Random Polynomials and Approximate Zeros of Newton's Method

Joel Friedman*
Princeton University

November 12, 1987

Contents

*Part of the work was done at UC Berkeley, which was supported by an IBM (Yorktown Heights) fellowship and an IBM (Almaden) R.S.A.

1 Introduction

The goal of this paper is to give average case analyses of a randomized version of Newton's method. For a function $f\colon \mathbf{C} \to \mathbf{C}$ we define Newton's map

$$T_f(z) = z - \frac{f(z)}{f'(z)}.$$

$z \in \mathbf{C}$ is said to be an *approximate zero* of f (for Newton's method) if the iterates of z, $z_0 = z$, $z_1 = T_f(z_0)$, $z_2 = T_f(z_1)$,... converge to a root of f, ζ, and converge quickly enough so that

$$|z_k - z_{k-1}| \le \left(\frac{1}{2}\right)^{2^{k-1}-1} |z_1 - z_0| \,.$$

It is easy to see that the above condition implies

$$|z_k - \zeta| \le \frac{7}{2}\left(\frac{1}{2}\right)^{2^{k-1}} |z_0 - \zeta| \,.$$

The following α*-test* was proven in was proven independently in [Kim85] and [Sma86a]

Lemma 1.1 *For some constant $\alpha_0 > 0$, $\alpha(f,z) < \alpha_0$ implies that z is an approximate zero of f, where*

$$\alpha(f,z) \equiv \frac{|f(z)|}{|f'(z)|} \sup_{k>1} \left| \frac{f^{(k)}(z)}{k! f'(z)} \right|^{\frac{1}{k-1}} .$$

Following [Sma86b], we consider the following randomized version of Newton's method. For $f \in P_d(1) = \{f(z) = z^d + a_1 z^{d-1} + \cdots + a_d \mid |a_i| \le 1\}$ choose z with $|z| \le 3$ at random and see if $\alpha(f,z) < \alpha_0$. If not, repeat the random choice until we find a z with $\alpha(f,z) < \alpha_0$. Then apply Newton's method, which is known to converge very quickly, some small number of times. Since Newton's method converges quickly there, the main cost of the algorithm will be the number of times needed to pick z's until we find one with $\alpha(f,z) < \alpha_0$ (times the cost of verifying this condition). Let

$$\Lambda_f \equiv \{z \in \mathbf{C} \mid \alpha(f,z) < \alpha_0\},$$

and let

$$\lambda(f) \equiv \frac{|\Lambda_f \cap B_3(0)|}{|B_3(0)|},$$

the density of Λ_f in $B_3(0)$ with respect to Lebesgue measure. If z is chosen uniformly in $B_3(0)$, then the expected time to a z with $\alpha(f,z) < \alpha_0$ is $1/\lambda(f)$.

Let $Q(\epsilon)$ be the set of polynomials in $P_d(1)$ with $\lambda(f) < \epsilon$. View $P_d(1)$ as a probability space with uniform distribution as a bounded subset of $\mathbf{C}^d$. In [Sma86a] Smale proves

$$\Pr\{Q(\epsilon)\} < cd^5\epsilon \tag{1.1}$$

for some absolute constant c. In this paper we use a different approach to estimate $\Pr\{Q(\epsilon)\}$, which gives estimates for various distributions of random polynomials. For uniform on $P_d(1)$ we prove that for any integer N we have a c such that

$$\Pr\{Q(\epsilon)\} < c\left(\epsilon^2 d^3 + \left(\epsilon \log\frac{1}{\epsilon}\right)^{\frac{N-1}{2}} d^N\right).$$

This shows that $\Pr\{Q(\epsilon)\}$ decays like ϵ^2 rather than ϵ, and that with arbitrarily high probability a function will have an approximate zero region of area $> cd^{-2-\beta}$ for any $\beta > 0$ (as opposed to $> cd^{-5}$ given by equation 1.1). For polynomials with roots chosen independently and uniformly in $B_1(0)$ we get

$$\Pr\{Q(\epsilon)\} < (c\epsilon d)^d.$$

The term *approximate zero* first appeared in [Sma81]. There Smale defined a weaker notion of approximate zero (exponential as opposed to doubly exponential convergence) and proved that an iterate of 0 under a relaxation of Newton's method[1] is an approximate zero (with bounds on how large an iterate). Related papers include [SS85] and [SS86]. Before that, double exponential convergence of Newton's method was proven under conditions on the values of f and f' at a point and of f'' in a region; this was done by Kantorovich in [Kan52]; see also [KA70] and [GT74]. Independently, Kim in [Kim85] and Smale in [Sma86b] discovered the α-test. Kim used Schlicht function theory and obtained $\alpha_0 = 1/54$. Smale

[1] Namely, $T_{f,h} = z - h\frac{f'(z)}{f(z)}$ with $0 < h < 1$.

proved the α-test in the more general Banach space setting (e.g. Newton's method for maps $:\mathbf{C}^n \to \mathbf{C}^n$) and obtained $\alpha_0 = .1307\ldots$. Royden, in [Roy86], has recently improved the best known α_0 value to $.15767\ldots$ for maps $\mathbf{C} \to \mathbf{C}$.

Our method of proof obtains an estime of $\Pr\{Q(\epsilon)\}$ in terms of the distribution of the roots, which is proven in §2[2]. In §3 we apply this to the distribution on f where we take the roots to be chosen independently with uniform distribution.

In §4-7 we estimate $\Pr\{Q(\epsilon)\}$ for f with coefficients chosen independently. This leads us to the problem of determining the distribution of the roots given independently chosen coefficients. This problem has received a lot of attention(see [BS86]), but most of it is concentrated on estimating the density function of one randomly chosen root of the polynomial (i.e. "the condensed distribution"). We are interested in the joint density of two or more roots. To do this we use a generalized formula of Hammersly (see [Ham60]) for the joint density of two or more roots. In §4 we calculate the joint density of two roots assuming the coefficients are distributed normally, and then prove a theorem about the density of approximate zeros. In §5 we show that if the coefficients are distributed uniformly similar results hold for the joint density of two roots and thus about the density of approximate zeros. These results also hold for a wider class of bounded distributions. In §6 we refine our estimate of $\Pr\{Q(\epsilon)\}$ in §4-5 by estimating the joint density of three or more roots. In §7 we use an estimate of Erdös and Turán on the distribution of the roots to improve our $\Pr\{Q(\epsilon)\}$ estimates further.

2 Distances of Roots

Lemma 2.1 *Let $x_1, \ldots, x_d$ be the roots of f. Let*

$$r = \frac{1}{\sum_{j>1} \frac{1}{|x_j - x_1|}}.$$

Then $|z - x_1| < cr$ implies $\alpha(f, z) < \alpha_0$ for some absolute constant c. Furthermore, Newton's method starting at such a z converges to x_1.

[2]Independently, Rengar (see [Ren87]) has discovered such an estimate, though his is weaker by a factor of anywhere from d^2 to d^4

Proof Consider $g(y) = f(y-z) = \sum_{i=0}^{d} b_i y_i$, which has roots $x_i - z$. We wish to bound

$$\alpha(f,z) = \left|\frac{b_0}{b_1}\right| \max_{k>1} \left|\frac{b_k}{b_1}\right|^{\frac{1}{k-1}}.$$

Consider $h(y) = \sum_{i=0}^{d} b_{d-i} y^i$. Since $h(y) = y^d g(1/y)$, h has roots $\frac{1}{x_i - z}$ and thus

$$\sum_{i=0}^{d} b_{d-i} y^i = h(y) = b_0 \prod_{i=1}^{d} \left(y - \frac{1}{x_i - z}\right).$$

Thus

$$b_i = (-1)^i b_0 \, \sigma_i\left(\frac{1}{x_1 - z}, \ldots, \frac{1}{x_d - z}\right)$$

where σ_k is the kth symmetric polynomial,

$$\sigma_k(w_1, \ldots, w_d) = \sum_{i_1 < \cdots < i_k} w_{i_1} w_{i_2} \ldots w_{i_d}.$$

Now

$$\begin{aligned}
\sigma_k\left(\frac{1}{x_1 - z}, \ldots, \frac{1}{x_d - z}\right) &= \sigma_k\left(\frac{1}{x_2 - z}, \ldots, \frac{1}{x_d - z}\right) + \\
&\quad \frac{1}{x_1 - z} \sigma_{k-1}\left(\frac{1}{x_2 - z}, \ldots, \frac{1}{x_d - z}\right).
\end{aligned}$$

Since $z \in B_{cr}(x_1)$ we have

$$\frac{c}{|x_1 - z|} > \sum_{j>1} \frac{1}{|x_j - z|}.$$

Thus

$$\left|\sigma_m\left(\frac{1}{x_2 - z}, \ldots, \frac{1}{x_d - z}\right)\right| \le \left(\sum_{i=2}^{d} \frac{1}{|x_i - z|}\right)^m \le \left(\frac{c}{|x_1 - z|}\right)^m$$

and

$$\begin{aligned}
\left|\sigma_m\left(\frac{1}{x_1 - z}, \ldots, \frac{1}{x_d - z}\right)\right| &\le \left|\sigma_m\left(\frac{1}{x_2 - z}, \ldots, \frac{1}{x_d - z}\right)\right| \\
&\quad + \frac{1}{|x_1 - z|} \left|\sigma_{m-1}\left(\frac{1}{x_2 - z}, \ldots, \frac{1}{x_d - z}\right)\right| \\
&\le \frac{c^m + c^{m-1}}{|x_1 - z|^m}.
\end{aligned}$$

On the other hand

$$\begin{aligned} \left|\sigma_1\Big(\frac{1}{x_1-z},\ldots,\frac{1}{x_d-z}\Big)\right| &= \left|\frac{1}{x_1-z}+\cdots+\frac{1}{x_d-z}\right| \\ &\geq \frac{1}{|x_1-z|}-\sum_{i=2}^{d}\frac{1}{|x_i-z|} \\ &\geq \frac{1-c}{|x_1-z|}. \end{aligned}$$

Hence

$$\left|\frac{b_m}{b_1}\right| \leq \frac{c^m+c^{m-1}}{1-c}\frac{1}{|x_1-z|^{m-1}}$$

and

$$\left|\frac{b_0}{b_1}\right| \leq \frac{|x_1-z|}{1-c}.$$

Thus

$$\begin{aligned} \alpha(f,z) &\leq \frac{|x_1-z|}{1-c}\max_{k>1}\left(\frac{c^k+c^{k-1}}{1-c}\right)^{\frac{1}{k-1}}\frac{1}{|x_1-z|} \\ &= \frac{1}{1-c}\max_{k>1} c\left(\frac{1+c}{1-c}\right)^{\frac{1}{k-1}} = \frac{c}{1-c}\sqrt{\frac{1+c}{1-c}}. \end{aligned}$$

And hence $\alpha(f,z) < \alpha_0$ for appropriate choice of c.

3 The Uniform Root Distribution

We can use lemma 2.1 to estimate the measure of $Q(\epsilon)$ for various distributions on the set of degree d polynomials. In this section we illustrate this by carrying out such an estimate in a case where the roots are distributed independently. In this case we can apply lemma 2.1 without much difficulty. Consider the distribution on polynomials

$$f(z) = (z-x_1)\ldots(z-x_d)$$

with $x_1,\ldots,x_d$ chosen independently, uniform in $B_1(0) \subset \mathbf{C}$. We begin by proving $\Pr\{Q(\epsilon)\} \leq cd\epsilon$ for some c, and then we refine the argument to get $\Pr\{Q(\epsilon)\} \leq (cd\epsilon)^d$ for some c.

Theorem 3.1 $\Pr\{Q(\epsilon)\} \le cd\epsilon$ *for all* ϵ *for some absolute constant* c.

Proof Viewing x_1 as fixed, we have for any fixed j

$$\Pr\{|x_j - x_1| < \rho\} = \frac{|B_\rho(0) \cap B_1(0)|}{|B_1(0)|} \le \rho^2.$$

Thus, for any $i_1 < \cdots < i_k$, we have

$$\Pr\{|x_{i_1} - x_1|, \ldots, |x_{i_k} - x_1| \text{ are } \le \rho\} \le \rho^{2k}.$$

Hence

$$\Pr\{k \text{ of } |x_2 - x_1|, |x_3 - x_1|, \ldots, |x_d - x_1| \text{ are} \le \rho\} \le \binom{d-1}{k}\rho^{2k} \le \left(\frac{ed}{k}\right)^k \rho^{2k},$$

(where $\binom{d-1}{k} \le \binom{d}{k} \le (ed/k)^k$ was used) which is $\le \eta/2^k$ if $\rho \le \sqrt{\frac{k}{2ed}\eta^{1/2k}}$. So if $a_1 \le \ldots \le a_{d-1}$ are the $|x_2 - x_1|, \ldots, |x_d - x_1|$ arranged in increasing order, we have

$$\Pr\left\{a_1 \le \sqrt{\frac{1}{2ed}\eta^{1/2}} \text{ or } a_2 \le \sqrt{\frac{2}{2ed}\eta^{1/4}} \text{ or } \ldots \text{ or } a_{d-1} \le \sqrt{\frac{d-1}{2ed}\eta^{1/2(d-1)}}\right\}$$

$$\le \frac{\eta}{2} + \frac{\eta}{4} + \cdots + \frac{\eta}{2^{d-1}} < \eta.$$

Hence with probability $\ge 1 - \eta$ we have

$$\sum_{i=1}^{d-1} \frac{1}{a_i} \le \sqrt{2ed}\left(\left(\frac{1}{\eta}\right)^{1/2} + \frac{1}{\sqrt{2}}\left(\frac{1}{\eta}\right)^{1/4} + \cdots \frac{1}{\sqrt{d-1}}\left(\frac{1}{\eta}\right)^{1/2(d-1)}\right). \quad (3.1)$$

Lemma 3.2 $N + \frac{1}{\sqrt{2}}N^{1/2} + \cdots + \frac{1}{\sqrt{m}}N^{1/m} \le 2N + 4\sqrt{m}$ *for* $N \ge 4$.

Proof Let $t > 1$ be the first integer for which $N^{1/t} \le 2$. Then $t-1 \le \log_2 N$ and so

$$\begin{aligned} N + \frac{1}{\sqrt{2}}N^{1/2} + \cdots + \frac{1}{\sqrt{t-1}}N^{1/(t-1)} &\le N + (t-2)N^{1/2} \\ &\le N + (\log_2 N)N^{1/2} \le 2N \end{aligned}$$

since $\log_2 N \le N^{1/2}$ for $N \ge 4$. Furthermore

$$\frac{1}{\sqrt{t}}N^{1/t} + \cdots + \frac{1}{\sqrt{m}}N^{1/m} \le 2\left(\frac{1}{\sqrt{2}} + \cdots \frac{1}{\sqrt{m}}\right)$$

$$\le \quad 2\int_1^m \frac{1}{\sqrt{x}}\,dx \le 4\sqrt{m}$$

and the lemma follows.

Applying lemma 3.2 to equation 3.1 yields that with probability $\le 1-\eta$,

$$\sum_{i=1}^{d-1} \frac{1}{a_i} \le \sqrt{2\pi d}(2\sqrt{1/\eta} + 4\sqrt{d})$$

(for $\sqrt{1/\eta} \ge 4$), and if $1/\eta \ge d$ this gives

$$\sum_{i=1}^{d-1} \frac{1}{a_i} \le \sqrt{2\pi d}(2\sqrt{1/\eta} + 2\sqrt{1/\eta}) \le \sqrt{8\pi}\sqrt{d/\eta}$$

and hence

$$\frac{1}{\sum \frac{1}{a_i}} \ge c\sqrt{\eta/d}$$

for some constant c. Hence, by lemma 2.1, $\alpha(f,z) < \alpha_0$ in a ball about x_1 of area

$$c\frac{\eta}{d}.$$

Applying the arguments with x_1 replaced by an arbitrary root, it follows that with probability $\ge 1 - d\eta$, each root has $\alpha(f,z) \le \alpha_0$ in a ball about it of area $c\eta/d$, for a total area of $c\eta$.

Lemma 3.3 *The probability that there are $d/2$ roots each having k other roots within a distance δ is*

$$\le \left(\delta^2 d^{1+\frac{1}{k}}/k\right)^{\frac{d}{2}\left(1-\frac{1}{k+1}\right)}.$$

Proof If there are $d/2$ roots each having k other roots within a distance δ, then for some distinct integers

$$i_1, i_1^1, i_1^2, \ldots i_1^k; i_2, i_2^1, \ldots, i_2^k; \ldots; i_j, i_j^1, \ldots i_j^k$$

with $j = d/2(k+1)$ we have

$$x_{i_n^m} \in B_\delta(x_{i_n}). \tag{3.2}$$

For any fixed set of integers this happens with probability

$$\leq \left(\delta^2\right)^{k\frac{d}{2(k+1)}}.$$

The number of ways of choosing a distinguished i_1 and distinct set of size k, $i_1^1 < \cdots < i_1^k$ is

$$d\binom{d-1}{k} \leq d\left(\frac{ed}{k}\right)^k.$$

The total number of ways of choosing j such sets of integers is

$$\leq \left[d\left(\frac{ed}{k}\right)^k\right]^{d/2(k+1)} = d^{d/2}(e/k)^{\frac{d}{2}\frac{k}{k+1}}.$$

Thus the probability that for some set of integers equation 3.2 holds is

$$\begin{aligned} &\leq \left(\delta^2\right)^{\frac{d}{2}\frac{k}{k+1}} d^{d/2}(e/k)^{\frac{d}{2}\frac{k}{k+1}} \\ &= \left(\delta^2 d^{1+\frac{1}{k}} e/k\right)^{\frac{d}{2}\left(1-\frac{1}{k+1}\right)}. \end{aligned}$$

Corollary 3.4 *For any $\delta_1, \ldots, \delta_{d-1}$ we have that with probability $\geq 1 - \eta_1 - \cdots - \eta_{d-1}$ the region of approximate zeros is*

$$\geq \frac{d}{2}c\left(\frac{1}{\sum 1/\delta_k}\right)^2$$

where

$$\eta_k = \left(\delta_k^2 d^{1+\frac{1}{k}} e/k\right)^{\frac{d}{2}\left(1-\frac{1}{k+1}\right)}.$$

Proof This follows from lemma 2.1.

Take $\eta_k = \eta/(d-1)$ in corollary 3.4 so that

$$\begin{aligned} \delta_k &= \sqrt{\frac{k}{ed^{1+\frac{1}{k}}}\,\eta_k^{\frac{1}{d\left(1-\frac{1}{k+1}\right)}}} \\ &\leq c\sqrt{\frac{1}{d}}\sqrt{\frac{k}{d^{1/k}}}\,\eta^{\frac{1}{d/2}}. \end{aligned}$$

Lemma 3.5 $\sqrt{d} + \sqrt{d^{1/2}/2} + \sqrt{d^{1/3}/3} + \cdots + \sqrt{d^{(d-1)}/(d-1)} = O\left(\sqrt{d}\right)$.

Proof Since $\sqrt{d^{1/x}/x}$ is monotone decreasing in x, we can estimate

$$\begin{aligned}
\sqrt{d^{1/2}/2} + \cdots + \sqrt{d^{1/(d-1)}/(d-1)} &\leq \int_2^d \sqrt{d^{1/x}/x}\, dx \\
&\leq \int_2^d \frac{1}{\sqrt{x}} e^{-\frac{\log d}{2x}}\, dx \\
&\leq \left[\sqrt{x} e^{-\frac{\log d}{2x}}\right]_2^d = O\left(\sqrt{d}\right).
\end{aligned}$$

This gives us that with probablity $\geq 1 - \eta$ we have an area of

$$\begin{aligned}
&\geq c\frac{d}{2}\left(\sqrt{\frac{1}{d}}\eta^{\frac{1}{d/2}} \frac{c}{\sqrt{d}}\right)^2 \\
&\geq c\eta^{\frac{1}{d}}.
\end{aligned}$$

Taking $\epsilon = c\eta^{1/d}/d$, i.e. $\eta = (\epsilon d/c)^d$ yields

$$\Pr\{Q(\epsilon)\} \leq \left(\frac{\epsilon d}{c}\right)^d.$$

4 Normally Distributed Coefficients

In the next sections we will estimate $\Pr\{Q_d(\epsilon)\}$ for distributions in which the coefficients a_i of

$$f(z) = a_d z^d + \cdots + a_0 \tag{4.1}$$

are chosed independently with fixed distributions. In this section we consider the case in which the a_i's are distibuted normally, i.e. $\Re(a_i)$ and $\Im(a_i)$ are independent random variables on $\mathbf{R}$ with density

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Here we have the problem that for any fixed value of d, there is some small probability that all the coefficients are large enough to enable $B_3(0)$ to lie completely within a sink of period 2. Hence $\Pr\{Q_d(0)\} > 0$ for each d; we cannot hope to prove $\Pr\{Q_d(\epsilon)\} \le \epsilon d^{O(1)}$. Instead, we shall prove

$$\Pr\{Q_d(\epsilon)\} \le c_1\epsilon^2 d^7 + 2^{-c_2 d}, \tag{4.2}$$

the $2^{-c_2 d}$ term taking $\Pr\{Q_d(0))\} > 0$ into account.

We begin by noting that for d large the roots tend to be located on the circle of radius 1.

Lemma 4.1 (Specht) *Let $z_1, \ldots, z_m$ be roots of $z^n + a_{n-1}z^{n-1} + \cdots + a_0$. Then*

$$|z_1 \ldots z_m| \le \sqrt{1 + |a_{d-1}|^2 + |a_{d-2}|^2 + \cdots + |a_0|^2}.$$

Proof See [Mar66].

Corollary 4.2 *If less that $d/2$ of the roots of $f(z) = a_d z^d + \cdots + a_0$ have absolute value between $1/2$ and 2 then either*

$$\frac{1}{|a_d|}\sqrt{|a_0|^2 + \cdots + |a_d|^2} \ge 2^{d/4} \tag{4.3}$$

or

$$\frac{1}{|a_0|}\sqrt{|a_0|^2 + \cdots + |a_d|^2} \ge (1/2)^{-d/4} = 2^{d/4}. \tag{4.4}$$

Proof Either $d/4$ roots have absolute value > 2 or $< 1/2$. Apply lemma 4.1 to either $\frac{1}{a_n}f(z)$ or $\frac{1}{a_0}f(1/z)z^d$.

Corollary 4.3 *With probability $\ge 1 - 2^{cd}$ there are at least $d/2$ roots z in the range $1/2 \le |z| \le 2$ for some constant $c > 0$.*

Proof For equation 4.3 of equation 4.4 to hold, one of the a_i's must be exponentially large or exponentially small (i.e. $\ge 2^{cd}$ or $\le 2^{-cd}$). For standard normal random variables, this occurs with probability $\le 2^{-cd}$ for some $c \ge 0$.

Next we deive a bound of the form $\Pr\{|z_1 - z_2| \le \epsilon\} \le O(\epsilon^4)$ ($O(\epsilon^4)$ for fixed d) where z_1, z_2 are randomly chosen roots of $f(z) = a_d z^d + \cdots + a_0$. Constrasting this to $\Pr\{|z_1 - z_2| \le \epsilon\} = O(\epsilon^2)$ when z_1, z_2 are disributed independently and uniformly explains why in equation 4.2 we estimate $\Pr\{Q(\epsilon)\}$ quadratically in ϵ rather than linearly (i.e. equation (3.1)).

In [Ham60], Hammersly gives a formula for the density function, [3] $P(z_1)$, of a randomly chosen root, z_1, of $f(z) = a_d z^d + \cdots + a_0$. Viewing $f(z_1)$ and $f'(z_1)$, for z_1 fixed, as sums of independent random variables $(z_1^d)a_d + \cdots + a_0$ and $(dz_1^{d-1})a_d + \cdots + a_1$, the formula for P can be written as

$$P(z_1) = \frac{1}{d}\, E\left\{|f'(z_1)|^2 \text{ subject to } f(z_1) = 0\right\},$$

where by

$$E\left\{|f'(z_1)|^2 \text{ s.t. } f(z_1) = 0\right\},$$

the expected value of $|f'(z_1)|^2$ subject to $f(z_1) = 0$, we mean

$$\int_{\mathbf{C}} \psi(0,t)|t|^2\, dt$$

where ψ is the joint density function of $f(z_1)$ and $f'(z_1)$. One can generalize this formula to the joint density of k randomly chosen roots

$$P(z_1, \ldots, z_k) = \frac{1}{d(d-1)\ldots(d-k+1)}\, E\left\{|f'(z_1)|^2 \ldots |f'(z_k)|^2 \text{ s.t. } f(z_1) = \cdots = f(z_k) = 0\right\};$$

see appendix B for the derivation. In particular

$$P(z_1, z_2) = \frac{1}{d(d-1)}\, E\left\{|f'(z_1)|^2 |f'(z_2)|^2 \text{ s.t. } f(z_1) = (z_2) = 0\right\}. \qquad (4.5)$$

We will estimate this expression for $\Delta z = z_2 - z_1$ with $|\Delta z| \leq 1/d^{5/4}$ (actually $\leq c/d$ for some constant c would give the same estimates) and $\frac{1}{2} < |z_i| < 2$.

For constants $b_0, \ldots, b_d$ we can write the random variable $\sum b_i a_i$ as $\langle \mathbf{b}, \bar{\mathbf{a}} \rangle$, where $\mathbf{b} = (b_0, \ldots, b_d) \in \mathbf{C}^{d+1}$, $\bar{\mathbf{a}} = (\bar{a}_0, \ldots, \bar{a}_d) \in \mathbf{C}^{d+1}$, $^-$ denoting complex conjugation, and $\langle\, , \rangle$ denotes the usual inner product on $\mathbf{C}^{d+1}$. Analogous to sums of real normal random variables, one can easily verify that $\langle \mathbf{b}, \bar{\mathbf{a}} \rangle$ and $\langle \mathbf{b}', \bar{\mathbf{a}} \rangle$ are independent random variables if $\Re \langle \mathbf{b}, \mathbf{b}' \rangle = 0$.

[3] Equivalently the chance of finding at least one root in $B_\epsilon(z_1)$ is $dP(z_1) \cdot \pi\epsilon^2 +$ (lower order terms).

For $i = 1, 2$, let $\mathbf{u_i}$ denote $(1, z_i, \ldots, z_i^d) \in \mathbf{C}^{d+1}$ and $\mathbf{v_i}$ denote

$$\mathbf{v_i} \equiv (0, 1, 2z_i, \ldots, dz_i^{d-1}) \in \mathbf{C}^{d+1}.$$

Let $\tilde{\mathbf{v}}_\mathbf{i}$ be the projection of $\mathbf{v_i}$ onto $(\mathbf{Cu_1} + \mathbf{Cu_2})^\perp$, i.e.

$$\tilde{\mathbf{v}}_\mathbf{i} = \mathbf{v_i} - \sum_{j=1}^{2} \frac{\left\langle \mathbf{v_i}, \mathbf{u_j} \right\rangle}{\left\langle \mathbf{u_j}, \mathbf{u_j} \right\rangle} \mathbf{u_j}.$$

We have

$$\begin{aligned} P(z_1, z_2) &= \frac{1}{\binom{d}{2}} E\left\{ |\langle \mathbf{v_1}, \bar{\mathbf{a}} \rangle|^2 |\langle \mathbf{v_2}, \bar{\mathbf{a}} \rangle|^2 \text{ s.t. } \langle \mathbf{u_1}, \bar{\mathbf{a}} \rangle = \langle \mathbf{u_2}, \bar{\mathbf{a}} \rangle = 0 \right\} \\ &= \frac{1}{\binom{d}{2}} E\left\{ |\langle \tilde{\mathbf{v}}_\mathbf{1}, \bar{\mathbf{a}} \rangle|^2 |\langle \tilde{\mathbf{v}}_\mathbf{2}, \bar{\mathbf{a}} \rangle|^2 \text{ s.t. } \langle \mathbf{u_1}, \bar{\mathbf{a}} \rangle = \langle \mathbf{u_2}, \bar{\mathbf{a}} \rangle = 0 \right\} \end{aligned}$$

(since $\left\langle \mathbf{v_j}, \bar{\mathbf{a}} \right\rangle = \left\langle \tilde{\mathbf{v}}_\mathbf{j}, \bar{\mathbf{a}} \right\rangle$ if $\langle \mathbf{u_i}, \bar{\mathbf{a}} \rangle = 0$)

$$= \frac{1}{\binom{d}{2}} E\left\{ |\langle \tilde{\mathbf{v}}_\mathbf{1}, \bar{\mathbf{a}} \rangle|^2 |\langle \tilde{\mathbf{v}}_\mathbf{2}, \bar{\mathbf{a}} \rangle|^2 \right\} \psi(0,0) \tag{4.6}$$

by independence, where ψ is the joint density of $\langle \mathbf{u_1}, \bar{\mathbf{a}} \rangle$, $\langle \mathbf{u_2}, \bar{\mathbf{a}} \rangle$. Similar to the case of real normal random variables,

$$\psi(0,0) = \left(\frac{1}{2\pi} \right)^2 \left| \det \begin{pmatrix} \langle \mathbf{u_1}, \mathbf{u_1} \rangle & \langle \mathbf{u_1}, \mathbf{u_2} \rangle \\ \langle \mathbf{u_2}, \mathbf{u_1} \rangle & \langle \mathbf{u_2}, \mathbf{u_2} \rangle \end{pmatrix} \right|^{-1}.$$

Add a note here giving a reference or explain real isomorphs of complex matrices. Letting $\mathbf{\Delta u} = \mathbf{u_2} - \mathbf{u_1}$, we have

$$\begin{aligned} \mathbf{\Delta u} &= (z_2 - z_1)(0, 1, z_1 + z_2, \ldots, z_1^{d-1} + z_1^{d-2} z_2 + \cdots + z_2^{d-1}) \\ &= \Delta z (0, 1, 2z_1, \ldots, dz_1^{d-1}) \Big(1 + O(d^{-1/4})\Big) \\ &= \Delta z \, \mathbf{v_1} \Big(1 + O(d^{-1/4})\Big) \end{aligned}$$

and so

$$\left| \det \begin{pmatrix} \langle \mathbf{u_1}, \mathbf{u_1} \rangle & \langle \mathbf{u_1}, \mathbf{u_2} \rangle \\ \langle \mathbf{u_2}, \mathbf{u_1} \rangle & \langle \mathbf{u_2}, \mathbf{u_2} \rangle \end{pmatrix} \right| = \left| \det \begin{pmatrix} \langle \mathbf{u_1}, \mathbf{u_1} \rangle & \langle \mathbf{u_1}, \mathbf{\Delta u} \rangle \\ \langle \mathbf{\Delta u}, \mathbf{u_1} \rangle & \langle \mathbf{\Delta u}, \mathbf{\Delta u} \rangle \end{pmatrix} \right|$$

$$= |\Delta z|^2\Big(|\mathbf{u_1}|^2|\mathbf{v_1}|^2\Big(1+O(d^{-1/4})\Big) - |\langle \mathbf{u_1},\mathbf{v_1}\rangle|^2\Big(1+O(d^{-1/4})\Big)\Big).$$

We have

$$\begin{aligned}
|\mathbf{u_1}|^2 &= 1+|z_1|^2+\cdots+|z_1|^{2d},\\
|\mathbf{v_1}|^2 &= 1+4|z_1|^2+\cdots+d^2|z_1|^{2d-2},\\
\langle \mathbf{u_1},\mathbf{v_1}\rangle &= z_1(1+2|z_1|^2+3|z_1|^4+\cdots+d|z_1|^{2d-2}),\\
|\langle \mathbf{u_1},\mathbf{v_1}\rangle|^2 &= |z_1|^2(1+2|z_1|^2+3|z_1|^4+\cdots+d|z_1|^{2d-2})^2.
\end{aligned}$$

Proposition 4.4 *$|\mathbf{u_1}|^2|\mathbf{v_1}|^2$, $|\langle \mathbf{u_1},\mathbf{v_1}\rangle|^2$, and $|\mathbf{u_1}|^2|\mathbf{v_1}|^2 - |\langle \mathbf{u_1},\mathbf{v_1}\rangle|^2$ are each $\Theta(d^4)$ for $1-(1/d)\le |z|\le 1$ and $\Theta((1-|z|^2)^{-4})$ for $|z|\le 1-(1/d)$. (For f to be $\Theta(g)$ means $c_1 g < f < c_2 g$ for some constants c_1 and c_2.)*

Proof Let $y=|z|^2$. We have

$$\begin{aligned}
|\mathbf{u_1}|^2|\mathbf{v_1}|^2 &= (1+y+\cdots+y^d)(1+4y+\cdots+d^2y^{d-1}),\\
|\langle \mathbf{u_1},\mathbf{v_1}\rangle|^2 &= y(1+2y+\cdots+dy^{d-1})^2,
\end{aligned}$$

and upon substraction

$$\begin{aligned}
|\mathbf{u_1}|^2|\mathbf{v_1}|^2 - |\langle \mathbf{u_1},\mathbf{v_1}\rangle|^2 &= 1+\binom{4}{3}y+\binom{5}{3}y^2+\cdots+\binom{d+2}{3}y^{d-1}+\\
&\quad c_dy^d+c_{d+1}y^{d+1}+\cdots+c_{2d-2}y^{2d-2},
\end{aligned}$$

where $c_d,\ldots,c_{2d-2}$ are positive integers. If $1-(1/d)<|z|<1$, then we have $e^{-2}\le |z|^j\le 1$ for any $j=1,\ldots,2d-1$ and the aforementioned estimates easily follow. If $|z|<1-(1/d)$, then the proposition follows using

$$\sum_{i=0}^{d} i^n r^i \text{ and } \sum_{i=0}^{\infty} i^n r^i \text{ are both } =\Theta\left(\frac{1}{1-r}\right)^n$$

for any n and any $r<1-c/d$ for any fixed c (this latter condition ensures that the former sum has enough terms to approximate its limiting infinite sum).

Corollary 4.5

$$\psi(0,0)\le \frac{c}{|\Delta z|^2}\begin{cases} d^4 & \text{if } 1-\frac{1}{d}\le |z_1|\le 1\\ \frac{d^2}{(1-|z_1|^2)^4} & \text{if } \frac{1}{2}\le |z_1|\le 1-\frac{1}{d}\end{cases}$$

To estimate

$$E\left\{ |\langle \tilde{\mathbf{v}}_\mathbf{1}, \bar{\mathbf{a}}\rangle|^2 |\langle \tilde{\mathbf{v}}_\mathbf{2}, \bar{\mathbf{a}}\rangle|^2 \right\}$$

note that

$$\mathbf{u_2} = \mathbf{u_1} + (\Delta z)\mathbf{v_1} + \frac{(\Delta z)^2}{2}\mathbf{w_1}\Big(1 + O(d^{-1/4})\Big)$$

where

$$\mathbf{w_1} = (0, 0, 2, 6z_1, \ldots, d(d-1)z_1^{d-2}).$$

Hence

$$\tilde{\mathbf{v}}_\mathbf{1} = \left(\mathbf{v_1} - \frac{\mathbf{u_2} - \mathbf{u_1}}{\mathbf{\Delta z}}\right)^{\sim} = \frac{\Delta z}{2}\tilde{\mathbf{w}}_\mathbf{1}\Big(1 + O(d^{-1/4})\Big)$$

(where $\tilde{}$ denotes the projection onto $(\mathbf{Cu_1} + \mathbf{Cu_2})^\perp$). Similarly we have

$$\tilde{\mathbf{v}}_\mathbf{2} = \frac{\Delta z}{2}\tilde{\mathbf{w}}_\mathbf{1}\Big(1 + O(d^{-1/4})\Big).$$

So we estimate

$$E\left\{ |\langle \tilde{\mathbf{v}}_\mathbf{1}, \bar{\mathbf{a}}\rangle|^2 |\langle \tilde{\mathbf{v}}_\mathbf{2}, \bar{\mathbf{a}}\rangle|^2 \right\} \leq c(\Delta z)^4 E\left\{ |\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}}\rangle|^2 |\langle \tilde{\mathbf{w}}_\mathbf{2}, \bar{\mathbf{a}}\rangle|^2 \right\}$$

(where c is an absolute constant replacing $(1 + O(d^{-1/4}))$

$$\leq c(\Delta z)^4 \left(E\left\{ |\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}}\rangle|^4 \right\} + E\left\{ |\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}}\rangle|^4 \right\} \right)$$

by Schwartz' inequality. To simplify estimating these fourth moments we use

Proposition 4.6 *Let α, β, and γ be independent complex valued random variables with $E\{\alpha^i\bar{\alpha}^j\} = 0$ for $i \neq j$ and similary for β and γ. Then*

$$E\left\{|\alpha|^4\right\} \leq E\left\{|\alpha + \beta + \gamma|^4\right\}.$$

Proof

$$E\left\{|\alpha + \beta + \gamma|^4\right\} = E\left\{(\alpha + \beta + \gamma)^2(\bar{\alpha} + \bar{\beta} + \bar{\gamma})^2\right\}$$

which, when expanded as sum of expectations of products has terms which are of the form $E\left\{\delta\bar{\delta}\right\} > 0$ or which drop out. One of these terms is $E\{\alpha^2\bar{\alpha}^2\}$.

Since $\langle \mathbf{w_1}, \bar{\mathbf{a}} \rangle$ is the sum of the three independent, radially symmetric random variables $\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}} \rangle$, $\alpha_1 \langle \mathbf{u_1}, \bar{\mathbf{a}} \rangle$, and $\alpha_2 \langle \mathbf{u_2}, \bar{\mathbf{a}} \rangle$ for appropriate α_1, α_2, we have

$$E\left\{ |\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}} \rangle|^4 \right\} \leq E\left\{ |\langle \mathbf{w_1}, \bar{\mathbf{a}} \rangle|^4 \right\}.$$

Now

$$\begin{aligned} E\left\{ |\langle \mathbf{w_1}, \bar{\mathbf{a}} \rangle|^4 \right\} &= E\left\{ \left| \sum_{i=0}^{d} i(i-1) z^{i-2} a_i \right|^4 \right\} \\ &= \sum_{i,j,k,l} i(i-1)j(j-1)k(k-1)l(l-1) z^{i+j+k+l-8} E\left\{ a_i a_j \bar{a}_k \bar{a}_l \right\}. \end{aligned}$$

The only terms not vanishing in the latter sum are those for which either $i = k$, $j = l$, or $i = l$, $j = k$. By the symmetry of these conditions, and since the $E\left\{ a_i a_j \bar{a}_k \bar{a}_l \right\}$ are bounded, we can estimate the above sum by

$$\leq c \sum_{i,j} i^4 j^4 |z|^{2(i+j)-8} \leq c' \sum_{m=0}^{2d} m^9 |z|^{2m-8} \tag{4.7}$$

where we have set $m = i + j$. If $|z| < 1 - (1/d)$, we can estimate equation 4.7 by using

$$\sum_{i=0}^{\infty} i^9 r^i \leq c \left(\frac{1}{1-r} \right)^{10}$$

to get

$$E\left\{ |\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}} \rangle|^4 \right\} \leq c \left(\frac{1}{1-|z|^2} \right)^{10}$$

which gives $O(d^{10})$ or better for $|z|$ in this range. For $1 - (1/d) \leq |z| \leq 1$ we simply use $|z|^m \leq 1$ in equation 4.7 to get

$$E\left\{ |\langle \tilde{\mathbf{w}}_\mathbf{1}, \bar{\mathbf{a}} \rangle|^4 \right\} \leq c \sum_{m=0}^{2d} m^9 \leq c' d^{10}.$$

Summing up, we have

Lemma 4.7

$$E\left\{ |\langle \tilde{v}_1, \bar{a} \rangle|^2 |\langle \tilde{v}_2, \bar{a} \rangle|^2 \right\} \leq c |\Delta z|^4 \begin{cases} d^{10} & \mathit{if}\ 1 - \frac{1}{d} \leq |z_1| \leq 1 \\ (1 - |z_1|^2)^{-10} & \mathit{if}\ \frac{1}{2} \leq |z_1| \leq 1 - \frac{1}{d} \end{cases}$$

Combining lemma 4.7, corollary 4.5, and equation 4.6 yields

Theorem 4.8

$$P(z_1, z_2) \le \frac{c}{d(d-1)}(\Delta z)^2 \begin{cases} d^6 & \text{if } 1 - \frac{1}{d} \le |z_1| \le 1 \\ (1 - |z_1|^2)^{-6} & \text{if } \frac{1}{2} \le |z_1| \le 1 - \frac{1}{d} \end{cases}$$

Corollary 4.9 *For $1 \le |z_1| \le 2$, the same estimates, as in theorem 4.8, hold (with slightly different c and θ).*

Proof Let $y_1 = 1/z_1$, $y_2 = 1/z_2$. Let $\tilde{P}(\ ,\ )$ be the density of two random roots of

$$a_0 y^d + a_1 y^{d-1} + \cdots + a_d = 0. \tag{4.8}$$

On the one hand, clearly $\tilde{P} = P$. On the other hand, y satisfied equation 4.8 iff $x = 1/y$ satisfies

$$a_d x^d + \cdots + a_0 = 0.$$

Thus

$$\begin{aligned} P(s,t) &= \tilde{P}\left(\frac{1}{s}, \frac{1}{t}\right) \left| \frac{\partial(1/s, 1/t)}{\partial(s,t)} \right|^2 \\ &= \tilde{P}\left(\frac{1}{s}, \frac{1}{t}\right) \frac{1}{|st|^4} \\ &= P\left(\frac{1}{s}, \frac{1}{t}\right) \frac{1}{|st|^4}. \end{aligned}$$

Thus, since $1 \le |z_1|, |z_2| \le 2$, we have

$$P(z_1, z_2) \le P(y_1, y_2) \frac{1}{4^4}.$$

Since $|y_1 - y_2| \le c/d$ can be ensured by requiring $|z_1 - z_2| \le c' d$, we can apply theorem 4.8 in this case to obtain the desired estimate (note that $(1 - |z_1|)^2$ and $(1 - |y_1|)^2$ differ from each other by some multiple in a bounded, positive range).

Lemma 4.10

$$\Pr\left\{ |z_1 - z_2| \le \delta \text{ and } \frac{1}{2} \le |z_1| \le 2 \right\} \le c\delta^4 d^3$$

Proof We have

$$\int_{t\in B_2(0)-B_{1/2}(0)} \int_{s\in B_\delta(t)} P(s,t)\,ds\,dt$$
$$\leq \int_{t\in B_2(0)-B_{1/2}(0)} \frac{c}{d^2}\delta^4 \left\{ \begin{array}{ll} d^6 & \text{if } 1-\frac{1}{d} \leq |t| \leq 1 \\ (1-|t|^2)^{-6} & \text{if } \frac{1}{2} \leq |t| \leq 1-\frac{1}{d} \end{array} \right\} dt. \quad (4.9)$$

The integral in equation 4.9, over the range $1-\frac{1}{d} \leq |t| \leq 1+\frac{1}{d}$ is

$$\begin{aligned} &\leq \frac{c}{d^2}\delta^4 d^6 \left| B_{1+\frac{1}{d}}(0) - B_{1-\frac{1}{d}}(0) \right| \\ &= \frac{c}{d^2}\delta^4 d^6 \frac{4\pi}{d} = c'\delta^4 d^3. \end{aligned}$$

Over the range $1/2 \leq |t| \leq 1-\frac{1}{d}$, setting $r=|t|$ the integral of equation 4.9 becomes

$$\begin{aligned} \int_t \frac{c}{d^2}\delta^4 \left(\frac{1}{1-|t|^2}\right)^6 dt &= \int_{r=1/2}^{1-\frac{1}{d}} \frac{c}{d^2}\delta^4 \left(\frac{1}{1-r^2}\right)^6 2\pi r\,dr \\ &= \frac{c'}{d^2}\delta^4 \left.(1-r^2)^{-5}\right|_{1/2}^{1-\frac{1}{d}} \\ &= \frac{c'}{d^2}\delta^4 \left[\left(\frac{2}{d}-\frac{1}{d^2}\right)^{-5} - \left(\frac{3}{4}\right)^{-5}\right] \leq c''\delta^4 d^3. \end{aligned}$$

Corollary 4.11 *The probability that there is a root $|z_i|$ with $1/2 \leq |z_i| \leq 2$ for which there is some other root, $|z_j|$ with $|z_i - z_j| \leq \delta$, is no more that $c\delta^4 d^5$ for some constant c.*

Proof For $\delta > d^{-5/4}$ the statement holds with $c=1$. For $\delta \leq d^{-5/4}$ we apply lemma 4.10 to each of the $d(d-1)$ pairs of roots, z_i, z_j, $i \neq j$; the total probability is no more than the sum of the $d(d-1)$ probabilities $c'\delta^4 d^3$.

Finally we arrive at our main theorem:

Theorem 4.12 *For a_i's distributed as independent standard normals,*

$$\Pr\{Q_d(\epsilon)\} \leq c\epsilon^2 d^7 + 2^{-c'd}$$

for some constants $c, c' > 0$.

Proof By corollary 4.11 and corollary 4.3 we have that with probability $\geq 1 - c\delta^4 d^3 - 2^{-c'd}$ we have at least $d/2$ of the roots z lying in the range $1/2 \leq |z| \leq 2$ and each such root is separated from the others by a distance δ. By lemma 2.1 this guarentees for each of these $d/2$ roots an approximate zero region of area $\geq c\delta^2/(d-1)^2$, for a total of $\geq c\delta^2/d$. Setting $\epsilon = c\delta^2/d$ we get $\delta^4 d^5 = c\epsilon^2 d^7$ and the theorem follows.

5 Uniform and Some Other Distributions

In this section we obtain estimates like those of the previous section for coefficients distributed independently according to some other distribution. We will assume the distribution is the uniform distribution in $B_1(0)$ for a_i with $i < d$ and $a_d = 1$. In fact, one can do the same estimates verbatum with $a_d = 1$ and for $i < d$ taking a_i according to any distribution supported in $B_1(0)$, possibly different for different i's, which satisfy equation 5.3 for $\ell = 14$ with uniform bounds on the $m_i's$.

In Smale's works, [Sma81] and [Sma86b], the polynomials $a_d z^d + \cdots + a_0$ are considered with $a_d = 1$ and a_i distributed uniformly in $B_1(0)$ for $i \neq d$. This has the advantage of guarenteeing that the roots lie in the ball of radius 2 (if $|z| > 2$, then clearly $|z^d| > \sum_{i<d} |a_i z^i|$ if $|a_i| \leq 1$ and so such a z cannot be a root of the polynomial). In contrast to the distribution of the previous section, (almost all) such polynomials have regions of approximate zeros in $B_3(0)$, and we will obtain estimates of the form

$$\Pr\{Q(\epsilon)\} \leq c d^7 \epsilon^2. \tag{5.1}$$

We begin by considering the probability measure on polynomials in which $a_d = 1$ and the remaining coefficients distributed uniformly in the unit ball, $B_1(0)$; i.e. with density

$$\psi(z) = \begin{cases} 1/\pi & \text{for } z \in B_1(0) \\ 0 & \text{otherwise} \end{cases}$$

For the density ψ, we have its characteristic function, $\widehat{\psi}$, satisfies

$$|\widehat{\psi}(\xi)| \leq \frac{k}{|\xi|} \; \forall \xi \in \mathbf{C} \tag{5.2}$$

for some k, and

$$\widehat{\psi}(\xi) = 1 - m_1|\xi|^2 - m_2|\xi|^4 - \cdots - m_\ell|\xi|^{2\ell} + O\left(|\xi|^{2\ell+2}\right) \qquad (5.3)$$

for any ℓ (see appendix A).

It will be easier to have all the a_i's radially symmetric, so we will take a_d to be distributed as

$$e^{2\pi i\theta}$$

with θ uniform random variable in $[0,1]$. We denote its characteristic function by

$$\hat{\psi}_1(\xi) = E\left\{e^{i\Re(\langle\xi,a_d\rangle)}\right\}.$$

For $\hat{\psi}_1$ we also have an expansion

$$\hat{\psi}_1(\xi) = 1 - M_1|\xi|^2 - M_2|\xi|^4 - \ldots - M_\ell|\xi|^{2\ell} + O\left(|\xi|^{2\ell+2}\right).$$

We begin by estimating $P(z_1,z_2)$ for $|z_1-z_2|$ small. As in the previous section, by $|z_1-z_2|$ *small* it suffices to take $|z_1-z_2| \leq c|z_1|/d$ for some constant c, but we will only be applying the estimate when $|z_1-z_2| \leq c|z_1|d^{-5/4}$; we will assume the latter for notational convenience.

To estimate $P(z_1,z_2)$, from equation 4.5 we see that it suffices to estimate

$$E\left\{|f'(z_1)|^4 \text{ s.t. } f(z_1)=f(z_2)=0\right\}. \qquad (5.4)$$

We can estimate it as

$$\leq \frac{c|\Delta z|^2}{d(d-1)} \int_{\mathbf{C}} |t|^4 \Upsilon(0,0,t)\,dt \qquad (5.5)$$

where Υ is the joint density of

$$\sum a_i u_i\ ,\ \sum a_i v_i\ ,\ \sum a_i w_i \qquad (5.6)$$

with

$$u_i = z_1^i\ ,\ v_i = iz_1^{i-1}\left(1+O(d^{-1/4})\right)\ ,\ w_i = i(i-1)z_1^{i-2}\left(1+O(d^{-1/4})\right).$$

Let $\tilde{a}_i$ be distributed as $\sqrt{2m_1}$ times the standard normal distribution, and Ξ the distribution of

$$\sum \tilde{a}_i u_i\ ,\ \sum \tilde{a}_i v_i\ ,\ \sum \tilde{a}_i w_i. \qquad (5.7)$$

The main task of this section is to prove:

Theorem 5.1 *There exist constants c and d_0 independent of d, t, z_1, and z_2 such that the following hold. For all z_1 and z_2 with $|z_1 - z_2| \le |z_1|\left(1 + O(d^{-5/4})\right)$, we have if $d \le d_0$ or $|z_1| \le \frac{1}{2}$ then*

$$\int_{\mathbf{C}} |t|^4 \Upsilon(0,0,t)\, dt \le c, \tag{5.8}$$

if $d > d_0$ and $1 - \frac{1}{d} \le |z_1| \le 1 + \frac{1}{d}$ then

$$\Upsilon(0,0,t) \le c\Xi(0,0,t) + c\Xi(0,0,t/2) + cd^{-15}, \tag{5.9}$$

and if $d > d_0$ and $\frac{1}{2} \le |z_1| \le 1 - \frac{1}{d}$ then

$$\Upsilon(0,0,t) \le c\Xi(0,0,t) + c\Xi(0,0,t/2) + c(1 - |z_1|^2)^{-12}. \tag{5.10}$$

This will give estimates on equation 5.5 and thus on equation 5.4 similar to those in §4.

For equation 5.8, notice that for any d_0 there is a c such that for and $d \le d_0$ or if $|z_1| \le \frac{1}{2}$ we have

$$\begin{aligned} \int |t|^4 \Upsilon(0,0,t)\, dt &= E\left\{\left|\sum a_i w_i\right|^4 \text{ s.t. } \sum a_i u_i = \sum a_i v_i = 0\right\} \\ &\le \max\left|\sum a_i w_i\right|^4 \Upsilon(0,0) \le c \end{aligned}$$

where $\Upsilon(0,0)$ is the joint density of $\sum a_i u_i$, $\sum a_i v_i$ at $(0,0)$, since

$$\sum |a_i||v_i| \le \sum |v_i|$$

which is bounded uniformly for such $|z_1|$, and

$$\begin{aligned} \sum a_i u_i &= a_0 + z_1 a_1 + \cdots \\ \sum a_i v_i &= a_1 + \cdots \end{aligned}$$

so that

$$\begin{aligned} & \Pr\left\{\sum a_i u_i \in B_\epsilon(0)\ ,\ \sum a_i v_i \in B_\epsilon(0)\right\} \\ = \ & \Pr\left\{a_0 \in B_\epsilon(l_0)\ ,\ a_1 \in B_\epsilon(l_1)\right\} \\ \le \ & \epsilon^4, \end{aligned}$$

where l_0, l_1 are linear combinations of $a_2, \ldots a_d$, and so

$$\Upsilon(0,0) \leq \left(\frac{1}{\pi}\right)^2 .$$

Next we estimate $\Upsilon(0,0,t)$ for $1 \leq |z_1| \leq 1 + \frac{1}{d}$.

It will be convenient to rescale u, v, and w via

$$U_i \equiv \frac{u_i}{z_1^d} \ , \ V_i \equiv \frac{v_i}{d z_1^{d-1}} \ , \ W_i \equiv \frac{w_i}{d(d-1) z_1^{d-2}},$$

and set

$$Y_i \equiv (U_i, V_i, W_i) \approx z_1^{i-d} \left(1, \frac{i}{d}, \frac{i^2}{d^2}\right).$$

Note that $|Y_i| \leq c$ for some constant c independent of i and d. We will obtain estimates as in equation 5.9, with Ψ being the density of

$$\sum a_i U_i \ , \ \sum a_i V_i \ , \ \sum a_i W_i \tag{5.11}$$

and Φ the density of equation 5.11 with a_i replaced by $\tilde{a}_i$.

Consider, for each j, $(a_j U_j, a_j V_j, a_j W_j) \in \mathbf{C}^3 \simeq \mathbf{R}^6$. Since the characteristic function of a_j for $j > d$ is

$$\widehat{\psi}(\xi) = E\left\{e^{i\Re(\xi \bar{a}_j)}\right\},$$

the characteristic function of $(a_j U_j, a_j V_j, a_j W_j)$ is

$$\begin{aligned} \tilde{\chi}(\eta, \sigma, \tau) &= E\left\{e^{i\Re(\eta \overline{a_j u U j} + \sigma \overline{a_j V_j} + \tau \overline{a_j W_j}}\right\} \\ &= E\left\{e^{i\Re\left[(\eta \bar{U}_j + \sigma \bar{V}_j + \tau \bar{W}_j)\overline{a}_j\right]}\right\} \\ &= \widehat{\psi}\left(|\eta \bar{U}_j + \sigma \bar{V}_j + \tau \bar{W}_j|\right). \end{aligned}$$

It follows that the characteristic function of the joint density of equation 5.6 is

$$\widehat{\Psi} = \widehat{\psi}_1\left(|\eta \bar{U}_j + \sigma \bar{V}_j + \tau \bar{W}_j|\right) \prod_{j=0}^{d-1} \widehat{\psi}\left(|\eta \bar{U}_j + \sigma \bar{V}_j + \tau \bar{W}_j|\right),$$

and its density is

$$\Psi(q,s,t) = \left(\frac{1}{2\pi}\right)^6 \int_{\mathbf{C}^3} e^{-i\Re(q\bar{\eta} + s\bar{\sigma} + t\bar{\tau})} \widehat{\Psi}(\eta, \sigma, \tau)\, d\eta\, d\sigma\, d\tau.$$

Let $\xi = (\eta, \sigma, \tau)$. Then

$$\widehat{\Psi}(\xi) = \widehat{\psi}_1 \left(| \langle \xi, Y_d \rangle |\right) \prod_{i=0}^{d-1} \widehat{\psi} \left(| \langle \xi, Y_i \rangle |\right).$$

From equation 5.3 it follows that

$$\widehat{\psi}(\xi) = e^{-m_1|\xi|^2 + m'_2|\xi|^4 + \cdots + m'_7|\xi|^{14}} + O\left(|\xi|^{16}\right)$$

and

$$\widehat{\psi}_1(\xi) = e^{-M_1|\xi|^2 + M'_2|\xi|^4 + \cdots + M'_7|\xi|^{14}} + O\left(|\xi|^{16}\right)$$

for $|\xi|$ small for some constants $m'_2, \ldots, m'_7$ and $M'_2, \ldots, M'_7$. Let δ be a small positive number with $22\delta \leq 1$. Let

$$B \equiv B_{d^{\delta - 1/2}}(0).$$

Let $\widehat{\omega}$ be defined by

$$\widehat{\omega}(\xi) = \begin{cases} e^{-m|\xi|^2 + m'_2|\xi|^4 + \cdots + m'_7|\xi|^{14}} & \text{for } z \in B \\ 0 & \text{otherwise} \end{cases},$$

and let $\widehat{\omega}_1$ be defined by replacing the m's by M's. Let Ω be given by

$$\widehat{\Omega}(\xi) = \widehat{\omega}_1 \left(| \langle \xi, Y_d \rangle |\right) \prod_{i=0}^{d-1} \widehat{\omega} \left(| \langle \xi, Y_i \rangle |\right).$$

Henceforth we will replace ϕ_1 and ω_1 by ψ and ω when we write our equations. It makes no difference in the analysis.

We reduce the study of Ψ to that of Ω.

Lemma 5.2 *For some constant c we have for any z_1 with $1 \leq |z_1| \leq 1 + \frac{1}{d}$,*

$$|\Psi(0,0,t) - \Omega(0,0,t)| \leq cd^{-9} \ \forall t \in \mathbf{C}.$$

Proof By the Fourier inversion formula,

$$\begin{aligned} |\Psi(0,0,t) - \Omega(0,0,t)| &\leq \left(\frac{1}{2\pi}\right)^6 \int_{\mathbf{C}^3} \left|[\widehat{\Psi}(\xi) - \widehat{\Omega}(\xi)] e^{-i\Re\langle x, \xi \rangle}\right| d\xi \\ &\leq \left(\frac{1}{2\pi}\right)^6 \int_{\mathbf{C}^3} \left|\widehat{\Psi}(\xi) - \widehat{\Omega}(\xi)\right| d\xi \end{aligned}$$

$$\leq \left(\frac{1}{2\pi}\right)^6 \left[\int_{\mathbf{C}^3-B} |\widehat{\Psi}(\xi)|\, d\xi + \int_B |\widehat{\Psi}(\xi) - \widehat{\Omega}(\xi)|\, d\xi\right]. \qquad (5.12)$$

To estimate the second integral of equation 5.12, we have

$$\int_B |\widehat{\Psi} - \widehat{\Omega}|\, d\xi = \int_B \left|\prod_{i=0}^{d} \widehat{\psi}(\langle \xi, Y_i\rangle) - \prod_{i=0}^{d} \widehat{\omega}(\langle \xi, Y_i\rangle)\right|\, d\xi$$

which, by lemma A.1, is

$$\leq \int_B \sum \left|\widehat{\psi}(\langle \xi, Y_i\rangle) - \widehat{\omega}(\langle \xi, Y_i\rangle)\right|\, d\xi \leq c \int_B \sum |\langle \xi, Y_i\rangle|^{16}\, d\xi$$

which, by Cauchy-Schwartz and since the $|Y_i|$ are bounded independent of i and d,

$$\begin{aligned} \leq c \int_B \sum |\xi|^{16}\, d\xi &\leq c|B| d \left(d^{\delta-1/2}\right)^{16} = c\left(d^{\delta-1/2}\right)^{22} d \\ &= c d^{22\delta-10} \leq c d^{-9} \end{aligned} \qquad (5.13)$$

since $22\delta \leq 1$, where we have used the fact that $|B_r(0)| = cr^6$ for balls in $\mathbf{C}^3$.

To estimate the first integral of equation 5.12, let

$$\begin{aligned} D_0 &\equiv \left\{ i \in \mathbf{Z} \,:\, \frac{5}{6}d - 2 \leq i < d \right\}, \\ D_1 &\equiv \left\{ i \in \mathbf{Z} \,:\, \frac{3}{6}d - 1 \leq i \leq \frac{4}{6}d \right\}, \\ D_2 &\equiv \left\{ i \in \mathbf{Z} \,:\, \frac{1}{6}d - 1 \leq i \leq \frac{2}{6}d \right\}. \end{aligned}$$

Note that for d sufficiently large the D_j's are disjoint and each D_j contains $\geq d/6$ integers. Let d_0 be an integer such that this is the case for $d > d_0$. This will be our d_0 in theorem 5.1. We will use

Sublemma 5.3 *Let $i \in D_0$, $j \in D_1$, $k \in D_2$. Then for each $\xi \in \mathbf{C}^3$, either*

$$|\langle \xi, Y_i\rangle| \ , \ |\langle \xi, Y_j\rangle| \ , \ \textit{or} \ |\langle \xi, Y_k\rangle| \ \textit{is} \ \geq c_2|\xi|$$

for some absolute constant c_2 (independent of ξ and z_1 with $1 \leq |z_1| \leq 1+\frac{1}{d}$).

Proof This is an easy calulation. For d small one can use compactness in $z_1 \in B_{1+\frac{1}{d}}(0) - B_1(0)$. For d large we have

$$Y_i \approx z_1^{i-d}(1, \theta, \theta^2)$$

with $\frac{1}{6} - \frac{1}{d} \le \theta \le \frac{2}{6}$; note that $z_1^{i-d} \in B_1(0) - B_{1/e}(0)$ since $|z_1|^{-d} \ge \left(1 + \frac{1}{d}\right)^{-d} \ge 1/e$. Similar estimates hold for Y_j and Y_k in which θ takes on a different range of values, leading to the desired estimate.

Let $A = B_{2kc_2}(0)$ where c_2 is the constant in lemma 5.3 and k is the constant in lemma 5.2. We write

$$\int_{\mathbf{C}^3 - B} \left|\widehat{\Psi}(\xi)\right| d\xi = \int_{\mathbf{C}^3 - A} \left|\widehat{\Psi}(\xi)\right| d\xi + \int_{A-B} \left|\widehat{\Psi}(\xi)\right| d\xi. \tag{5.14}$$

In $\mathbf{C}^3 - A$ we use

$$|\widehat{\Psi}(\xi)| = \prod_{i=0}^{d} |\widehat{\psi}(\langle \xi, Y_i \rangle)| \le \left(\frac{k}{c_2|\xi|}\right)^{d/6},$$

which follows from equation 5.2 and lemma 5.3, to obtain

$$\int_{\mathbf{C}^3 - A} |\widehat{\Psi}(\xi)| \, d\xi \le \int_{\mathbf{C}^3 - A} \left(\frac{k}{c_2|\xi|}\right)^{d/6} d\xi \le c \left(\frac{1}{2}\right)^{d/6}. \tag{5.15}$$

In $A - B$, we take a constant k'' with the property that

$$|\widehat{\psi}(\xi)| \le e^{-|\xi|^2 k''} \ \forall \xi \in A,$$

and estimate

$$|\widehat{\Psi}(\xi)| \le \prod_{i=0}^{d} e^{-k''|\langle \xi, Y_i \rangle|^2} \le e^{-k'' \frac{d}{6} |\xi|^2 c_2^2} = e^{-cd|\xi|^2}$$

so that

$$\int_{A-B} \left|\widehat{\Psi}(\xi)\right| d\xi \le \int_{A-B} e^{-cd|\xi|^2} \, d\xi \le c' e^{-cd^{2\delta}}. \tag{5.16}$$

Combining equations 5.12, 5.13, 5.14, 5.15, and 5.16 yields lemma 5.2.

To deal with Ω, note that for $\xi \in B$,

$$\widehat{\Omega}(\xi) = e^{-Q_2(\xi) - \cdots - Q_{14}(\xi)},$$

where the Q_i are homogeneous polynomials of degree i in $\xi = (\eta, \sigma, \tau)$ given by

$$Q_{2k}(\xi) = \sum_{i=0}^{d} m'_k \left|\langle \xi, Y_i \rangle\right|^{2k},$$

where $m'_1 = m_1$. Note that

$$Q_2(\xi) \geq m_1 c_2^2 \frac{d}{6} |\xi|^2 \geq cd|\xi|^2$$

and that

$$|Q_{2k}(\xi)| \leq \sum_{i=0}^{d} m'_k |\xi|^{2k} |Y_i|^{2k} \leq cd|\xi|^{2k}.$$

Expanding by power series we get

$$\begin{aligned} \widehat{\Omega}(\xi) &= e^{-Q_2(\xi) - \cdots - Q_{14}(\xi)} \\ &= e^{-Q_2(\xi)} \left(1 + R_4(\xi) + \cdots + R_{14}(\xi)\right) + O\left(|\xi|^{16}\right), \end{aligned}$$

where $R_k(\xi)$ are homogeneous polynomials of degree $2k$ in ξ. Since

$$|Q_{2k}(\xi)| \leq cd^{k/2} |\xi|^{2k}$$

for $2k \geq 4$, we have

$$|R_{2k}(\xi)| \leq c' d^{k/2} |\xi|^{2k} \tag{5.17}$$

for some c', and hence

$$|R_{2k}(\xi)| \leq cd^3 |\xi|^{2k} \tag{5.18}$$

for $k = 2, \ldots, 7$ (in equation 5.17, we can replace fractional powers of d by the nearest lower integral power of d).

Let Θ be given by

$$\widehat{\Theta}(\xi) = e^{-Q_2(\xi)} \left(1 + R_4(\xi) + \cdots + R_{14}(\xi)\right).$$

We finish the proof of lemma 5.2 with:

Lemma 5.4

$$|\Omega(0,0,t) - \Theta(0,0,t)| \leq cd^{-9}.$$

Lemma 5.5

$$|\Theta(0,0,t)| \le c\Phi(0,0,t).$$

The proof of lemma 5.4 is the same as the proof of lemma 5.2. Note

$$\int_{\mathbf{C}^3} \left|\widehat{\Omega} - \widehat{\Theta}\right| d\xi = \int_{\mathbf{C}^3 - B} \left|\widehat{\Theta}\right| d\xi + \int_B \left|\widehat{\Omega} - \widehat{\Theta}\right| d\xi.$$

Similar to the proof of lemma 5.2, we can estimate

$$\int_B \left|\widehat{\Omega} - \widehat{\Theta}\right| d\xi \le cd^{-9}$$

and estimate

$$\begin{aligned} \int_{\mathbf{C}^3 - B} \left|\widehat{\Theta}\right| d\xi &\le \int_{\mathbf{C}^3 - A} + \int_{A - B} \\ &\le cd^3 \left[\left(\frac{1}{2}\right)^{d/6} + e^{-c' d^{2\delta}} \right], \end{aligned}$$

the d^3 coming from equation 5.18.

The proof of lemma 5.5 is a straightforward calculation. We have

$$\widehat{\Phi}(\xi) = \prod_{i=0}^{d} e^{-m_1 |\langle \xi, Y_i \rangle|^2} = e^{-Q_2(\xi)}.$$

It follows that

$$\Theta(x) = \left(1 + R_4\left(i\frac{\partial}{\partial x}\right) + \cdots + R_{14}\left(i\frac{\partial}{\partial x}\right)\right)\Phi(x).$$

Sublemma 5.6 *Let B be an $n \times n$, complex Hermitian matrix, and let*

$$g(x) = e^{-\left\langle B^{-1}x, x\right\rangle}.$$

Then for any muti-index α we have

$$\left|\frac{\partial^\alpha g}{\partial x^\alpha}(t)\right| \le \left\|B^{-1}\right\|^{|\alpha|/2} cg(t/2) \quad \forall t$$

for some constant $c = c(\alpha, n)$ independent of B.

Proof Since B is Hermitian, it suffices to prove it assuming B is diagonal. To prove it for B diagonal it suffices to prove it for the one variable case. In the one variable case,

$$g(t) = e^{-t^2/b}$$

and

$$\frac{\partial^\alpha g}{\partial x^\alpha}(t) = \left(\frac{1}{b}\right)^{|\alpha|/2} P_\alpha\left(\frac{t}{\sqrt{b}}\right) e^{-t^2/b}$$

where P_α is a polynomial of degree α. We have

$$P_\alpha\left(\frac{t}{\sqrt{b}}\right) e^{-\frac{3}{4}t^2/b} < c(\alpha)$$

for some constant $c(\alpha)$ for each α, and thus the sublemma follows.

Taking B to be the matrix given by

$$\langle B\xi, \xi\rangle = Q_2(\xi)$$

we get that since $Q_2(\xi) \geq cd|\xi|^2$ that $\|B^{-1}\| \leq c/d$ and thus

$$R_{2k}\left(i\frac{\partial}{\partial x}\right)\Phi(x) \leq cd^{k/2}\left|\left(i\frac{\partial}{\partial x}\right)^{2k}\Phi(x)\right| \leq cd^{k/2}d^{-k}\Phi(x/2)$$

for $k \geq 2$ and lemma 5.5 follows.

Combining lemmas 5.4, 5.5, and 5.2 we get

$$\Psi(0,0,t) \leq c\Phi(0,0,t) + c\Phi(0,0,t/2) + cd^{-9}$$

for all t for some c for $1 \leq |z_1| \leq 1+\frac{1}{d}$. Changing from U_i, V_i, W_i to u_i, v_i, w_i sums yields the desired estimate.

The same estimates hold for $1-\frac{1}{d} \leq |z_1| \leq 1$. One can see this directly, or by using the same trick as in corollary 4.9.

Next we estimate for $\frac{1}{2} \leq |z_1| \leq 1-\frac{1}{d}$. Let m be the largest integer such that

$$|z_1|^m \geq \frac{1}{2}.$$

We remark that

$$m = \theta(-\log|z_1|) = \theta(1-|z_1|) = \theta(1-|z_1|^2)$$

(where $f = \theta(g)$ means $c_1 g \leq f \leq c_2 g$ for some positive constants c_1 and c_2). We claim that

$$\Upsilon(0,0,t) \leq c\Xi(0,0,t) + c\Xi(0,0,t/2) + cm^{-12}$$

which is the same as equation 5.10. To see this we go through estimates similar to those for $1 \leq |z_1| \leq 1 + \frac{1}{d}$. We rescale

$$U_i \equiv u_i \ , \ V_i \equiv \frac{z_1 v_i}{m} \ , \ W_i \equiv \frac{z_1^2 w_i}{m^2}$$

and let

$$Y_i = (U_i, V_i, W_i).$$

Then we have

$$|Y_i| \leq ce^{-c'(i/m)}$$

for positive constants c and c' independent of i and z_1. We define Ψ and Φ as before. Set

$$B \equiv B_{m^{\delta-1/2}}(0)$$

and define ω and Ω as before. As before we get

$$|\Psi(0,0,t) - \Omega(0,0,t)| \leq cm^{-9}$$

using

$$\sum_{i=0}^{d} |Y_i|^{16} \leq c \sum_{i=0}^{d} |z_1|^{16i} \leq c \sum_{i=0}^{\infty} |z_1|^{16i} \leq cm.$$

Next let

$$\begin{aligned} D_0 &\equiv \{i \in \mathbf{Z} \ : \ m \leq i \leq 2m\}\,, \\ D_1 &\equiv \{i \in \mathbf{Z} \ : \ 3m \leq i \leq 4m\}\,, \\ D_2 &\equiv \{i \in \mathbf{Z} \ : \ 5m \leq i \leq 6m\}\,. \end{aligned}$$

Then sublemma 5.3 holds for these D_0, D_1, D_2. Defining Θ and Q_i and R_i as before we have

$$Q_2(\xi) \geq cm|\xi|^2$$

and

$$|Q_{2k}(\xi)| \leq cm|\xi|^{2k}$$

and all the estimates go through as before to yield

$$\Psi(0,0,t) \le c\Psi(0,0,t) + c\Psi(0,0,t/2) + cm^{-9}.$$

Upon rescaling to get Υ and Ξ we get the desired result.

Corollary 5.7 *If $|z_1 - z_2| \le |z_1| O\left(d^{-5/4}\right)$, then we have*

$$P(z_1, z_2) \le \frac{c}{d(d-1)} (\Delta z)^2 \begin{cases} d^6 & \text{if } 1 - \frac{1}{d} \le |z_1| \le 1 + \frac{1}{d} \\ (1 - |z_1|^2)^{-6} & \text{for other } |z_1| \in [0, 2] \end{cases} . \tag{5.19}$$

Proof If $1 - \frac{1}{d} \le |z_1| \le 1 + \frac{1}{d}$ then

$$|\sum a_i w_i| \le c \sum |w_i| \le cd^3$$

so that $\Upsilon(0,0,t) = 0$ for $|t| > cd^3$ and so

$$\begin{aligned} \int \Upsilon(0,0,t)|t|^4\, dt &= \int_{B_{cd^3}(0)} \Upsilon(0,0,t)|t|^4\, dt \\ &\le c \int_{B_{cd^3}(0)} \left(\Xi(0,0,t) + \Xi(0,0,t/2) + d^{-12}\right) |t|^4\, dt \\ &\le c \int_{\mathbf{C}} \left(\Xi(0,0,t) + \Xi(0,0,t/2)\right) |t|^4\, dt + cd^6 \end{aligned}$$

and using the estimates on Ξ in §4 the above is

$$\le cd^6.$$

For $\frac{1}{2} \le |z_1| \le 1 - \frac{1}{d}$ we have

$$\sum |w_i| \le cm^3$$

and the same estimates as before yield the theorem. For $1 + \frac{1}{d} \le |z_1| \le 2$ we use the same trick as in corollary 4.9 and note that the same estimates hold for $1/z_1$ and the equation

$$\sum_{i=0}^{d} a_{d-i} z^i = 0$$

(note that the fact that $a_d = 1$ doesn't affect the estimates since D_0, D_1, D_2 never contain a_0 or a_d). For $0 \le |z_1| \le \frac{1}{2}$ we use theorem 5.1 to get the desired result.

We can finally prove

Theorem 5.8

$$\Pr\{Q_d(\epsilon)\} \le c\epsilon^2 d^7.$$

Proof By integrating corollary 5.7 as in §4 we get that for any constant k there is a constant c such that

$$\Pr\Big\{|z_1 - z_2| \le \delta \text{ and } k\delta d^{5/4} \le |z_1| \le 2\Big\} \le c\delta^4 d^3.$$

We need $|z_1| \ge k\delta d^{5/4}$ to apply corollary 5.7. By lemma 4.1 we see that for $f(z)$ to have $\ge d/2$ roots of absolute value $\le k\delta d^{5/4}$ would imply

$$(k\delta d^{5/4})^{d/2} \le \frac{1}{|a_0|}\sqrt{|a_0|^2 + \cdots + |a_d|^2} \le \frac{d+1}{|a_0|}$$

so that

$$|a_0| \le \frac{d+1}{(k\delta d^{5/4})^{d/2}}$$

which happens with probabilty

$$\le \frac{cd^2}{(k\delta d^{5/4})^d}$$

which is dominated by $\delta^2 d^5$ if $\delta < d^{-5/4}$ and if we take, say, $k = \frac{1}{2}$. Thus the probability that some root in $B_2(0) - B_{k\delta d^{5/4}}(0)$ is within δ of another or that there are less that $d/2$ roots in $B_2(0) - B_{k\delta d^{5/4}}(0)$ is $\le c\delta^4 d^5$. When this is not the case then lemma 2.1 guarentees a total approximate zero region of area $\ge c'\delta^2 d$. Setting $\epsilon = c''\delta^2/d$ yields the theorem.

6 A Refined Estimate

In this section improve the estimates of the previous two sections by considering the joint density of three or more roots, similar to the latter part of §3.

Theorem 6.1 *Let $z_1, \ldots, z_k \in \mathbf{C}$ satisfy $|z_i - z_1| = |z_1| O(d^{-1-\beta})$ for some fixed β. Then*

$$P(z_1, \ldots, z_k) \leq \frac{c}{d(d-1)\ldots(d-k+1)} \prod_{i<j} |z_j - z_i|^2 \left\{ \begin{array}{ll} d^{k(k+1)} & \textit{if } 1-\frac{1}{d} \leq |z_1| \leq 1+\frac{1}{d} \\ (1-|z_1|)^{-k(k+1)} & \textit{for other } |z_1| \in [0,2] \end{array} \right. . \quad (6.1)$$

Proof The calculations are similar to the ones done in §4 and §5. We wish to estimate

$$E\left\{|f'(z_1)|^2 \ldots |f'(z_k)|^2 \text{ s.t. } f(z_1) = \ldots = f(z_k) = 0\right\}.$$

As in theorem 5.1, it suffices to prove the theorem for $\frac{1}{2} \leq |z_1| \leq 1$ and d sufficiently large; for $0 \leq |z_1| \leq \frac{1}{2}$ and small d the theorem will be clear from the estimates used elsewhere and the bounded sums of the linear combinations of the a_i's involved, and for $1 \leq |z_1| \leq 2$ we have that the $1/z_i$'s satisfy the equation with coefficients reversed and the $\frac{1}{2} \leq |z_1| \leq 1$ estimates can be invoked.

For convenience, let

$$m = \left\{ \begin{array}{ll} d & \text{if } 1-\frac{1}{d} \leq |z_1| \leq 1 \\ \lfloor -\log_2 |z_1| \rfloor & \text{if } \frac{1}{2} \leq |z_1| < 1-\frac{1}{d} \end{array} \right. ,$$

where $\lfloor a \rfloor$ denotes the largest integer $\leq a$.

We first deal with the case of a_i being distributed normally. It suffices to estimate

$$E\left\{|F(z_1)|^2 \ldots |F(z_k)|^2\right\}$$

and

$$\psi(0, \ldots, 0)$$

where ψ is the density of $f(z_1), \dots, f(z_k)$ and where $F(z_i)$ is the random variable $f'(z_i)$'s projection in $(\mathbf{C}f(z_1) + \cdots + \mathbf{C}f(z_k))^{\perp}$.

For the latter, note that setting

$$u_i = (1, z_i, z_i^2, \dots, z_i^d)$$

we have, first of all,

$$u_1 = (1, z_1, \dots, z_1^d).$$

Secondly, if we set $u'_j = \frac{u_j - u_1}{z_j - z_1}$ for $j > 1$, then we have

$$u'_2 = (0, 1, 2z_1, \dots, dz_1^{d-1}) \left(1 + O(d^{-\beta})\right).$$

Next, if we set $u''_j = \frac{u'_j - u'_2}{z_j - z_2}$ for $j > 2$, then we have

$$u''_3 = (0, 0, 2, 6z_1, \dots, d(d-1)z_1^{d-2}) \left(1 + O(d^{-\beta})\right).$$

Continuing in this fashion, we see that

$$\begin{aligned} & \det \begin{pmatrix} \langle u_1, u_1 \rangle & \cdots & \langle u_1, u_k \rangle \\ \vdots & \ddots & \vdots \\ \langle u_k, u_1 \rangle & \cdots & \langle u_k, u_k \rangle \end{pmatrix} \\ = & \prod_{j>i} (z_j - z_i)^2 \det \begin{pmatrix} \langle u_1, u_1 \rangle & \langle u_1, u'_1 \rangle & \cdots & \left\langle u_1, u_1^{(k-1)} \right\rangle \\ \langle u'_1, u_1 \rangle & \langle u'_1, u'_1 \rangle & \cdots & \left\langle u'_1, u_1^{(k-1)} \right\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \left\langle u_1^{(k-1)}, u_1 \right\rangle & \left\langle u_1^{(k-1)}, u'_1 \right\rangle & \cdots & \left\langle u_1^{(k-1)}, u_1^{(k-1)} \right\rangle \end{pmatrix} \end{aligned}$$

where

$$u_1^{(j)} = (0, \dots, 0, j!, \dots, d(d-1)\dots(d-j+1)z_1^{d-j}) \left(1 + O(d^{-\beta})\right).$$

We claim that

$$\det \begin{pmatrix} \langle u_1, u_1 \rangle & \langle u_1, u'_1 \rangle & \cdots & \left\langle u_1, u_1^{(k-1)} \right\rangle \\ \langle u'_1, u_1 \rangle & \langle u'_1, u'_1 \rangle & \cdots & \left\langle u'_1, u_1^{(k-1)} \right\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \left\langle u_1^{(k-1)}, u_1 \right\rangle & \left\langle u_1^{(k-1)}, u'_1 \right\rangle & \cdots & \left\langle u_1^{(k-1)}, u_1^{(k-1)} \right\rangle \end{pmatrix} \geq cm^{k^2}. \quad (6.2)$$

If we expand the above determinant into a sum of products of the entries of the matrix, each entry has size proportional to m^{k^2}. It suffices to show that equaiton 6.2 with the $u_1^{(j)}$ replaced by

$$v_j = (0, \dots, 0, j!, \dots, d(d-1)\dots(d-j+1)z_1^{d-j})$$

(i.e. the old $u_1^{(j)}$'s without the error term $\left(1+O(d^{-\beta})\right)$). Next note that

$$\begin{aligned}
\det\begin{pmatrix} \langle v_0, v_0\rangle & \cdots & \langle v_0, v_{k-1}\rangle \\ \vdots & \ddots & \vdots \\ \langle v_{k-1}, v_0\rangle & \cdots & \langle v_{k-1}, v_{k-1}\rangle \end{pmatrix} &= \det\begin{pmatrix} \langle v_0, v_0\rangle & \langle v_0, \tilde{v}_1\rangle & \cdots & \langle v_0, \tilde{v}_{k-1}\rangle \\ \langle \tilde{v}_1, v_0\rangle & \langle \tilde{v}_1, \tilde{v}_1\rangle & \cdots & \langle \tilde{v}_1, \tilde{v}_{k-1}\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \tilde{v}_{k-1}, v_0\rangle & \langle \tilde{v}_{k-1}, \tilde{v}_1\rangle & \cdots & \langle \tilde{v}_{k-1}, \tilde{v}_{k-1}\rangle \end{pmatrix} \\
&= \langle v_0, v_0\rangle \langle \tilde{v}_1, \tilde{v}_1\rangle \dots \langle \tilde{v}_{k-1}, \tilde{v}_{k-1}\rangle
\end{aligned}$$

where $\tilde{v}_j$ is the projection of v_j onto $(\mathbf{C}v_0 + \cdots + \mathbf{C}v_{j-1})^{\perp}$, i.e.

$$\tilde{v}_j = v_j - \sum_{i=0}^{j-1} \frac{\langle v_j, v_i\rangle}{\langle v_i, v_i\rangle} v_i.$$

Lemma 6.2 *For any n there exist $c, d_0 > 0$ such that for any $d > d_0$ and $\alpha_1, \dots, \alpha_n$ we have*

$$|t^n - \alpha_1 t^{n-1} - \cdots - \alpha_n| \geq cd^n \tag{6.3}$$

for at least $d/4$ of the integers $t = 1, 2, \dots, d$.

Proof We can write

$$t^n - \alpha_1 t^{n-1} - \cdots - \alpha_n = (t-\gamma_1)\dots(t-\gamma_n)$$

for some $\gamma_i \in \mathbf{C}$. For each i we have

$$|t-\gamma_i| \geq |\Re(t-\gamma_i)| \geq \frac{1}{9n}d$$

for all integers t, $\frac{d}{2} - 1 \leq t \leq d$, except for possibly $\frac{2}{9n}d + 1$ values of t. If d is sufficiently large we have

$$\frac{2}{9n}d + 1 \leq \frac{1}{4n}d,$$

and thus for $d/4$ values of t, $\frac{d}{2} - 1 \leq t \leq d$, we have

$$|t^n - \alpha_1 t^{n-1} - \cdots - \alpha_n| \geq cd^n$$

where $c = 1/(9n)^n$.

Corollary 6.3 *For any k there is a $c > 0$ such that*

$$\langle \tilde{v}_j, \tilde{v}_j \rangle \geq cm^{2j+1} \tag{6.4}$$

for $j = 0, 1, \ldots, k-1$.

Proof We have

$$\langle \tilde{v}_j, \tilde{v}_j \rangle = \sum_{i=j}^{d} |i^j - \alpha_1 i^{j-1} - \cdots - \alpha_j|^2 |z_1|^{2(i-j)}$$

for some $\alpha_1, \ldots, \alpha_j$. For sufficiently large d we can estimate this sum as

$$\geq \sum_{\frac{m}{2}-1 \leq i \leq m} (ci^j)^2 |z_1|^{2m} \geq c' m^{2j+1}.$$

Corollary 6.3 establishes equation 6.2, and thus

$$\begin{aligned} \psi(0, \ldots, 0) &\leq c \left| \prod_{j>i} (z_j - z_i)^2 \; d^{1+3+\cdots+(2k+1)} \right|^{-1} \\ &= c \left| \prod_{j>i} (z_j - z_i)^2 \; d^{k^2} \right|^{-1}. \end{aligned}$$

To estimate

$$E\left\{ |F(z_1)|^2 \ldots |F(z_k)|^2 \right\}$$

we notice that the equations

$$\begin{aligned} \frac{f(z_j) - f(z_1)}{z_j - z_1} &= f'(z_1) + \frac{z_j - z_1}{2} f''(z_1) + \cdots + \frac{(z_j - z_1)^{k-2}}{(k-1)!} f^{(k-1)}(z_1) \\ &+ \frac{(z_j - z_1)^{k-1}}{k!} f^{(k)}(z_1) \left(1 + O(d^{-\beta})\right) \end{aligned}$$

enables us to write

$$|f'(z_1) - L| \leq c \left(\prod_{j>1} |z_j - z_1| \right) |f^{(k)}(z_1)|$$

where L is a linear combination of $f(z_1), \ldots, f(z_n)$. This is true because the linear combination of the above $k-1$ equations which eliminates the

$f''(z_1), \ldots, f^{(k-1)}(z_1)$ terms and gives an $f'(z_1)$ term with coefficient 1 on the right hand side is the linear combination gotten by taking α_j times the z_j equation and adding them, where the α_j's satisfy

$$\begin{pmatrix} 1 & \cdots & 1 \\ \frac{z_2-z_1}{2} & \cdots & \frac{z_k-z_1}{2} \\ \vdots & \ddots & \vdots \\ \frac{(z_2-z_1)^{k-2}}{(k-1)!} & \cdots & \frac{(z_k-z_1)^{k-2}}{(k-1)!} \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Using Kramer's rule and solving van der Monde determinants yields

$$\alpha_i = \prod_{j \neq i,1} \frac{z_j - z_1}{z_j - z_i}.$$

The $f^{(k)}(z_1)$ term in the linear combination therefore has coefficient

$$\begin{aligned} & \frac{1}{k!} \sum_{i>1} (z_i - z_1)^{k-1} \prod_{j \neq i,1} \frac{z_j - z_1}{z_j - z_i} \\ = \;& \frac{1}{k!} \left(\prod_{i>1} (z_i - z_1) \right) \sum_{i>1} \frac{(z_i - z_1)^{k-2}}{\prod_{j \neq i,1} (z_j - z_i)} \\ = \;& \frac{1}{k!} \left(\prod_{i>1} (z_i - z_1) \right) 1 \end{aligned}$$

since

$$\sum_{i>1} \frac{(z_i - z_1)^{k-2}}{\prod_{j \neq i,1} (z_j - z_i)} = \frac{\sum_{i>1} \left[(-1)^{i+1} (z_i - z_1)^{k-2} \prod_{1<n<l,\, n,l \neq i} (z_l - z_n) \right]}{\prod_{1<n<l} (z_l - z_n)}$$

$= 1$ since the numerator is a polynomial, with the same $z_i^{k-2} \prod_{1<n<l,\, n,l\neq i}(z_l - z_n)$ coefficient as in $\prod_{1<n<l}(z_l - z_n)$, and the numerator vanishes whenever $z_l = z_n$ for some $l \neq n$.

Hence we may write

$$E\left\{ |F(z_1)|^2 \ldots |F(z_k)|^2 \right\} \leq c \left(\prod_{j>i} |z_j - z_i|^4 \right) E\left\{ |f^{(k)}(z_1)|^2 \ldots |f^{(k)}(z_k)|^2 \right\}$$

where the analogue of proposition 4.6 for $2k$th powers was used, and by Minkowski's inequality the above is

$$\begin{aligned} &\le c \left(\prod_{j>i} |z_j - z_i|^4 \right) \sum_{i=1}^{k} E\left\{ |f^{(k)}(z_i)|^{2k} \right\} \\ &\le c \left(\prod_{j>i} |z_j - z_i|^4 \right) m^{k+2k^2}, \end{aligned}$$

the bound on $E\left\{|f^{(k)}(z_i)|^{2k}\right\}$ coming from expanding the expression as in equation 4.7 and the preceding equation. Combining the above and the estimate on $\psi(0,\ldots,0)$ yields

$$E\left\{|f'(z_1)|^2 \ldots |f'(z_k)|^2 \text{ s.t. } f(z_1) = \ldots = f(z_k) = 0\right\} \le c \prod_{j>i} |z_j - z_i|^2 \, m^{k(k+1)}$$

which gives the desired result.

For the a_i's distributed uniformly, we do the same estimates as in §5. From the above discussion we see that it suffices to estimate

$$\int_{\mathbf{C}} |t|^{2k} \Upsilon(0,\ldots,0,t)\, dt$$

where Υ is the joint density of $\sum a_i u_i$, $\sum a_i v_i$, ..., $\sum a_i s_i$ with

$$\begin{aligned} u_i &= z_1^i, \\ v_i &= i z_1^{i-1} \left(1 + O(d^{-\beta})\right), \\ &\vdots \\ s_i &= i(i-1)\ldots(i-k+1) z_1^{i-k} \left(1 + O(d^{-\beta})\right). \end{aligned}$$

From here the arguments are just like those in §5.

Corollary 6.4 *The probability that there are k roots within distance $\delta \le d^{-1-\beta}$ of each other in $B_2(0) - B_{\delta d^{-1-\beta}}(0)$ is*

$$\le c \delta^{k^2+k-2} d^{k^2+k-1}.$$

Proof By integrating theorem 6.1

Applying corollary 6.4 to the case of $k = 1$ and to some other value of k yields that each root in $B_2(0) - B_{\max(\delta_1,\delta_2)d^{-1-\beta}}(0)$ has no roots within a distance δ_1 and at most $k-1$ roots within a distance δ_2 with probability $\geq 1 - \tau_1 + \tau_2$, where

$$c\delta_1^4 d^5 = \tau_1$$

and

$$c'\delta_2^{k^2+k-2} d^{k^2+k-1} = \tau_2.$$

If we have at least $d/2$ such roots we get an approximate zero region of area

$$\geq cd \left(\frac{1}{\frac{k-1}{\delta_1} + \frac{d-k}{\delta_2}} \right)^2 \geq cd \min(\delta_1^2, \frac{\delta_2^2}{d}).$$

Setting $\epsilon = d\delta_1^2 = \delta_2^2/d$ yields

$$\tau_1 = c\epsilon^2 d^3,\ \tau_2 = cd(\epsilon d^3)^{\frac{k^2+k-2}{2}}.$$

The probability of having $< d/2$ roots in $B_2(0) - B_{\max(\delta_1,\delta_2)d^{-1-\beta}}(0)$ can be estimated as in the proof of theorem 5.8 and is dominated by $\tau_1 + \tau_2$. Hence we can replace the

$$\epsilon^2 d^7$$

term in equations 4.2 and 5.1 by

$$\epsilon^2 d^3 + \epsilon^{\frac{k^2+k-2}{2}} d^{\frac{3k^2+3k-4}{2}}.$$

We summarize the improved results:

Theorem 6.5 *For any fixed $N \geq 2$, there are positive constants c and c' such that*

$$\Pr\{Q_d(\epsilon)\} \leq c\left(\epsilon^2 d^3 + \epsilon^N d^{3N+1}\right) + 2^{-c'd}$$

if the a_i's are normally distributed, and

$$\Pr\{Q_d(\epsilon)\} \leq c\left(\epsilon^2 d^3 + \epsilon^N d^{3N+1}\right)$$

if the a_i's are uniform.

7 Consequences of the Erdös-Turán Estimate

In this section we give two types of improvements of the previous estimates using the following theorem of Erdös and Turán:

Lemma 7.1 *Let $N(\alpha,\beta)$ be the number of roots of $\sum_{i=0}^{d} a_i z^i = 0$ of the form $re^{i\theta}$ with r,θ real and $\alpha \le \theta \le \beta$ (and $\beta-\alpha \le 2\pi$). Then*

$$\left| N(\alpha,\beta) - \frac{(\beta-\alpha)d}{2\pi} \right| \le 16\sqrt{d\log\frac{\sum_{i=0}^{d}|a_i|}{|a_0||a_d|}}.$$

Proof See [ET50].

We will now argue assuming the a_i's are distributed uniformly in $B_1(0)$. The same estimates hold for the a_i's distributed normally, with minor modifications in the arguments. From the above it follows that with probability $\ge 1-\tau$ we have $|a_0|$ is $> c\sqrt{\tau}$ and thus

$$\left| N(\alpha,\beta) - \frac{(\beta-\alpha)d}{2\pi} \right| \le c\sqrt{d\log\frac{d}{\sqrt{\tau}}}.$$

This implies that the k-th closest root to a given root, z, has distance

$$\ge c'\frac{k - c\sqrt{d\log\frac{d}{\sqrt{\tau}}}}{d}|z|.$$

By lemma 4.1 we have that more than $d/2$ of the roots lying in $B_\rho(0)$ implies

$$|a_0| \le \rho^{d/2}$$

so that

$$\rho \ge \left(\frac{\tau}{\pi}\right)^{1/d}.$$

If follows that with probability $\ge 1-\tau-\tau_1-\tau_2$ we have a region of approxiamte zeros of size

$$\begin{aligned}
&\ge\ cd\left[\frac{k-1}{\delta_1} + \frac{\sqrt{d\log\frac{d}{\sqrt{\tau}}}}{\delta_2} + \sum_{k=1}^{d}\frac{d}{k}\left(\frac{\tau}{\pi}\right)^{-1/d}\right]^{-2} \\
&\ge\ cd\left[\min\left(\delta_1, \delta_2/\sqrt{d\log\frac{d}{\sqrt{\tau}}}, \tau^{1/d}/(d\log d)\right)\right]^2
\end{aligned}$$

where

$$\begin{aligned}\tau_1 &= c'\delta_1^4 d^5\\ \tau_2 &= c''\delta_2^{N-1} d^N, \qquad N = k^2+k-1.\end{aligned}$$

Setting

$$\epsilon = d\delta_1^2 = \delta_2^2/\log\frac{d}{\sqrt{\tau}}$$

we get

$$\tau_1 = c\epsilon^2 d^3, \qquad \tau_2 = c\left(\epsilon\log\frac{d}{\sqrt{\tau}}\right)^{\frac{N-1}{2}} d^N.$$

Choosing $\tau = \epsilon^2 d^2$ and assuming $\epsilon < 1/d^2$ we get that with probability

$$\geq 1 - c\left(\epsilon^2 d^3 + \left(\epsilon\log\frac{1}{\epsilon}\right)^{\frac{N-1}{2}} d^N\right)$$

we have an approximate zero region of area $\geq c'\epsilon$ since $\delta_1 = \sqrt{\epsilon/d}$ is smaller than $\tau^{1/d}/(d\log d) > c\epsilon^{2/d}/(d\log d)$.

Hence we can replace the

$$\epsilon^2 d^7$$

terms in equations 4.2 and 5.1 by

$$\epsilon^2 d^3 + \left(\epsilon\log\frac{1}{\epsilon}\right)^{\frac{N-1}{2}} d^N$$

for any fixed N.

Second, we claim that the $\epsilon^2 d^3$ term can be replaced by $(\epsilon^2 d^3)^M$ for any fixed integer M for all $\epsilon > e^{-cd}$, c depending on M. This is because lemma 7.1 gives that with probability $> 1 - e^{-cd}$ we have

$$N(0,\pi/M), N(2\pi/M, 3\pi/M), \ldots, N((2M-2)\pi/M, (2M-1)\pi/M)$$

each contain $c'd/M$ roots. For $i = 1, \ldots, M$,

$$z_1^i, \ldots, z_k^i \in N((2i-2)\pi/M, (2i-1)\pi/M) \cap (B_2(0) - B_{\tau^{1/d}}(0))$$

with $|z_j^i - z_1^i| \leq c|z_1^i|/d$ we have that

$$P(z_1^1, \ldots, z_k^1, z_1^2, \ldots, z_k^2, \ldots, z_1^M, \ldots, z_k^M) \approx \prod_{i=1}^{M} P(z_1^i, \ldots, z_k^i), \tag{7.1}$$

i.e. the events of finding roots at $z_1^i, \ldots, z_k^i$ for $i = 1, \ldots, M$ are approximately independent. This can be seen by considering for

$$v_j^i \equiv (0, \ldots, 0, j!, \ldots, d(d-1)\ldots(d-j+1)(z_1^i)^{d-j})$$

the matrix

$$\begin{pmatrix} \langle v_0^1, v_0^1 \rangle & \langle v_0^1, v_1^1 \rangle & \cdots & \left\langle v_0^1, v_{k-1}^M \right\rangle \\ \langle v_1^1, v_0^1 \rangle & \langle v_1^1, v_1^1 \rangle & \cdots & \left\langle v_1^1, v_{k-1}^M \right\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \left\langle v_{k-1}^M, v_0^1 \right\rangle & \left\langle v_{k-1}^M, v_1^1 \right\rangle & \cdots & \left\langle v_{k-1}^M, v_{k-1}^M \right\rangle \end{pmatrix}$$

and noting that its determinant is approximately

$$\prod_{i=1}^{M} \det \begin{pmatrix} \langle v_0^i, v_0^i \rangle & \cdots & \left\langle v_0^i, v_{k-1}^i \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle v_{k-1}^i, v_0^i \right\rangle & \cdots & \left\langle v_{k-1}^i, v_{k-1}^i \right\rangle \end{pmatrix} = cd^{Mk^2} \tag{7.2}$$

since for $i \neq j$ the term

$$\left\langle v_l^i, v_m^j \right\rangle = O\left(d^{l+m}\right)$$

instead of proportional to d^{l+m+1}. Hence any term in the expansion of the determinant involving at least one and therefore at least two such terms has size $O(1/d^2)$ times the term in equation 7.2.

From here on, calculations similar to those done previously yield equation 7.1 and thus the ability to replace $\epsilon^2 d^3$ by $(\epsilon^2 d^3)^M$.

A Complex Random Variables

In this appendix we give some basic facts about complex random variables and their Fourier transforms.

The *real isomorph* of a $m \times n$ complex matrix $M = U + iV$, with U, V real, is the real $2m \times 2n$ matrix given in block form as

$$\widehat{M} \text{ or } M\hat{\ } = \begin{pmatrix} U & -V \\ V & U \end{pmatrix}.$$

It is easy to see that $(M_1 + M_2)\hat{\ } = \widehat{M_1} + \widehat{M_2}$, $(M_1 M_2)\hat{\ } = \widehat{M_1}\widehat{M_2}$, $\widehat{M}^\top = \widehat{M^*}$ where $M^\top$ denotes the transpose of M and M^* denotes the complex conjugate transpose of M, and

$$\det \widehat{M} = |\det M|^2.$$

For $v = (v_1, \ldots, v_n) \in \mathbf{C}^n$, let

$$\tilde{v} = (\Re(v_1), \ldots, \Re(v_n), \Im(v_1), \ldots, \Im(v_n)) \in \mathbf{R}^{2n}.$$

One can check that $v^\top = Au^\top \Leftrightarrow \tilde{v}^\top = \hat{A}\tilde{u}^\top$, and that $\Re\langle u, v\rangle = (\tilde{u}, \tilde{v})$, where $\langle\ ,\ \rangle$ and $(\ ,\ )$ denote the usual inner products on $\mathbf{C}^n$ and $\mathbf{R}^{2n}$ respectively.

We say that u is a *normally distributed* complex random variable if $\Re(u)$ and $\Im(u)$ are independent, identically and normally distributed, real random variables. The *standard* complex normal u has distribution

$$\phi(z) = \frac{1}{2\pi} e^{-|z|^2/2}.$$

If $w_1, \ldots, w_m$ are independently, normally distributed real random variables with mean 0, and $v_1, \ldots, v_k$ are linear combinations of them, then the $v_1, \ldots, v_k$ have distribution $\phi \colon \mathbf{R^k} \to \mathbf{R}$

$$\phi(x) = \frac{1}{2\pi^{k/2}} \frac{1}{\sqrt{\det C}} e^{-(C^{-1}x, x)},$$

where C is the variance-covariance matrix for the v's. If $v_j = \sum a_{ij} w_i$, then

$$C = \begin{pmatrix} (a_1, a_1) & (a_1, a_2) & \cdots & (a_1, a_k) \\ (a_2, a_1) & (a_2, a_2) & \cdots & (a_2, a_k) \\ \vdots & \vdots & \ddots & \vdots \\ (a_k, a_1) & (a_k, a_2) & \cdots & (a_k, a_k) \end{pmatrix}$$

where $a_j = (a_{1j}, \ldots, a_{nj})$. Writing $w = Au$, we have $C = AA^\top$.

If $u = (u_1, \ldots, u_m)$ are standard complex normals, and $w = Au$, then one has $\tilde{v}^\top = \widehat{A}\tilde{u}^\top$. Thus $\tilde{v}^\top$ are real normals with distribution

$$\left(\frac{1}{2\pi}\right)^k \frac{1}{\sqrt{\det C}} e^{-(C^{-1}x,x)/2}.$$

with $C = \widehat{A}\widehat{A}^\top$. Hence $C = \widehat{B}$, where

$$B = AA^* = \begin{pmatrix} \langle a_1, a_1\rangle & \cdots & \langle a_1, a_k\rangle \\ \vdots & \ddots & \vdots \\ \langle a_k, a_1\rangle & \cdots & \langle a_k, a_k\rangle \end{pmatrix}$$

Also $C^{-1} = \widehat{B^{-1}}$, and so $(C^{-1}x, x) = \Re\langle B^{-1}z, z\rangle$, where $x = \tilde{z}$. Since $B^{-1} = (AA^*)^{-1} = (A^{-1})^* A^{-1}$, we have $\langle B^{-1}z, z\rangle = \langle A^{-1}z, A^{-1}z\rangle$, which is real, and so $\Re\langle B^{-1}z, z\rangle = \langle B^{-1}z, z\rangle$. Thus w has the distribution

$$\begin{aligned} \phi(z) &= \left(\frac{1}{2\pi}\right)^k \frac{1}{\sqrt{\det C}} e^{-(C^{-1}x,x)/2} \\ &= \left(\frac{1}{2\pi}\right)^k \frac{1}{|\det B|} e^{-\langle B^{-1}z,z\rangle/2}. \end{aligned}$$

In particular, v_1 and v_2 are independent iff $\langle a_1, a_2\rangle = 0$.

Next we recall some facts about the characteristic functions (Fourier transform) of complex random variables. For an $\mathbf{R}^n$ valued random variable, u, with density ϕ, its characteristic function $\widehat{\phi}\colon \mathbf{R}^n \to \mathbf{R}$ is

$$\hat{\phi}(\xi) \equiv E\left\{e^{i(\xi,u)}\right\} = \int_{\mathbf{R}^n} e^{i(\xi,x)}\phi(x)\,dx.$$

By the Fourier inversion formula,

$$\phi(x) = \left(\frac{1}{2\pi}\right)^n \int_{\mathbf{R}^n} e^{-i(\xi,x)}\hat{\phi}(\xi)\,d\xi.$$

For a complex valued random variable u with density $\phi\colon \mathbf{C} \to \mathbf{R}$, we can view u as two real random variables and define its characteristic function

$$\hat{\phi}(\xi_1, \xi_2) = E\left\{e^{i[\xi_1\Re(u)+\xi_2\Im(u)]}\right\} = \int_{\mathbf{C}} e^{i\Re\langle\xi,x\rangle}\phi(x)\,dx$$

where $\xi = \xi_1 + i\xi_2$.

If ϕ is radially symmetric, i.e. $\phi(z) = \tilde{\phi}(|z|)$, then we claim that $\hat{\phi}$ is radially symmetric. To see this, note

$$\begin{aligned}
\hat{\phi}(\xi) &= \int_{\mathbf{R}^2} \phi(x,y) e^{-i(x\xi_1 + y\xi_2)}\, dx\, dy \\
&= \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} \tilde{\phi}(r) e^{-ir(\xi_1 \cos\theta + \xi_2 \sin\theta)}\, d\theta\; r\, dr.
\end{aligned}$$

Writing $\xi = |\xi| e^{i\psi}$ and substituting we get

$$\hat{\phi}(\xi) = \iint \tilde{\phi}(r) e^{ir|\xi|\cos(\theta - \psi)}\, d\theta\; r\, dr$$

which is independent of ψ.

Assuming that the fourth moments of u are finite we have

$$\begin{aligned}
\hat{\phi}(\xi) &= \iint \tilde{\phi}(r) \left(1 + (ir|\xi|\cos\theta) + \frac{(ir|\xi|\cos\theta)^2}{2} + \cdots \right) d\theta\; r\, dr \\
&= 1 - m_1|\xi|^2 + O(|\xi|^4)
\end{aligned}$$

since the integrals involving odd powers of $\cos\theta$ vanish. Furthermore, if the $2\ell + 2$-th moments of u are finite, we have

$$\widehat{\phi}(\xi) = 1 - m_1|\xi|^2 - \cdots - m_\ell |\xi|^{2\ell} + O(|\xi|^{2\ell+2}).$$

Furthermore, for any $a \in \mathbf{C}$, it is easy to see that the characteristic function of ua is $\hat{\phi}(|a\xi|^2)$.

The characteristic function of the standard normal v, with density

$$\psi(z) = \frac{1}{2\pi} e^{-|z|^2/2}$$

is

$$\hat{\psi}(\xi) = e^{-|\xi|^2/2} = 1 - \frac{1}{2}|\xi|^2 + O\left(|\xi|^4\right)$$

for ξ small.

In §5 we will need to make estimates similar to those used in proving the central limit theorem. For these estimates, we recall the following facts. If u is a complex random variable with density ϕ then

$$|\hat{\phi}(\xi)| \leq 1\ \forall \xi$$

If ϕ is bounded, of bounded support, and, say, has $\phi(x, y)$ for fixed y of bounded total variation in x, then

$$|\hat{\phi}(\xi)| < \frac{a}{|\xi|} \ \forall \xi \text{ for some } a.$$

To see this, given ϕ is supported in $[-B, B]^2$ we estimate

$$\begin{aligned}
|\widehat{\phi}(\xi)| &= \left| \int_{\mathbf{R}^2} \phi(x,y) e^{ix|\xi|} \, dx \, dy \right| \\
&\leq 2B \max_y \left| \int_{x=-B}^{x=B} \phi(x,y) e^{ix|\xi|} \, dx \right| \\
&= 2B \max_y \left| \phi(x,y) \frac{e^{ix|\xi|}}{i|\xi|} \right|_{-B}^{B} - \int_{-B}^{B} \frac{\partial \phi}{\partial x}(x,y) \frac{e^{ix|\xi|}}{i|\xi|} \, dx \Bigg| \\
&\leq \frac{2B}{|\xi|} \left[\max_{x,y} |\phi(x,y)| + \int_{-B}^{B} \left| \frac{\partial \phi}{\partial x}(x,y) \right| dx \right] \\
&\leq \frac{c}{|\xi|} \left[\max_{x,y} |\phi(x,y)| + \text{T.V.}_{\cdot x} \phi(x,y) \right].
\end{aligned}$$

One can calculate the density of $u = u_1 + \cdots + u_n$ by taking the characteristic functions and using

$$\hat{\phi}(\xi) = E\left\{ e^{i\Re\langle \xi, u \rangle} \right\} = \prod_{j=1}^{n} E\left\{ e^{i\Re\langle \xi, u_i \rangle} \right\} = \prod_{j=1}^{n} \hat{\phi}_i(\xi).$$

Finally, the following lemma is useful

Lemma A.1 *If* $z_1, \ldots, z_n, {z'}_1, \ldots, {z'}_n \in B_1(0) \subset \mathbf{C}$, *then*

$$|z_1 \cdots z_n - {z'}_1 \cdots {z'}_n| \leq \sum_{i=1}^{n} |z_i - {z'}_i|.$$

Proof See [Bil79].

B Hammersley's Formula

In [H], Hammersley gives the formula

$$P(z_1, \dots, z_r) = \int_{f(z_1)=\dots=f(z_r)=0} \frac{|f'(z_1)|^2 \dots |f'(z_r)|^2}{\prod_{1 \le p < q < r} |z_p - z_q|^2} \omega(c)\, dc_r \dots dc_d \quad (2.1)$$

where $P(z_1, \dots, z_r)\, dz_1 \dots dz_r$ is the probability of finding a root in the region of volume $dz_1 \dots dz_r$ at $(z_1, \dots, z_r)$,

$$f(z) = \sum_{i=0}^{d} c_j z^j$$

and $\omega(c)$ is the density function of the coefficients. Note that this P differs from the P defined in §4 by a factor of

$$d(d-1)\dots(d-r+1)$$

since in §4 we first choose r roots at random to calculate P. Then he claims that

$$P(z_1) = \int_{\mathbf{C}} |t|^2 \psi(0, t)\, dt$$

where ψ is the joint density of f and f'. More generally we have

$$P(z_1, \dots, z_r) = E\left\{ |f'(z_1)|^2 \dots |f'(z_r)|^2 \text{ s.t. } f(z_1) = \dots = f(z_r) = 0 \right\}$$

where by the right hand side of the above we mean

$$\int_{\mathbf{C}^r} |t_1|^2 \dots |t_r|^2 \psi(0, \dots, 0, t_1, \dots, t_r)\, dt_1 \dots dt_r$$

where ψ is the joint density of $f(z_1), \dots, f(z_r), f'(z_1), \dots, f'(z_r)$. This can be derived from equation 2.1 by setting

$$s_i = \sum_{j=0}^{d} c_j z_i^j$$

$(= f(z_i))$ and writing

$$P(z_1, \dots, z_r) = \lim_{\epsilon \to 0} \left(\frac{1}{\pi \epsilon^2} \right)^r \int_{s_i \in B_\epsilon(0), i=1,\dots,r} \frac{|f'(z_1)|^2 \dots |f'(z_r)|^2}{\prod_{1 \le p < q < r} |z_p - z_q|^2} \omega(c)\, dc_r \dots dc_d\, ds_1 \dots ds_r$$

and noting that

$$\frac{\partial(s_1,\ldots,s_r,c_r,\ldots c_d)}{\partial(c_0,\ldots,c_d)} = \frac{\partial(s_1,\ldots,s_r)}{\partial(c_0,\ldots,c_{r-1})} = \prod_{i<j<r}(z_j - z_i)$$

and so

$$\begin{aligned} P(z_1,\ldots,z_r) &= \lim_{\epsilon\to 0}\left(\frac{1}{\pi\epsilon^2}\right)^r \int_{s_i\in B_\epsilon(0), i=1,\ldots,r} |f'(z_1)|^2\ldots|f'(z_r)|^2\omega(c)\,dc_0\ldots dc_d \\ &= \int_{\mathbf{C}^r} |t_1|^2\ldots|t_r|^2\psi(0,\ldots,0,t_1,\ldots,t_r)\,dt_1\ldots dt_r \end{aligned}$$

where ψ is the joint density of $f(z_1),\ldots,f(z_r),f'(z_1),\ldots,f'(z_r)$.

References

[Bil79] Patrick Billingsley. *Probability and Measure*. John Wiley & Sons, 1979.

[BS86] A.T. Bharucha-Reid and M. Sambandham. *Random Polynomials. Probability and Mathematical Statistics*, Academic Press, New York, 1986.

[ET50] P. Erdös and P. Turán. On the distribution of roots of polynomials. *Annals of Math.*, 51(1):105–119, 1950.

[GT74] W. Greg and R. Tapia. Optimal error bounds for the newton-kantorovich theorem. *SIAM J. Numer. Anal.*, 11:10–13, 1974.

[Ham60] J.M. Hammersley. The zeros of a random polynomial. In *The Third Berkeley Symp. ...*, pages 89–111, 1960.

[KA70] L. Kantorovich and G. Akilov. *Functional Analysis in Normed Spaces*. MacMillan, 1970.

[Kan52] L. Kantorovich. *Functional Analysis and Applied Mathematics*. National Bureau of Standards, 1952.

[Kim85] M. Kim. *Ph.D. Thesis*. PhD thesis, City University of New York, 1985.

[Mar66] M. Marden. *Geometry of Polynomials.* Amer. Math. Soc., 1966.

[Ren87] James Renegar. On the worst-case arithmetic complexity of approximating zeros of polynomials. 1987. preprint.

[Roy86] H.L. Royden. Newton's method. *Preprint*, 1986.

[Sma81] Steve Smale. The fundamental theorem of algebra and complexity theory. *Bull. of the AMS*, 4(1):1–35, 1981.

[Sma86a] Steve Smale. Algorithms for solving equations. In *written for the International Congress of Mathematicians*, 1986.

[Sma86b] Steve Smale. Newton's method estimates from data at one point. In R.E. Ewing, K.I. Gross, and C.F. Martin, editors, *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics*, pages 185–196, Springer-Verlag, 1986.

[SS85] Mike Shub and Steve Smale. Computational complexity: on the geometry of polynomials and a theory of cost: part i. *Ann. Scient. École Norm. Sup. 4 serie t*, 18:107–142, 1985.

[SS86] Mike Shub and Steve Smale. Computational complexity: on the geometry of polynomials and a theory of cost: part ii. *SIAM J. on Computing*, 15(1):145–161, 1986.