in  $\mathcal{R}_F^\circ$ .)

**Theorem 5.5.** *Let  $P$  be a convex integral polytope and let  $\psi_P$  be the limit expected zero current of Theorem 1.3. Then  $\psi_P$  vanishes along the orbits of the normal flow.*

*Proof.* Let

$$O := \tilde{C}_F \cdot z^0 = \{e^{\tau+i\theta} \cdot z^0 : \tau \in C_F, \theta \in T_F^\perp\}$$

be a maximal orbit of the normal flow, where  $p\mu(z^0) \in F$ . For  $z = e^{\tau+i\theta} \cdot z^0 \in O$ , we have

$$\mu \circ \xi(z) = \frac{1}{p}q(z) = \frac{1}{p}q(e^\tau \cdot z^0) = \frac{1}{p}q(z^0) = \mu(z^0)$$

and hence  $\mu \circ \xi$  is constant on  $O$ . It then follows from (150) and (155) that  $\psi_P|_O = 0$ .  $\square$

To relate Theorem 5.5 to the explanation above, suppose for example that  $z^0$  is in the flow-out of a 1-dimensional face (edge). Choose coordinates  $w_1, \dots, w_m$  so that the normal flow near  $z^0$  is given by  $w_1 = \text{constant}$ . Since  $\psi_P$  vanishes along the orbits of the flow and has rank 1 near  $z^0$ , we have

$$(166) \quad \frac{1}{N} \mathbf{E}_{|NP}(Z_f) \rightarrow \psi_P = ic(w)dw_1 \wedge d\bar{w}_1, \quad c(w) > 0,$$

near  $z^0$ . For a regular point  $z$  of  $|Z_f|$ , we let  $\eta_f(z) \in T^{*1,0}((\mathbb{C}^*)^m)$  be a unit vector (unique up to the  $S^1$  action) annihilating  $T^{1,0}(|Z_f|)$ . We can then write  $Z_f = \delta_{Z_f}(\frac{1}{2}\eta \wedge \bar{\eta})$ , where  $\delta_{Z_f}$  is the measure given by  $(\delta_{Z_f}, \varphi) = \int_{|Z_f|} \varphi d\text{Vol}_{2m-2}$ . Writing  $\eta = \sum_{j=1}^m a_j(w)dw_j$ , we see by (166) that

$$\frac{1}{N} \mathbf{E}_{|NP} \int_{|Z_f|} |a_j|^2 d\text{Vol}_{2m-2} \rightarrow 0, \quad \text{for } j \geq 2.$$

Thus the expected value of the average distance between the tangent spaces of  $|Z_f|$  and of the normal flow approaches 0 as  $N \rightarrow \infty$ .

**5.5. Amoebas in the plane.** The term ‘amoeba’ was introduced by Gelfand, Kapranov and Zelevinsky [GKZ] to refer to the image under the moment map of a zero set  $Z_{f_1, \dots, f_k}$  of polynomials and has been studied in various contexts (see [FPT, GKZ, Mil, PR] and the references in the survey article by Mikhalkin [Mi2]). The image of a zero set under the moment map  $\mu$  is called a *compact amoeba*, while the image under the map

$$(167) \quad \text{Log} : (\mathbb{C}^*)^m \rightarrow \mathbb{R}^m, \quad (z_1, \dots, z_m) \mapsto (\log|z_1|, \dots, \log|z_m|)$$

is a *noncompact amoeba*, or simply an *amoeba*. Note that  $\text{Log}$  is the moment map for the  $\mathbf{T}^m$  action with respect to the Euclidean symplectic form  $\sum dx_j \wedge dy_j$ , and  $\text{Log} = \frac{1}{2}\mathcal{L} \circ \mu$ , where  $\mathcal{L} : \Sigma^\circ \approx \mathbb{R}^m$  is the diffeomorphism given by (99).

To illustrate what our statistical results can say about amoebas, we consider zero sets in  $(\mathbb{C}^*)^2$ . An amoeba in  $\mathbb{R}^2$  is the image of a plane algebraic curve under the Euclidean moment map  $\text{Log}$ . An example of an amoeba of the form  $\text{Log}(Z_f)$ , where  $f$  is a quartic polynomial in two variables with (full) Newton polytope  $4\Sigma$ , is given in the illustration from [Th] reproduced in Figure 7.

One notices that this amoeba contains 12 ‘tentacles’. By definition, a tentacle on a compact amoeba  $A$  is a connected component of a small neighborhood in  $A$  of  $A \cap \partial\Sigma$ ; the tentacles of a noncompact amoeba correspond to those of the compact amoeba under the diffeomorphism  $\Sigma^\circ \approx \mathbb{R}^m$ .

It is known that there is a natural injective map from the set of connected components (which are convex sets) of the complement of a noncompact amoeba  $A$  to the set  $P \cap \mathbb{Z}^2$  of lattice points of the polytope and that there are amoebas for which each lattice point is assigned to a component of the complement. (This

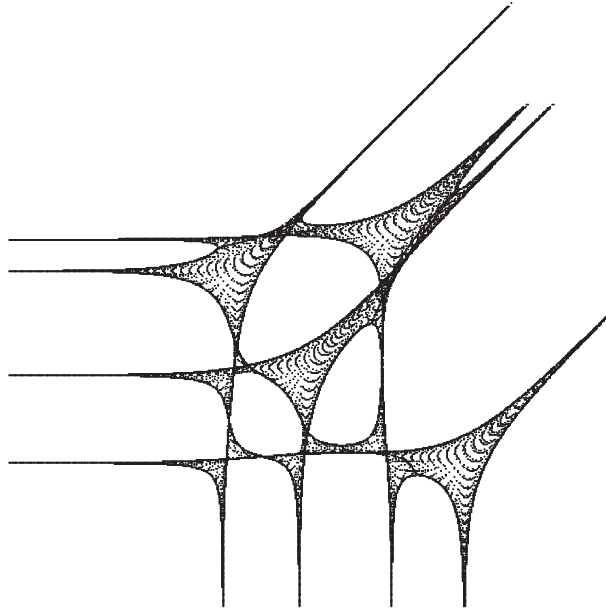


FIGURE 7. An amoeba with polytope  $4\Sigma$ . [From T. Theobald, Computing amoebas, *Experimental Math.* 11:4 (2002), 513–526, with permission of AK Peters.]

fact is also valid in higher dimensions; see [FPT, Mi1].) For a generic 2-dimensional noncompact amoeba with polytope  $P$ , each lattice point in  $\partial P$  corresponds to a distinct unbounded component of  $\mathbb{R}^2 \setminus A$ , and adjacent lattice points correspond to adjacent unbounded components. (The correspondence is given in [Mi1, §3.1] or [FPT].) Each tentacle thus corresponds to a segment connecting 2 adjacent lattice points on  $\partial P$ . Hence the number of tentacles of  $A$  equals the number of points of  $\partial P \cap \mathbb{Z}$ ; this number is called the *length* of  $\partial P$ . (For example, the 12 tentacles in Figure 7 correspond to the 12 segments connecting the 12 lattice points in  $\partial(4\Sigma)$ .)

We can decompose  $\partial P$  into two pieces:  $\partial^\circ P = \partial P \cap p\Sigma^\circ$  and  $\partial^e P = P \cap \partial(p\Sigma)$ . Tentacles corresponding to segments of  $\partial^\circ P$  end (in the compact picture  $\Sigma$ ) at a vertex of  $\Sigma$ , and tentacles corresponding to segments of  $\partial^e P$  are free to end anywhere on the face of  $\Sigma$  containing the segment. We call the latter *free tentacles*, and we say that a free tentacle is a *classically allowed tentacle* if its end is in the classically allowed region  $\mathcal{A}_P$ . For an amoeba  $A$ , we let  $\nu_{\text{AT}}(A)$  denote the number of classically allowed tentacles of  $A$ . It is clear from the above that

$$\nu_{\text{AT}}(A) \leq \#\{\text{free tentacles}\} = \text{Length}(\partial^e P)$$

and that this bound can be attained for any polytope  $P$ . Here, ‘Length’ means the length in the above sense; i.e., the diagonal face of  $p\Sigma$  is scaled to have length  $p$ . As a consequence of Theorem 1.3 (for  $m = 1$ ), we conclude that this maximum is asymptotically the average:

**Corollary 5.6.** *For a convex integral polytope  $P$  in  $\mathbb{R}^2$ , we have*

$$\frac{1}{N} \mathbf{E}_{|NP} \left( \nu_{\text{AT}}(\text{Log}(Z_f)) \right) \rightarrow \text{Length}(\partial^e P).$$

*Proof.* Let  $F_1, F_2, F_3$  denote the facets of  $p\Sigma$  and apply Theorem 1.3(ii) to the 1-dimensional polytopes  $P \cap \bar{F}_j$ ,  $j = 1, 2, 3$ .  $\square$

## 6. APPENDIX: APPROACH USING TORIC GEOMETRY

Although our results and proofs do not depend on the theory of toric varieties, the subject motivated some of our ideas. We describe briefly in this appendix how toric varieties can be used to give a geometric derivation of the Szegő kernel asymptotics of Theorem 4.1.

**6.1. Relationship to toric geometry.** A *toric variety* is a complex algebraic variety  $M$  containing the complex torus  $(\mathbb{C}^*)^m$  as a Zariski-dense open set such that the group action of  $(\mathbb{C}^*)^m$  extends to a  $(\mathbb{C}^*)^m$  action on  $M$ . A toric variety can be constructed from a fan by gluing together the affine varieties arising from the cones in the fan (see [Fu, Chapter 1]). The toric variety  $M_P$  constructed in this way from the fan of a convex integral polytope  $P$  is a projective variety. If  $P$  is simple, then  $M_P$  has orbifold singularities; if  $P$  is Delzant, then  $M_P$  is smooth [De] (see also [Fu, §2.1–2.2]).

For example,  $\mathbb{C}\mathbb{P}^m$  is the toric variety corresponding to the simplex  $\Sigma \subset \mathbb{R}^m$ . Hence, the toric variety corresponding to the square from Example 1 in §1.2 is the product  $\mathbb{C}\mathbb{P}^1 \times \mathbb{C}\mathbb{P}^1$  of projective lines. When  $P$  is the trapezoid in Example 2, the associated toric variety  $M_P$  is the blow up of  $(0, 0, 1) \in \mathbb{C}\mathbb{P}^2$ , i.e.,  $M_P$  is the Hirzebruch surface  $F_1$ . The toric variety corresponding to the polytope in Example 3 is the Hirzebruch surface  $F_n$  (see [Fu, §1.1]).

We shall now assume for simplicity that  $P$  is Delzant and has nonempty interior. In this case,  $M_P$  can be given as the closure of the image of a monomial embedding  $\Phi_P : (\mathbb{C}^*)^m \rightarrow \mathbb{C}\mathbb{P}^{\#P-1}$  (see [GKZ, Chapter 5] or [STZ1]). We also give  $M_P$  the structure of a symplectic or Kähler manifold with symplectic/Kähler form  $\omega_P := \Phi_P^* \omega_{\text{FS}}$ . (The symplectic form  $\omega_P$  depends on the choice of constants defining the monomial embedding  $\Phi_P$ ; see [STZ1].) We define the line bundle  $L_P := \Phi_P^* \mathcal{O}(1)$ , where  $\mathcal{O}(1)$  denotes the hyperplane section bundle on  $\mathbb{C}\mathbb{P}^{\#P-1}$ , and we give  $L_P$  the Hermitian metric obtained by pulling back the Fubini-Study metric on  $\mathcal{O}(1)$  so that  $L_P$  has curvature form  $\omega_P$ .

For example, for the case  $P = \Sigma$ , the toric variety  $M_\Sigma = \mathbb{C}\mathbb{P}^m$ ,  $L_\Sigma = \mathcal{O}(1)$ , and

$$\text{Poly}(N\Sigma) \cong H^0(\mathbb{C}\mathbb{P}^m, \mathcal{O}(N)) = H^0(M_\Sigma, L_\Sigma^N).$$

(Recall that  $H^0(M, L)$  denotes the space of holomorphic sections of  $L$ .) More generally,

$$(168) \quad \text{Poly}(NP) \simeq H^0(M_P, L_P^N) = \Phi_P^* H^0(\mathbb{C}\mathbb{P}^{\#P-1}, \mathcal{O}(1)).$$

The underlying real torus  $\mathbf{T}^m$  action on the symplectic manifold  $(M_P, \omega_P)$  is Hamiltonian with moment map  $\mu_P : M_P \rightarrow P$  fibering  $M_P$  as a singular torus bundle over the polytope  $P$  (see, e.g., [Fu, §4.2]). The volume form on  $M_P$  can be written in terms of the moment map:  $\frac{1}{m!} \omega_P^m = dI \wedge d\theta$ , where  $dI = \mu^*(dx_1 \wedge \cdots \wedge dx_m)$ . Integrating, we obtain

$$(169) \quad \text{Vol}(P) = \int_{M_P} dI \wedge d\theta = \int_{M_P} \frac{1}{m!} \omega_P^m = \frac{1}{m!} \deg[c_1(L_P)^m],$$

which yields Kouchnirenko's Theorem, since  $\deg[c_1(L_P)^m]$  equals the number of points in the intersection of the divisors of  $m$  generic sections in  $H^0(M_P, L_P) \simeq \text{Poly}(P)$ . (This proof of Kouchnirenko's Theorem is from Atiyah [At].)

*Remark.* Using the isomorphism (168), one can view  $\chi_P(e^{i\theta})$  as the character of the torus  $\mathbf{T}^m$  action on the space  $H^0(M_P, L_P)$  of holomorphic sections of the line bundle  $L_P$  over the toric variety  $M_P$  (or equivalently as the equivariant index of the  $\bar{\partial}$  operator on sections of  $L_P$ ). The characters  $\{\chi_{NP}(e^{i\theta})\}$  are the characters of the powers  $L_P^N$  of  $L_P$  and are given by the equivariant Riemann-Roch formula [BV3, Gu] (see also [BP, §8]). When  $\theta = 0$ , the Riemann-Roch formula gives another form of the Ehrhart formula for the number of lattice points of  $NP$ :

$$\begin{aligned} \chi_{NP}(1) &= \#(NP) = \dim \text{Poly}(NP) = \dim H^0(M_P, L_P^N) \\ &= \chi(M_P, L_P^N) = \sum_{k=0}^m \deg [c_1(L_P)^k \cup \text{Todd}_{m-k}(M_P)] \frac{N^k}{k!} \\ (170) \quad &= \frac{\deg[c_1(L_P)^m]}{m!} N^m + \dots + \deg \text{Todd}_m(M_P). \end{aligned}$$

(Note that equations (57) and (170) provide an alternate derivation of (169).)

**6.2. Geometric proof of Theorem 4.1.** We outline here an alternative approach to evaluating the integral (89) by lifting it to  $M_P$ , obtaining a complex oscillatory integral over  $\mathbf{T}^m \times M_P$ . By (64), we have

$$(171) \quad \mathcal{I}_N^1 = \frac{N^m}{(2\pi)^m} \int_{\mathbf{T}^m} d\varphi \cdot \xi_1(\varphi) e^{N\Psi_0(\varphi; \tau, z)} \text{Todd}(\mathcal{F}, N^{-1}\partial/\partial h) \left( \int_{P(h)} e^{N\langle \tau + i\varphi, x \rangle} dx \right) \Big|_{h=0}.$$

We can analyze this integral by lifting to the toric variety  $(M_h, \omega_h)$  associated to the deformed polytope  $P(h)$  given by (60). The underlying complex variety  $M_h = M_P$  is fixed, while the symplectic form is given by

$$\omega_h = \omega_P + \sum_{k=1}^d h_k c_1(L_k),$$

which is affine in  $h$ . Here,  $L_k$  is the line bundle associated to the divisor  $\mu_P^{-1}(\bar{F}_k)$  in  $M_P$ , and  $c_1(L_k)$  is a  $\mathbf{T}^m$ -invariant  $(1, 1)$ -form in the Chern class of  $L_k$ .

The moment map  $\mu_h = (I_1^h, \dots, I_m^h) : M_h \rightarrow P(h)$  is also an affine function in  $h$ . Indeed, let  $X_j$  denote the vector field on  $M_P$  giving the infinitesimal action on  $\mathbf{T}^m$  corresponding to the  $j$ -th unit vector in the Lie algebra  $\mathfrak{t}^m \approx \mathbb{R}^m$ . Then we have

$$\begin{aligned} dI_j^h &= d\mu_h(X_j) = \omega_h(X_j, \cdot) \\ &= \omega_P(X_j, \cdot) + \sum_{k=1}^d h_k c_1(L_k)(X_j, \cdot) = dI_j + \sum_{k=1}^d h_k c_1(L_k)(X_j, \cdot) \end{aligned}$$

and hence

$$I_j^h = I_j + \sum_{k=1}^d h_k J_j^k.$$

Using the fact that

$$\int_{P(h)} f(x) dx = \int_{M_P} f \circ \mu_h \frac{1}{m!} \omega_h^m,$$

we then lift (171) to  $M_h$  to obtain

$$(172) \quad \mathcal{I}_N^1 = \frac{N^m}{(2\pi)^m m!} \text{Todd}(\mathcal{F}, N^{-1}\partial/\partial h) \Big|_{h=0} \int_{\mathbf{T}^m} \int_{M_P} e^{N\Psi_h(\varphi, w; \tau, z)} \xi_1(\varphi) \omega_h^m(w) d\varphi.$$



The phase is now the function on  $\mathbf{T}^m \times M_P$  given by

$$(173) \quad \begin{aligned} \Psi_h(\varphi, w; \tau, z) = & \langle \tau + i\varphi, \mu_P(w) \rangle \\ & + \sum_{k=1}^d h_k \langle \tau + i\varphi, J^k(w) \rangle + p \log \left( \frac{1 + \sum_{j=1}^m e^{-\tau_j - i\varphi_j} |z_j|^2}{1 + \|z\|^2} \right). \end{aligned}$$

The amplitude is a polynomial in  $h$ :

$$(174) \quad \omega_h^m = \omega_P^m + \sum_{1 \leq |\alpha| \leq m} h^\alpha \gamma_\alpha = G(w, h) \omega_P^m.$$

Interchanging the order of integration and Todd differentiation in (172) and then using (69) as before, we obtain

$$(175) \quad \mathcal{I}_N^1 = \frac{N^m}{(2\pi)^m m!} \int_{\mathbf{T}^m} \int_{M_P} e^{N\Psi(\varphi, \mu_P(w); \tau, z)} A(N, \varphi, w; \tau) \omega_P^m(w) d\varphi,$$

with amplitude

$$A(N, \varphi, w; \tau) = \xi_1(\varphi) G(w, N^{-1} \partial / \partial q) \text{Todd}(\mathcal{F}, q)|_{q_k = \langle \tau + i\varphi, J^k(w) \rangle}.$$

The phase is our familiar phase function  $\Psi$  given by (91) lifted to  $M_P$ .

We need to find the critical points where the phase  $\Psi$  has maximal real part. As we determined before, the phase has maximal real part where  $\varphi = 0$  and  $\tau$  is in the normal cone to  $P$  at  $\mu_P(w)$ . As before, we find that the critical point equation  $d_\varphi \Psi(\varphi, \mu_P(w); \tau, z)|_{\varphi=0} = 0$  is equivalent to

$$\frac{1}{p} \mu_P(w) = \mu(e^{-\tau/2} \cdot z).$$

The second critical point equation  $d_w \Psi(0, \mu_P(w); \tau, z) = 0$  reduces to

$$d_w \langle \mu_P(w), \tau \rangle = 0.$$

Since  $\frac{\partial}{\partial w_j} \mu_P|_{w^0}$  is always tangent to the face of  $P$  at  $\mu_P(w^0)$ , this is just the condition that  $\tau$  is orthogonal to the face of  $P$  containing  $\mu_P(w)$ , which is automatically satisfied when  $\tau$  is in the normal cone to  $P$  at  $\mu_P(w)$ . Hence, in order to have critical points,  $\tau$  must equal  $\tau_z$  as before, and we obtain a critical submanifold

$$\mathcal{C}_z := \{(0, w) : \mu_P(w) = q(z)\}.$$

Note that  $\dim \mathcal{C}_z = \text{codim } F$ , where  $F$  is the face of  $P$  containing  $q(z)$ .

By a calculation using the methods of [SZ2, §3.2.2] (where a more complicated phase function was used), one can show that the normal Hessian is nondegenerate whenever  $z$  is not a transition point. The asymptotic expansion of Theorem 4.1 then follows by the method of stationary phase with nondegenerate critical submanifolds.

*Remark.* We may also express polytope characters through the Szegő kernel  $\Pi^{M_P}$  of  $H^0(M_P, L_P^N)$  by means of the obvious identity

$$(176) \quad \chi_{NP}(e^{i\varphi}) = \int_{M_P} \Pi_N^{M_P}(e^{i\varphi} \cdot w, w) d\text{Vol}_{M_P}(w).$$

In an article with T. Tate [STZ1], we describe an explicit construction of the Szegő kernel of a toric variety, which we use to obtain a formula for the polytope character  $\chi_{NP}(e^{i\varphi})$  as a complex oscillatory integral over the toric variety  $M_P$ .

## INDEX OF NOTATION

|                                |            |   |           |
|--------------------------------|------------|---|-----------|
| $\binom{p}{\alpha}$            | (48)       | $\mathcal{E}_Q$                                 | (98)      |
| $\langle \cdot, \cdot \rangle$ | (44)       | $\text{Flow}(x)$                                | (12)      |
| $\gamma_P$                     | (6)        | $ f(z) _{\text{FS}}$                            | (11)      |
| $\gamma_{P P}$                 | (7)        | $\mathcal{I}_N^1, \mathcal{I}_N^2$              | (89)      |
| $\kappa_N$                     | (158)      | $\mathcal{J}(x)$                                | (42)      |
| $\mu$                          | (8)        | $\ell_j$  | (40)      |
| $\Pi_P$                        | (50), (52) | $\mathcal{L}$                                   | (99)      |
| $\Pi_{ P}$                     | (53), (54) | $\text{Log}$                                    | (167)     |
| $\Sigma$                       | §1.1       | $\mathcal{M}_x$                                 | (17)      |
| $\tau_z$                       | (14)–(15)  | $\mathcal{N}(Q)$                                | §4.2.1    |
| $\chi_\alpha$                  | (1)        | $P_f$   | (3)       |
| $\ \chi_\alpha\ $              | (46)       | $P(h)$  | (60)      |
| $\widehat{\chi}_\alpha^p$      | (49), (51) | $\text{Poly}(P)$                                | (4)       |
| $\chi_{NP}$                    | (55)       | $q(z)$  | (14)–(15) |
| $\psi_P$                       | (23)       | $\mathcal{R}_F$                                 | (13)      |
| $\Psi$                         | (91)       | $S(\varepsilon)$                                | (56)      |
| $\Psi_{\mathcal{A}}$           | (82)       | $S_f$   | (2)       |
| $\omega_{\text{FS}}$           | (45)       | $\mathbf{T}^m$                                  | (34)      |
| $\mathcal{A}_P$                | (9)        | $\text{Todd}(\mathcal{F}, \partial/\partial h)$ | (61)      |
| $b_P$                          | (16)       | $u_N$   | (134)     |
| $b_{\{x\}}$                    | (77), (78) | $u_\infty$                                      | (135)     |
| $C_F$                          | (97)       | $Z_{f_1, \dots, f_k}$                           | (21)      |
| $C_x = C_x^P$                  | (41)       | $ Z_{f_1, \dots, f_k} $                         | (20)      |
| $E_{ NP}$                      | (10)       | $\ Z_{f_1, \dots, f_k}\ $                       | (22)      |
| $E_{ P_1, \dots, P_k}$         | §1.1.3     |   |           |

## REFERENCES

- [Ag] S. Agmon, *Lectures on exponential decay of solutions of second-order elliptic equations: bounds on eigenfunctions of  $N$ -body Schrödinger operators*. Mathematical Notes 29, Princeton University Press, Princeton, NJ, 1982. MR **85f**:35019
- [AGV] V. I. Arnold, S. M. Gusein-Zade and A. N. Varchenko, *Singularities of differentiable maps. Vol. II. Monodromy and asymptotics of integrals*, Monographs in Math. 83, Birkhäuser, Boston, 1988. MR **89g**:58024
- [At] M. F. Atiyah, Angular momentum, convex polyhedra and algebraic geometry, *Proc. Edinburgh Math. Soc.* 26 (1983), 121–138. MR **85a**:58027
- [BP] A. Barvinok and J. E. Pommersheim, An algorithmic theory of lattice points in polyhedra, *New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97)*, Math. Sci. Res. Inst. Publ. 38, Cambridge Univ. Press, Cambridge, 1999, pp. 91–147. MR **2000k**:52014
- [BT] E. Bedford and B. A. Taylor, A new capacity for plurisubharmonic functions, *Acta Math.* 149 (1982), 1–40. MR **84d**:32024
- [Be] D. N. Bernstein, The number of roots of a system of equations, *Functional Anal. Appl.* 9 (1975), 183–185. MR **55**:8034
- [BSZ] P. Bleher, B. Shiffman and S. Zelditch, Universality and scaling of correlations between zeros on complex manifolds, *Invent. Math.* 142 (2000), 351–395. MR **2002f**:32037
- [BV1] M. Brion and M. Vergne, Lattice points in simple polytopes, *J. Amer. Math. Soc.* 10 (1997), 371–392. MR **98a**:11132
- [BV2] M. Brion and M. Vergne, Residue formulae, vector partition functions and lattice points in rational polytopes, *J. Amer. Math. Soc.* 10 (1997), 797–833. MR **98e**:52008
- [BV3] M. Brion and M. Vergne, An equivariant Riemann-Roch theorem for complete, simplicial toric varieties, *J. Reine Angew. Math.* 482 (1997), 67–92. MR **98a**:14067

- [De] T. Delzant, Hamiltoniens périodiques et image convexe de l'application moment, *Bull. Soc. Math. France* 116 (1988), 315–339. MR **90b**:58069
- [Eh] E. Ehrhart, Sur un problème de géométrie diophantienne linéaire, I: Polyèdres et réseaux, *J. Reine Angew. Math.* 226 (1967), 1–29. MR **35**:4184
- [Fan] X. Fan, The necessary and sufficient conditions for Lipschitz local homeomorphism, *Chinese Ann. Math. Ser. B* 13 (1992), 40–45. MR **93h**:58020
- [Fe] H. Federer, *Geometric measure theory*, Springer-Verlag, New York, 1969. MR **41**:1976
- [FPT] M. Forsberg, M. Passare and A. Tsikh, Laurent determinants and arrangements of hyperplane amoebas, *Adv. in Math.* 151 (2000), 45–70. MR **2001m**:32060
- [Fu] W. Fulton, *Introduction to Toric Varieties*, Ann. Math. Studies 131, Princeton Univ. Press, Princeton (1993). MR **94g**:14028
- [GKZ] I. M. Gelfand, M. M. Kapranov and A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Mathematics: Theory and Applications, Birkhäuser, Boston, 1994. MR **95e**:14045
- [Gu] V. Guillemin, Riemann-Roch for toric orbifolds, *J. Diff. Geometry* 45 (1997), 53–73. MR **98a**:58075
- [Hö] L. Hörmander, *The Analysis of Linear Partial Differential Operators, I*, Springer Verlag, N.Y., 1983. MR **85g**:35002a
- [HS] B. Huber and B. A. Sturmfels, A polyhedral method for solving sparse polynomial systems, *Math. Comp.* 64 (1995), 1541–1555. MR **95m**:65100
- [KP] A. G. Khovanskii and A. V. Pukhlikov, A Riemann-Roch theorem for integrals and sums of quasi-polynomials over virtual polytopes, *St. Petersburg Math J.* 6 (1993), 789–812. MR **94c**:14044
- [Kl] M. Klimek, *Pluripotential Theory*, London Math. Soc. Monographs, New Series 6, Oxford University Press, New York, 1991. MR **93h**:32021
- [Ko1] A. G. Kouchnirenko, Polyèdres de Newton et nombres de Milnor, *Invent. Math.* 32 (1976), 1–31. MR **54**:7454
- [Ko2] A. G. Kouchnirenko, Newton Polytopes and the Bezout theorem, *Functional Anal. Appl.* 10 (1976), 233–235.
- [MR] G. Malajovich and J. M. Rojas, Random Sparse Polynomial Systems, E-print (2000), [arxiv.org/math.NA/0012104](http://arxiv.org/math.NA/0012104).
- [Mi1] G. Mikhalkin, Real algebraic curves, the moment map and amoebas, *Ann. Math.* 151 (2000), 309–326. MR **2001c**:14083
- [Mi2] G. Mikhalkin, Amoebas of algebraic varieties, e-print archive, [math.AG/0108225](http://math.AG/0108225).
- [PR] M. Passare and H. Rullgård, Ameobas, Monge-Ampere measures and triangulations of the Newton polytope, *Duke Math. J.*, to appear.
- [Ro] J. M. Rojas, On the average number of real roots of certain random sparse polynomial systems. In: *The mathematics of numerical analysis (Park City, UT, 1995)*, 689–699, Lectures in Appl. Math. 32, Amer. Math. Soc., Providence, RI, 1996. MR **97j**:14060
- [Sh] B. V. Shabat, *Distribution of values of holomorphic mappings*, Amer. Math. Soc., Providence, RI, 1985. MR **86k**:32023
- [STZ1] B. Shiffman, T. Tate and S. Zelditch, Harmonic analysis on toric varieties, Contemporary Mathematics, Vol. 332, Amer. Math. Soc, Providence, RI, 2003, pp. 267–286.
- [STZ2] B. Shiffman, T. Tate and S. Zelditch, Distribution laws for integrable eigenfunctions, E-print (2003), [arxiv.org/math.CV/0306189](http://arxiv.org/math.CV/0306189).
- [SZ1] B. Shiffman and S. Zelditch, Distribution of zeros of random and quantum chaotic sections of positive line bundles, *Comm. Math. Phys.* 200 (1999), 661–683. MR **2001j**:32018
- [SZ2] B. Shiffman and S. Zelditch, Random polynomials with prescribed Newton polytope I, E-print (2002), [arxiv.org/math.AG/0203074](http://arxiv.org/math.AG/0203074).
- [SZ3] B. Shiffman and S. Zelditch, Self-averaging of distributions of zeros of random polynomials (in preparation).
- [SZ4] B. Shiffman and S. Zelditch, Random complex fewnomials (in preparation).
- [St] B. Sturmfels, On the number of real roots of a sparse polynomial system, *Hamiltonian and gradient flows, algorithms and control*, Fields Inst. Commun. 3, Amer. Math. Soc., Providence, RI, 1994, pp. 137–143. MR **95h**:12002
- [Th] T. Theobald, Computing amoebas, *Experimental Math.* 11 (2002), 513–526.
- [Ve] J. Verschelde, Toric Newton method for polynomial homotopies, *J. Symbolic Computation* 29 (2000), 777–793. MR **2001e**:65090

DEPARTMENT OF MATHEMATICS, JOHNS HOPKINS UNIVERSITY, BALTIMORE, MARYLAND 21218  
*E-mail address:* [shiffman@math.jhu.edu](mailto:shiffman@math.jhu.edu)

DEPARTMENT OF MATHEMATICS, JOHNS HOPKINS UNIVERSITY, BALTIMORE, MARYLAND 21218  
*E-mail address:* [szelditch@jhu.edu](mailto:szelditch@jhu.edu)