THE ZEROS OF A RANDOM POLYNOMIAL

J. M. HAMMERSLEY

OXFORD UNIVERSITY

1. Introduction

If the coefficients of an algebraic equation are subject to random error, the roots of this equation will also be subject to random error; and it is natural to enquire how the latter errors depend upon the former. This is clearly a question of some practical importance. It arises not only when the coefficients result from experimental data, but also, for example, when the coefficients are rounded off to some specified number of decimal places before commencing a numerical solution. Yet it is a question which has so far received rather scant attention, apart from the treatment of three special instances.

In the first of these special instances the equation is of a particular type, namely the characteristic equation of a variance-covariance matrix pencil whose elements are real and distributed in Wishart's form. Under these circumstances the roots are all real, and their joint sampling distribution is well known. Amongst the several textbooks, which discuss this question, the reader may consult Wilks [19].

In the second special instance the equation is linear and the coefficients are real and distributed normally (though not necessarily independently). The distribution of the root of this equation is therefore that of the quotient of two real correlated normal variates. Geary [9] gives the required result. This special instance arises in bio-assay work under the name of Fieller's theorem (see pp. 27–29 in Finney [8]). The interpretation of this theorem is, however, open to question, and I discuss this matter further in section 9.

In the third special instance the equation is of general degree and its coefficients are real and distributed independently and symmetrically about zero either

  (i) normally, or
  (ii) rectangularly, or
  (iii) discretely into the pair of classes ± 1.

Littlewood and Offord [14], [15] enquired how many real roots on the average such a random equation might be expected to have, and they gave asymptotic approximations for the result valid for equations of large degree. Kac [12], [13] improved their results by showing that the average number of real roots of an equation of degree $n - 1$ is

$$(1.1) \qquad N_n = \frac{4}{\pi} \int_0^1 \frac{(1 - h_n^2)^{1/2}}{1 - x^2} \, dx \leq \frac{2}{\pi} \log n + \frac{14}{\pi},$$

where

$$(1.2) \qquad h_n = n x^{n-1} \frac{1 - x^2}{1 - x^{2n}}$$

when (i) applies; and for both (i) and (ii) $N_n \sim (2/\pi) \log n$ as $n \to \infty$. For the latest results, see Erdös and Offord [20]. Littlewood and Offord noticed that the three cases (i), (ii), and (iii) led to results which were (as far as they could tell from their approximations) all of the same order of magnitude for large n. This is not surprising in view of the result obtained in section 6 that the average number of real roots depends only on the distribution of the real and imaginary parts of the polynomial and its derivatives (regarding its argument as fixed); for by the central limit theorem these four parts will (under wide conditions) be asymptotically normal. However, as Kac [13] discovered, it is not easy to put this reasoning into rigorous terms; and we shall not attempt it here.

Some quotations from the papers of Littlewood and Offord and of Kac indicate certain differences of approach in the present article. From [15]: "In 'selecting' a random equation of a particular given degree n (for example, $n = 10$) it is natural to treat each coefficient on the same basis." This is a condensation of a longer passage in [14], where two players are betting on the number of real roots of an equation selected by a referee, who "would presumably treat each coefficient on the same basis." Similarly from [13]: ". . . let [the coefficients] be independent random variables each having $\sigma(u)$ as its distribution function." It is, of course, reasonable to treat the coefficients as independent and having the same distribution function when the problem is posed as one of pure mathematics; and, even when entering upon a wider sphere than that of pure mathematics, it is doubtless wise to regard the simple case of similarly distributed coefficients as a starting point in one's investigations. But in a problem of applied mathematics it is far from likely that the coefficients will be independent or have similar distributions. My own interest in the problem sprang from a question put to me concerning the precision of an estimate of the growth rate of an insect population, which growth rate is the root of an algebraic equation whose coefficients are determined from experimental data in a manner that certainly vitiates the assumption of independence. So in the present article I allow the coefficients to be distributed in a quite general manner. Although I have been able to determine a formal expression for the distribution of the roots in this general case, I am still very far from having solved the practical problem; for my result involves the parameters of the distribution function of the coefficients, and except in a trivial special case I cannot at present see how to "Studentize" the result in such a way that these parameters may be replaced by their estimates. Concerning this problem of "Studentization," the reader may care to consider the papers by Creasy [4] and Fieller [7].

Another result of attacking the problem from the point of view of generally distributed coefficients was to throw light upon certain essential points of difficulty in the problem. In [13] Kac writes: "Upon closer examination it turns out that the proof I had in mind [in [12]] . . . is inapplicable to the [case when the coefficients have a discrete distribution]. . . . This situation tends to emphasize the particular interest of the discrete case, which surprisingly enough turns out to be the most difficult." Something more than this is in fact surprising: namely, that the simplest case arises when the coefficients are allowed to assume complex values, the joint distribution of all real and imaginary parts having a continuous frequency function with finite moments. When any discontinuities are introduced into the distribution function of the coefficients, the corresponding distribution of the roots is in danger of acquiring discontinuities as well. Thus when all real and imaginary parts are distributed normally with strictly positive variances (that is to say in nondegenerate multivariate Gaussian form) the distribution of the roots pos-

sesses a frequency function over the complex plane; but when the coefficients are made real (so that the distribution function of the imaginary part contains a step function) the distribution function of the roots, instead of being a well-behaved surface distribution in the complex plane, partly gathers itself up into a line distribution on the real axis, so that a frequency function no longer exists over the whole complex plane. The results I have just quoted upon normally distributed coefficients will be discussed in sections 8 and 9. Also in the case of normally distributed coefficients, I shall fulfill the requirement in [12]: ". . . the problem of the exact determination of the average distribution of real roots on the real axis will, of course, depend on a more delicate treatment."

In all these three special instances, cited above, the preoccupation is with real roots. As soon as we attempt to tackle the more general enquiry, which embraces complex roots as well, we come up against a difficulty. In order to give the joint distribution of a set of roots it is necessary to keep track of each individual root. When the roots are real we may do this (as is done in the first special instance) by arranging them in order of magnitude. However, while it is possible to arrange complex numbers in an order of magnitude by means of appropriate conventions, it seems difficult to arrange them in an order which is continuous with respect to the geometry of the complex plane. It is conceivable that, by some device or other, one might manage to label individual complex roots in such a manner that their resulting joint distribution possesses a coherent and useful interpretation; but I have not succeeded in doing this, and I am indeed doubtful whether it is possible. Alternatively, one may abandon the conventional idea of a joint distribution. In this article, I have replaced it by a somewhat similar concept, which I shall call a condensed distribution. We shall see in due course that this concept leads to a particularly simple solution of the problem.

Another branch of investigation, connected with the present problem, is followed by Erdös and Turán [5]. They consider the uniformity of the distribution of the roots; they prove that, if $N(\alpha, \beta)$ denotes the number of zeros of $g(z) = \sum_{j=0}^{n} c_j z^j = 0$ which lie in the sector $0 \leq \alpha < \arg z < \beta \leq 2\pi$, then

$$(1.3) \qquad \left| N(\alpha, \beta) - \frac{(\beta - \alpha) n}{2\pi} \right| < 16 \left(n \log \frac{\sum_{j=0}^{n} |c_j|}{\sqrt{|c_0 c_n|}} \right)^{1/2} .$$

2. Notation

Capital German letters will denote finite-dimensional Euclidean spaces. Small German letters and small boldface letters will respectively denote subsets and points (column or row vectors) of these spaces. Capital boldface letters will denote matrices. Italic letters will denote scalars (real or complex according to the context). It will be convenient sometimes to regard a complex scalar as a two-dimensional vector; and we utilize the foregoing conventions by writing $z = x + iy$ or $z = \{x, y\}$ according to requirements. A dagger attached to a boldface letter will denote the transposed matrix or vector. A prime attached to a German letter will denote the complementary set (with respect to the Euclidean space containing the set in question), while a prime attached to an italic or boldface letter will denote a derivative. $|x|$ will denote the length of a vector x, while $|X|$ will denote the determinant of a (square) matrix X. An asterisk attached to a letter will indicate that the quantity in question is a random variate.

Sets of points will always be Borel sets, either by hypothesis or as the result of proofs. Except when the contrary is explicitly stated, functions will always be single valued. All many-valued functions will be finitely many valued. All functions will be Borel measurable (except when explicitly stated) either by hypothesis or as the result of proofs.

3. The extended Slutzky-Fréchet theorem

The extended Slutzky-Fréchet theorem is the principal tool of this paper. Before enunciating it, we recall some familiar properties of one-valued random variates.

A *random variate* x^* in a space $\mathfrak{X}$ is defined by a real nonnegative countably additive set function

$$(3.1) \qquad\qquad F[\mathfrak{x}] = Pr\{x^* \in \mathfrak{x}\}$$

defined for all Borel sets $\mathfrak{x}$ of $\mathfrak{X}$ and satisfying $F[\mathfrak{X}] = 1$. $F[\mathfrak{x}]$ is called the *probability set function* of x^*. In the particular case when $\mathfrak{x}$ consists of all points whose coordinates do not exceed the corresponding coordinates of a prescribed point x, we write $F[\mathfrak{x}] = F(x)$, and call $F(x)$ the *cumulative distribution function* of x^*. Evidently $F[\mathfrak{x}]$ uniquely determines $F(x)$, and the converse is a consequence of Lebesgue's theory of integration. Thus we can uniquely specify a random variate by means of its cumulative distribution function. If x_ν^*, $\nu = 1, 2, \cdots$, is a sequence of random variates in $\mathfrak{X}$ specified by the respective cumulative distribution functions $F_\nu(x)$, and if $\lim_{\nu \to \infty} F_\nu(x) = F(x)$ except perhaps at the hyperplanes of discontinuity of the cumulative distribution function $F(x)$, we say that x_ν^* *converges in distribution* to x^* and write $\operatorname{dlim}_{\nu \to \infty} x_\nu^* = x^*$. If $y = y(x)$ is a (one-valued Borel-measurable) function carrying points x of $\mathfrak{X}$ into points y of another space $\mathfrak{Y}$, the *function of a random variate* $y(x^*)$ is defined to be the random variate specified by $F[y^{-1}(\mathfrak{y})]$, where $y^{-1}(\mathfrak{y})$ is the set of all x satisfying $y(x) \in \mathfrak{y}$. Cramér [3] gives a proof that this is a consistent definition, that is, that $F[y^{-1}(\mathfrak{y})]$ is indeed a probability set function on $\mathfrak{Y}$).

We now consider the terminology for many-valued functions of a random variate which will be required in the extended Slutzky-Fréchet theorem.

An n-valued function $y(x)$ is a set of n points (not necessarily distinct) in $\mathfrak{Y}$ corresponding to each given point x of $\mathfrak{X}$. An *indexing* of $y(x)$ is a system of n one-valued functions $y_j(x), j = 1, 2, \cdots, n$, such that, for every given x of $\mathfrak{X}$, the n points $y(x)$ coincide with the n points $y_j(x)$ with due regard to multiplicity. If there exists at least one such indexing in which each $y_j(x)$ is a Borel-measurable function, then $y(x)$ is called an *n-valued Borel-measurable function*, and the indexing in question is called a *Borel-measurable indexing*. Notice that an n-valued Borel-measurable function may possess, besides this Borel-measurable indexing, other indexings which are not Borel measurable. The n-valued function $y(x)$ is said to be *continuous at the point* x_0 if, given any prescribed $\epsilon > 0$, we can find

   (i) an indexing $y_j(x)$ of $y(x)$,
   (ii) a number $\eta = \eta(\epsilon, x_0) > 0$, and
   (iii) a permutation $\pi_1, \pi_2, \cdots, \pi_n$ of the integers $1, 2, \cdots, n$
such that

$$(3.2) \qquad\qquad |y_j(x_0) - y_{\pi_j}(x)| < \epsilon$$

simultaneously for $j = 1, 2, \cdots, n$, whenever x satisfies $|x - x_0| < \eta$. The permutation in (iii) is allowed to depend on ϵ, x_0, and x. Then n-valued function $y(x)$ is said to be *almost certainly continuous with respect to the one-valued random variate* x^* if there exists a Borel set $\mathfrak{x}_0$ satisfying $F[\mathfrak{x}_0] = 1$, where F is the probability set function of x^*, such that $y(x)$ is continuous at each $x_0 \in \mathfrak{x}_0$. If $y_j(x)$, $j = 1, 2, \cdots, n$, is a Borel-measurable indexing of an n-valued Borel-measurable function, and if x^* is a one-valued random variate with a probability set function $F[\mathfrak{x}]$, the one-valued random variate in $\mathfrak{Y}$ speci-

fied by the probability set function $n^{-1} \sum\limits_{j=1}^{n} F[y_j^{-1}(\mathfrak{y})]$ is denoted by $\bar{y}(x^*)$ and called the

condensation of the many-valued random variate $y(x^*)$. Evidently this definition is independent of the particular Borel-measurable indexing used therein.

We can now state the extended Slutzky-Fréchet theorem.

THEOREM 3.1. *If* $y(x)$ *is a given n-valued Borel-measurable function which is almost certainly continuous with respect to a given random variate* x^*, *then* $\operatorname*{dlim}\limits_{\nu \to \infty} x_\nu^* = x^*$ *implies* $\operatorname*{dlim}\limits_{\nu \to \infty} \bar{y}(x_\nu^*) = \bar{y}(x^*)$.

See Hammersley [10] for a proof of this theorem.[1]

The motive for condensing a many-valued random variate is to overlook its indexing, since the indexing is an extraneous artifice for handling the different values. In some respects, however, condensation is too drastic a way of disregarding the indexing. To obtain a less extreme concept consider any given Borel set $\mathfrak{y}$ in $\mathfrak{Y}$. For an indexing $y_j(x)$, let $n^* = n(x^*)$ denote the number of values of j such that $y_j(x^*) \in \mathfrak{y}$. Clearly n^* is independent of the particular indexing used. The *moment-generating set function* is defined to be $M(t, \mathfrak{y}) = E[\exp (tn^*)]$. The greater flexibility of this concept is paid for by increased mathematical difficulties; for, although $M(t, \mathfrak{y})$ is a set function of $\mathfrak{y}$, it is not an additive set function. The mathematical theory of nonadditive set functions is not well explored.

The equation

$$(3.3) \qquad M(t, \mathfrak{y}) = \left[\prod_{j=1}^{n} \left(\int_{y_j \in \mathfrak{y}'} + e^t \int_{y_j \in \mathfrak{y}} \right) \right] dG(y_1, y_2, \cdots, y_n)$$

provides a formal expression for M. Here G is the joint cumulative distribution function

---

[1] I take this opportunity of correcting a flaw in the proof of a preliminary lemma (theorem 1 of [10]). The flaw is concealed by the notation used for the joint determination G, which in fact depends on ν, δ, and ϵ. Instead of the final relation

(i) $\qquad\qquad\qquad G[\mathfrak{z}_0] > 1 - \epsilon, \quad \nu \geqq \nu_0 (\delta, \epsilon),$

it would have been better to write

(ii) $\qquad\qquad\qquad G[\mathfrak{z}_0 | \nu, \delta, \epsilon] > 1 - \epsilon, \quad \nu \geqq \nu_0(\delta, \epsilon).$

What had to be proved was the existence of some G, depending upon ν but independent of δ and ϵ, such that

(iii) $\qquad\qquad\qquad G[\mathfrak{z}_0 | \nu] > 1 - \epsilon, \quad \nu \geqq \nu_0(\delta, \epsilon).$

The original proof established (ii) but not (iii). However a *reductio ad absurdum* argument suffices to yield (iii). For suppose the desired result false. Then there exist fixed strictly positive δ, ϵ such that (iii) is false for infinitely many ν for *every* joint determination G. Since δ, ϵ are fixed, (ii) provides the necessary contradiction.

for the indexing $y_1(x^*)$, $y_2(x^*)$, $\cdots$, $y_n(x^*)$; and the product of the integral signs is to be expanded formally before integration. Thus for $n = 2$

$$(3.4) \quad M(t, \mathfrak{y}) = \iint_{\substack{y_1 \in \mathfrak{y}' \\ y_2 \in \mathfrak{y}'}} dG(y_1, y_2) + e^t \iint_{\substack{y_1 \in \mathfrak{y}' \\ y_2 \in \mathfrak{y}}} dG(y_1, y_2)$$

$$+ e^t \iint_{\substack{y_1 \in \mathfrak{y} \\ y_2 \in \mathfrak{y}'}} dG(y_1, y_2) + e^{2t} \iint_{\substack{y_1 \in \mathfrak{y} \\ y_2 \in \mathfrak{y}}} dG(y_1, y_2).$$

Indeed, the justification of (3.3) is apparent as soon as the formal expansion is made. Theorem 3.1 is a special case of

THEOREM 3.2. *If $y(x)$ is a given n-valued Borel-measurable function which is almost certainly continuous with respect to a random variate x^*, if $\operatorname*{dlim}_{\nu \to \infty} x_\nu^* = x^*$, and if $M(t, \mathfrak{y})$ and $M_\nu(t, \mathfrak{y})$ are the moment-generating set functions of $y(x^*)$ and $y(x_\nu^*)$ respectively, then, for every open set $\mathfrak{y}$, $\lim_{\nu \to \infty} M_\nu(t, \mathfrak{y}) = M(t, \mathfrak{y})$.*

This theorem follows without difficulty from the result established by Hammersley (see 1. 10, p. 256 in [10]). The requirement that $\mathfrak{y}$ should be an open set can be relaxed somewhat; roughly speaking, all that is needed is that $M(t, \mathfrak{y})$ should not be a discontinuous function of $\mathfrak{y}$ at $\mathfrak{y}$, but a precise formulation of this idea is rather cumbrous.

Evidently the probability set function of $\bar{y}(x^*)$ is the coefficient of t in the expansion of $M(t, \mathfrak{y})/n$ in powers of t.

4. Borel measurability and continuity of the zeros of a polynomial

In this section we prove that the zeros of a polynomial are Borel measurable and (with a single exception) continuous functions of the coefficients of the polynomial.

LEMMA 4.1. *There exists a fixed indexing $z_1, z_2, \cdots, z_m$ of any given fixed set of m distinct points in the complex plane such that*

$$(4.1) \qquad |z_1 - e^{\pi i/s}| > |z_2 - e^{\pi i/s}| > \cdots > |z_m - e^{\pi i/s}| > 0$$

for all sufficiently large positive integers s, say $s \geqq S_0(z_1, z_2, \cdots, z_m)$.

In enunciating this lemma we make no attempt to claim that S_0 is a Borel-measurable function of $z_1, z_2, \cdots, z_m$. The reason for this remark will appear during the proof of theorem 4.1 below.

Lemma 4.1 effectively states that an ordering can be set up for the m points, and it is therefore enough to consider the case $m = 2$. The result is obvious when z_1 and z_2 are at different distances from $z = 1$. When $|z_1 - 1| = |z_2 - 1|$, z_1 is the name given to the point with the smaller imaginary part, except that, when the two imaginary parts are equal, z_1 is the name of the point with the larger real part.

THEOREM 4.1. *There exists an indexing $z_1, z_2, \cdots, z_n$ of the zeros of the polynomial*

$$\sum_{j=0}^{n} c_j z^j,$$ *where $c_j = a_j + ib_j$, such that, for each $k = 1, 2, \cdots, n$, $z_k = \{x_k, y_k\}$ is a one-valued Borel-measurable function of $c = \{a_0, a_1, \cdots, a_n, b_0, b_1, \cdots, b_n\}$. Moreover, for this indexing $|z_k - 1|$ is a nonincreasing function of k.*

First and until further notice assume that c is fixed and $c_n = 1$. Then the zeros of $\sum_{j=0}^{n} c_j z^j$ are fixed and all finite. Suppose that there are m distinct zeros amongst them,

and use the indexing of lemma 4.1 to name these distinct zeros $z_1, z_2, \cdots, z_m$. Since $z_1, z_2,$ $\cdots, z_m$ are uniquely determined by c, the function S_0 of lemma 4.1 is a function of c, say $S_0(z_1, z_2, \cdots, z_m) = S(c)$. We have not yet proved that $z_1, z_2, \cdots, z_m$ are Borel-measurable functions of c, so that we cannot and do not assert that $S(c)$ is a Borel-measurable function of c. This explains why we did not bother to prove that S_0 was a Borel-measurable function of $z_1, z_2, \cdots, z_m$.

For each integer $s = 1, 2, \cdots$ define numbers C_{js} by means of the identity in z

$$(4.2) \qquad \sum_{j=0}^{n} c_j (z + e^{\pi i/s})^j = \sum_{j=0}^{n} C_{js} z^j.$$

For $t = 1, 2, \cdots, n$ let u_t be an arbitrary real number satisfying $0 \leq u_t \leq 1$. For $s, t = 1, 2, \cdots$ define Z_{st} as a function of $u = \{u_1, u_2, \cdots, u_n\}$ by means of the recurrence relations

$$(4.3) \qquad Z_{st} = \begin{cases} u_t & \text{for } t \leq n, \\ -\sum_{j=0}^{n-1} C_{js} Z_{s, \, t-n+j} & \text{for } t > n. \end{cases}$$

Then define

$$(4.4) \qquad U_{st} = \begin{cases} \dfrac{Z_{s, \, t+1}}{Z_{st}} & \text{for } Z_{st} \neq 0, \\ 0 & \text{for } Z_{st} = 0. \end{cases}$$

Finally write

$$(4.5) \qquad f(c) = \lim_{s \to \infty} \int_0^1 \int_0^1 \cdots \int_0^1 \lim_{t \to \infty} (e^{\pi i/s} + U_{st}) \, du_1 du_2 \cdots du_n.$$

We shall prove that $f(c)$ exists, is a Borel-measurable function of c subject to $c_n = 1$, and equals z_1.

By (4.2), $c_n = 1$ implies $C_{ns} = 1$. Hence the linear difference equation (4.3) has the solution

$$(4.6) \qquad Z_{st} = \sum_{k=1}^{m} P_k(t) (z_k - e^{\pi i/s})^t, \qquad\qquad t = 1, 2, \cdots,$$

where $P_k(t)$ is a polynomial in t whose coefficients are (not necessarily Borel measurable) functions of u, for the distinct roots of the right-hand side of (4.2) are $z = z_k - \exp(\pi i/s)$, $k = 1, 2, \cdots, m$. The polynomials $P_k(t)$ have degree less than n and are all identically zero if and only if $u = 0$. When $u \neq 0$, let K be the smallest value of k such that $P_k(t)$ is not identically zero. Consider any fixed value of $s \geq S(c)$. By lemma 4.1

$$(4.7) \qquad |z_K - e^{\pi i/s}| > |z_{K+1} - e^{\pi i/s}| > \cdots > |z_m - e^{\pi i/s}| > 0$$

Hence, from (4.6),

$$(4.8) \qquad \lim_{t \to \infty} U_{st} = \begin{cases} z_K - e^{\pi i/s} & \text{for } u \neq 0, \\ 0 & \text{for } u = 0. \end{cases}$$

Now K is a function of u, but we show that $K = 1$ for almost all u. From (4.3) and (4.6)

$$(4.9) \qquad u_t = \sum_{k=1}^{m} P_k(t) (z_k - e^{\pi i/s})^t, \qquad\qquad t = 1, 2, \cdots, n.$$

There are altogether n coefficients in all the polynomials P_k taken together. If any one of these coefficients is zero, (4.9) shows that there is a linear relation between $u_1, u_2, \cdots, u_n$; that is, u lies on a hyperplane. Hence, except perhaps when u lies on one or more of n hyperplanes in the unit hypercube, none of the polynomials P_k is zero. Hence, from (4.8),

$$(4.10) \qquad \int_0^1 \int_0^1 \cdots \int_0^1 \lim_{t \to \infty} (U_{st} + e^{\pi i/s}) \, du_1 du_2 \cdots du_n = z_1, \qquad s \geqq S(c) .$$

By (4.5), $f(c)$ exists and $f(c) = z_1$ for each fixed c subject to $c_n = 1$. It remains to prove that $f(c)$ is a Borel-measurable function of c. We therefore allow c to vary subject to the restriction $c_n = 1$. Evidently, by (4.2), the C_{js} are Borel-measurable functions of c and s; hence, by (4.3) and (4.4), U_{st} is a Borel-measurable function of c, s, and u. The result follows from (4.5), since the limit and the integral of Borel-measurable functions are Borel measurable.

Next we remove the restriction $c_n = 1$. If $c_n \neq 0$, we may divide all the coefficients of $\sum_{j=0}^{n} c_j z^j$ by c_n without affecting z_1. If $c_n = 0$, we define $z_1 = \{+\infty, +\infty\}$. It follows that $z_1 = \{x_1, y_1\}$ is a Borel-measurable function of c without restriction upon c.

For the next zero, we repeat the foregoing process for the polynomial $\sum_{j=0}^{n-1} \gamma_j z^j$, whose zeros are those of $\sum_{j=0}^{n} c_j z^j$ excepting one zero at $z = z_1$. The γ_j are defined by

$$(4.11) \qquad \gamma_j = \begin{cases} 0 & \text{for } j = n , \\ c_{j+1} - z_1 \gamma_{j+1} & \text{for } j = n-1, n-2, \cdots, 1, 0 , \end{cases}$$

provided $c_n \neq 0$, and by $\gamma_j = c_j, j = 0, 1, \cdots, n-1$, if $c_n = 0$. Since a Borel-measurable function of a Borel-measurable function is Borel measurable, the γ_j are Borel-measurable functions of c. Hence the next zero of $\sum_{j=0}^{n} c_j z^j$ is a Borel-measurable function of c.

Further repetition of this process provides the required indexing and shows that the z_k are Borel-measurable functions of c; the method of procedure allied with lemma 4.1 shows that $|z_k - 1|$ is a nonincreasing function of k.

LEMMA 4.2. *Let n be a given integer greater than zero, and let R and ϵ_1 be given real numbers satisfying $0 < \epsilon_1 \leqq 2 \leqq R < \infty$. Define*

$$(4.12) \qquad \epsilon_{j+1} = \left(\frac{\epsilon_j}{(2R)^i} \right)^i , \qquad i = 1, 2, \cdots, n .$$

Let

$$(4.13) \qquad P_n(z) \equiv \sum_{j=0}^{n} p_{nj} z^j = 0 , \qquad Q_n(z) \equiv \sum_{j=0}^{n} q_{nj} z^j = 0 , \qquad p_{nn} = q_{nn} = 1 ,$$

be two algebraic equations (with complex coefficients) having no roots outside the circle $|z| = R$. Then, if

$$(4.14) \qquad |p_{nj} - q_{nj}| \leqq \epsilon_{n+1} , \qquad j = 0, 1, \cdots, n-1$$

it is possible to arrange the roots α of $P_n(z) = 0$ and the roots β of $Q_n(z) = 0$ into some set of n pairs (α_j, β_j) such that

$$(4.15) \qquad\qquad |\alpha_j - \beta_j| < \epsilon_j, \qquad\qquad j = 1, 2, \cdots, n.$$

Since $\frac{1}{2} \geq \epsilon_1/2R = \epsilon_2 > 0$ it is clear from (4.12) that $\frac{1}{2} \geq \epsilon_2 > \epsilon_3 > \cdots > \epsilon_{n+1} > 0$; hence

$$(4.16) \qquad\qquad 2\epsilon_{j+1} \leq 2^{j-1}(2\epsilon_{j+1})^{1/j}, \qquad\qquad j = 1, 2, \cdots, n.$$

If λ is any root of $P_n(z) = 0$ or of $Q_n(z) = 0$ we have $|\lambda| \leq R$, and hence

$$(4.17) \qquad \sum_{k=0}^{j} |\lambda^k| \leq \sum_{k=0}^{j} R^k = \frac{R^j(1-R^{-j-1})}{1-R^{-1}} < 2R^j$$

since $R \geq 2$.

Let β_n be any specified root of $Q_n(z) = 0$. Then

$$(4.18) \quad \left| \prod_{j=1}^{n} (\beta_n - \alpha_j) \right| = |P_n(\beta_n)| = |P_n(\beta_n) - Q_n(\beta_n)| = \left| \sum_{j=0}^{n} (p_{nj} - q_{nj}) \beta_n^j \right|$$

$$\leq \epsilon_{n+1} \sum_{j=0}^{n} |\beta_n^j| < 2\epsilon_{n+1}R^n$$

by virtue of (4.14) and (4.17). It follows that there is at least one root of $P_n(z) = 0$, which we denote by α_n, such that

$$(4.19) \qquad\qquad |\beta_n - \alpha_n| < (2\epsilon_{n+1})^{1/n}R.$$

Now define the polynomials $P_{n-1}(z)$ and $Q_{n-1}(z)$ by the identities

$$P_n(z) \equiv (z - \alpha_n)P_{n-1}(z) \equiv (z - \alpha_n) \sum_{j=0}^{n-1} p_{n-1, j} z^j, \qquad p_{n-1, n-1} = 1,$$

$$(4.20)$$

$$Q_n(z) \equiv (z - \beta_n)Q_{n-1}(z) \equiv (z - \beta_n) \sum_{j=0}^{n-1} q_{n-1, j} z^j, \qquad q_{n-1, n-1} = 1.$$

The roots of $P_{n-1}(z) = 0$ and $Q_{n-1}(z) = 0$, being a subset of the roots of $P_n(z) = 0$ and $Q_n(z) = 0$, do not lie outside the circle $|z| = R$. Upon identification of the coefficients in (4.20),

$$(4.21) \quad (p_{n-1, j} - q_{n-1, j}) = (p_{n, j+1} - q_{n, j+1}) + \alpha_n(p_{n-1, j+1} - q_{n-1, j+1})$$

$$+ (\alpha_n - \beta_n) q_{n-1, j+1}.$$

Now $\pm q_{n-1, j+1}$ is the sum of the products of the roots of $Q_{n-1}(z) = 0$ taken $n - j - 2$ at a time, so

$$(4.22) \qquad\qquad |q_{n-1, j+1}| \leq \binom{n-1}{j+1} R^{n-j-2}.$$

Take the modulus of (4.21) and employ (4.14), (4.19), and (4.22) to yield

$$(4.23) \quad |p_{n-1, j} - q_{n-1, j}| < \epsilon_{n+1} + R|p_{n-1, j+1} - q_{n-1, j+1}|$$

$$+ (2\epsilon_{n+1})^{1/n}R \binom{n-1}{j+1} R^{n-j-2}.$$

In this inequality write $j + k$ for j, multiply through by R^k, and sum over $0 \leq k \leq n - j - 2$. This yields

$$(4.24) \quad \sum_{k=0}^{n-j-2} R^k |p_{n-1,\,j+k} - q_{n-1,\,j+k}| < \sum_{k=0}^{n-j-3} R^{k+1} |p_{n-1,\,j+k+1} - q_{n-1,\,j+k+1}|$$

$$+ \sum_{k=0}^{n-j-2} \left\{ \epsilon_{n+1} R^k + \binom{n-1}{j+k+1} (2\epsilon_{n+1})^{1/n} R^{n-j-1} \right\},$$

the first sum on the right-hand side having the upper limit $n - j - 3$ instead of $n - j - 2$ since $p_{n-1,\,n-1} = q_{n-1,\,n-1} = 1$. Cancelling like terms from each side of the inequality, we find

$$(4.25) \quad |p_{n-1,\,j} - q_{n-1,\,j}| < \sum_{k=0}^{n-j-2} \left\{ \epsilon_{n+1} R^k + \left(\frac{n-1}{j+k+1} \right) (2\epsilon_{n+1})^{1/n} R^{n-j-1} \right\}$$

$$< \epsilon_{n+1} \sum_{k=0}^{n-2} R^k + (2\epsilon_{n+1})^{1/n} R^{n-1} \sum_{k=0}^{n-1} \binom{n-1}{k}$$

$$< 2\epsilon_{n+1} R^{n-2} + (2\epsilon_{n+1})^{1/n} (2R)^{n-1}$$

$$\leq 2 (2\epsilon_{n+1})^{1/n} (2R)^{n-1} \leq (2R)^n \epsilon_{n+1}^{1/n} = \epsilon_n ,$$

upon using (4.17), (4.16), and (4.12). But (4.25) is simply (4.14) with $n - 1$ for n. Hence we may repeat the process successively until, for any given order of the roots β_j, we have picked out the roots a_j in such an order that

$$(4.26) \qquad\qquad |\beta_j - a_j| < (2\epsilon_{j+1})^{1/j} R < \epsilon_j$$

by virtue of (4.19), (4.12), and $R \geq 2$.

The proof given above follows the general lines of an earlier attack by Coolidge [2], but care has been taken to ensure the pairing of the roots. For our purposes it would *not* have sufficed to show that every root of $P_n(z) = 0$ is near some root of $Q_n(z) = 0$ and vice versa.

LEMMA 4.3. *None of the roots of the equation $g(z) \equiv \sum_{j=0}^{n} c_j z^j = 0$ lie outside the circle* $|z| = 1 + M/L$, *where $L = |c_n|$ and $M = \max_{0 \leq j < n} |c_j|$.*

For suppose, on the contrary, that $g(z_0) = 0$ and $H = |z_0| > 1 + M/L \geq 1$. Then

$$0 = \left| \sum_{j=0}^{n} c_j z_0^j \right| \geq LH^n - M \sum_{j=0}^{n-1} H^j = LH^n - M(H^n - 1)/(H - 1) .$$ Since $H > 1$, we

deduce $0 \geq L(H - 1) - M(1 - H^{-n}) \geq L(H - 1) - M$, and thence the contradiction $H \leq 1 + M/L$.

THEOREM 4.2. *The n-valued function consisting of the zeros of the polynomial $\sum_{j=0}^{n} c_j z^j = 0$ is continuous at any point c such that $c_n \neq 0$ and $c_{n-1}, c_{n-2}, \cdots, c_0$ are finite.*

Suppose $\epsilon > 0$ is prescribed. Consider any fixed c such that $c_n \neq 0$ and $c_{n-1}, c_{n-2}, \cdots, c_0$ are finite. Choose $\delta > 0$ so that

$$(4.27) \qquad \epsilon > \delta > 0 , \qquad |c_n| \geq \delta , \qquad \sum_{j=0}^{n-1} |c_j| \leq \frac{1}{\delta}, \qquad \delta \leq \tfrac{1}{2} .$$

Define

(4.28) $$\eta = \eta\,(\epsilon,\,c) = \tfrac{1}{2}\,\delta^3\left(\frac{\delta^{4\epsilon+1}}{2^{4\epsilon}}\right)^{n!}.$$

Let

(4.29) $$g\,(z) = \sum_{j=0}^{n} c_j z^j, \qquad h\,(z) = \sum_{j=0}^{n} d_j z^j, \qquad |c - d| < \eta.$$

Then

(4.30) $$|c_j - d_j| < \eta.$$

In view of the definition of continuity in section 3, it is enough to show that the zeros of $g(z)$ and $h(z)$ can be arranged in pairs such that the distance in the complex plane between the two members of any pair is less than δ.

Lemma 4.3, in conjunction with (4.27), (4.28), and (4.29), shows that none of the roots of $g(z) = 0$ and $h(z) = 0$ lie outside the circle $|z| = 1 + (\delta^{-1} + \eta)/(\delta - \eta) < 2/\delta^2$. In the notation of lemma 4.2, we take $R = 2/\delta^2$ and $\epsilon_1 = \delta$. This ensures $0 < \epsilon_1 \leq 2 \leq R < \infty$. The roots of $g(z) = 0$ and $h(z) = 0$ lie inside the circle $|z| = R$. Write $p_{nj} = c_j/c_n$ and $q_{nj} = d_j/d_n$; so that $p_{nn} = q_{nn} = 1$, and $P_n(z) = 0$ has the same roots as $g(z) = 0$ while $Q_n(z) = 0$ has the same roots as $h(z) = 0$. Now, for $0 \leq j \leq n - 1$,

(4.31) $$|p_{nj} - q_{nj}| = \left| \frac{c_j\,[\,c_n + (d_n - c_n)\,] - c_n\,[\,c_j + (d_j - c_j)\,]}{c_n\,[\,c_n + (d_n - c_n)\,]} \right|$$

$$\leq \left| \frac{c_j\,(d_n - c_n)}{c_n\,[\,c_n + (d_n - c_n)\,]} \right| + \left| \frac{d_j - c_j}{c_n + (d_n - c_n)} \right|$$

$$\leq \frac{\delta^{-1}\eta}{\delta\,(\delta - \eta)} + \frac{\eta}{\delta - \eta} \leq \frac{2\eta}{\delta^3}$$

by virtue of (4.30) and (4.27) and $\delta \leq \tfrac{1}{2}$. Finally, from (4.12),

(4.32) $$\log \epsilon_{j+1} = j \log \epsilon_j - j^2 \log\,(2R).$$

Multiply this equation by $n!/j!$ and sum over $1 \leq j \leq n$, to yield

(4.33) $$\log \epsilon_{n+1} = n!\left[\log \epsilon_1 - \{\log\,(2R)\,\}\sum_{j=1}^{n} \frac{j}{(j-1)!}\right]$$

$$\geq n!\left[\log \epsilon_1 - \{\log\,(2R)\,\}\sum_{j=1}^{\infty} \frac{j}{(j-1)!}\right] = n!\,[\log \epsilon_1 - 2\,e \log\,(2R)\,]$$

$$\geq n!\,\log\left(\frac{\delta^{4\epsilon+1}}{2^{4\epsilon}}\right) = \log\left(\frac{2\eta}{\delta^3}\right).$$

The result now follows from (4.31), (4.33), and lemma 4.2, since the ϵ_j form a decreasing sequence.

5. Real isomorphs of complex matrices

Any complex matrix Z can be expressed in the form $Z = X + iY$, where X and Y are real matrices. We define the *real isomorph* of Z to be

(5.1) $$\hat{Z} = \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}.$$

Wedderburn (see p. 100 in [18]) uses this device in the study of Hermitian matrices: a matrix is Hermitian if and only if its real isomorph is symmetric. We use the word "isomorph" in view of the easily verified relations

$$(5.2) \qquad \widehat{Z_1 + Z_2} = \hat{Z}_1 + \hat{Z}_2 \, ,$$

$$\widehat{Z_1 Z_2} = \hat{Z}_1 \hat{Z}_2 \, ,$$

which are valid in the sense that, when one side of one of the equations (5.2) exists, the other side of that equation exists and the two sides are equal. The matrix identity

$$(5.3) \qquad \begin{pmatrix} I & iI \\ 0 & I \end{pmatrix} \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \begin{pmatrix} I & -iI \\ 0 & I \end{pmatrix} = \begin{pmatrix} X+iY & 0 \\ Y & X-iY \end{pmatrix}$$

implies

THEOREM 5.1. *If Z is a square matrix, the determinant of $\hat{Z}$ equals the square of the modulus of the determinant of Z.*

We have already mentioned that the complex number $z = x + iy$ may be regarded as a vector $z = \{x, y\}$. Corresponding to the latter form, we shall write dz as an abbreviation for $dxdy$, while dz will represent the ordinary complex differential element. If a multiple integral in a number of complex variables $w_1, w_2, \cdots, w_m$ is to be transformed to an integral in the complex variables $z_1, z_2, \cdots, z_m$ by means of the relations

$$(5.4) \qquad w_j = w_j (z_1, z_2, \cdots, z_m), \qquad\qquad j = 1, 2, \cdots, m \, ,$$

the appropriate Jacobian $\partial(w_1, \cdots, w_m)/\partial(z_1, \cdots, z_m)$ will consist, in the usual way, of the determinant whose elements are $\partial w_j/\partial z_k$. Let us suppose that, for each j and k, w_j is an analytic function of z_k when $z_1, \cdots, z_{k-1}, z_{k+1}, \cdots, z_m$ are held fixed. The familiar Cauchy-Riemann equations assert, for an analytic function $w(z) = u(z) + iv(z)$, that

$$(5.5) \qquad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = \Re \frac{dw}{dz}, \qquad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} = \Im \frac{dw}{dz},$$

where $\Re$ and $\Im$ denote the real and imaginary parts of the ensuing expression. From (5.5) and (5.1), it follows that, if J is the Jacobian matrix of the transformation from $dw_1 \cdots dw_m$ to $dz_1 \cdots dz_m$, then $\hat{J}$ is the Jacobian matrix of the corresponding transformation from $dw_1 \cdots dw_m$ to $dz_1 \cdots dz_m$. Hence, theorem 5.1 implies

THEOREM 5.2. *If*

$$(5.6) \qquad dw_1 \cdots dw_m = \frac{\partial (w_1, \cdots, w_m)}{\partial (z_1, \cdots, z_m)} \, dz_1 \cdots dz_m,$$

then

$$(5.7) \qquad dw_1 \cdots dw_m = \left| \frac{\partial (w_1, \cdots, w_m)}{\partial (z_1, \cdots, z_m)} \right|^2 dz_1 \cdots dz_m.$$

Further than this, the isomorphism exhibited in (5.2) enables us to construct directly the analogues of all the familiar theorems on Jacobians, which rest on the ordinary properties of matrix multiplication. For instance,

THEOREM 5.3. *If the variables $w_1, \cdots, w_m$ are related to the variables $z_1, \cdots, z_m$ by analytic equations*

$$(5.8) \qquad F_j(w_1, \cdots, w_m, z_1, \cdots, z_m) = 0 \, , \qquad j = 1, 2, \cdots, m \, ,$$

then

$$(5.9) \qquad dw_1 \cdots dw_m = \left| \frac{\partial (F_1, \cdots, F_m)}{\partial (z_1, \cdots, z_m)} \bigg/ \frac{\partial (F_1, \cdots, F_m)}{\partial (w_1, \cdots, w_m)} \right|^2 dz_1 \cdots dz_m.$$

6. **The moment-generating set function of the zeros of a polynomial with generally distributed coefficients**

We derive a formula for the moment-generating set function of the zeros of a polynomial $g^*(z) = \sum_{j=0}^{n} c_j^* z^j$, when the joint cumulative distribution function of the coefficients has the arbitrary form

$$(6.1) \qquad W(c) = W(a_0, a_1, \cdots, a_n, b_0, b_1, \cdots, b_n).$$

We first obtain this moment-generating set function when certain restrictions are placed on W, and then later invoke theorems 3.1 and 3.2 to relax these restrictions. Accordingly we suppose until further notice that W possesses a continuous frequency function

$$(6.2) \qquad w(c) = \frac{\partial^{2n+2} W(c)}{\partial a_0 \partial a_1 \cdots \partial a_n \partial b_0 \partial b_1 \cdots \partial b_n},$$

and that all moments of c^* exist.

Let $\mathfrak{C}$ denote the $(2n+2)$-dimensional Euclidean space consisting of points c, and $\mathfrak{Z}$ the complex plane. For $p = 1, 2, \cdots, r$ let dz_p be the small rectangle

$$(6.3) \qquad x_p < x \leqq x_p + dx_p, \qquad y_p < y \leqq y_p + dy_p,$$

where $z_p = x_p + iy_p$ are prescribed. We are going to evaluate the probability that $g^*(z) = 0$ has precisely one root in each of the small rectangles dz_p. The probability will be written in the form

$$(6.4) \qquad P(z_1, \cdots, z_r) \, dz_1 \cdots dz_r.$$

This expression corresponds to Ramakrishnan's *product density* [16], introduced in the treatment of cosmic ray phenomena. However, in Ramakrishnan's work there was a certain degree of independence between behaviours of distinct quanta; whereas in the present context we do not enjoy any such simplification because, when one root of g^* is known, the conditional distribution of the remaining roots is affected.

For fixed $z = x + iy$, the equations

$$(6.5) \qquad \Re g(z) = \Im g(z) = 0$$

define a $2n$-flat $\mathfrak{f}_z$ in $\mathfrak{C}$. As z varies, the flats $\mathfrak{f}_z$ develop a twisted regulus $\mathfrak{r}$. The generator of $\mathfrak{r}$ lying in $\mathfrak{f}_z$ is a $(2n-2)$-flat $\mathfrak{g}_z$ with equation

$$(6.6) \qquad \Re g(z) = \Im g(z) = \Re g'(z) = \Im g'(z) = 0.$$

The Cauchy-Riemann equations provide the reason why $\mathfrak{g}_z$ is a $(2n-2)$-flat and not a $(2n-4)$-flat, as one might at first sight imagine.

If $g^*(z) = 0$ has at least one root in each of the rectangles dz_p, $p = 1, 2, \cdots, r$, then c^* must lie at the intersection of the r $2n$-flats $\mathfrak{f}_{z_p}$. If, further, one of the rectangles dz_p (say dz_r) contains more than one root of $g^*(z) = 0$, then c^* must also lie on $\mathfrak{g}_{z_r}$. Since $w(c)$ is supposed continuous, it follows that the probability that each of the dz_p contains precisely one root differs by terms of higher order than $dz_1 \cdots dz_r$. Hence

$$(6.7) \qquad P(z_1, \cdots, z_r) \, dz_1 \cdots dz_r = \int_c w(c) \, da_0 \cdots da_n db_0 \cdots db_n,$$

where c is the intersection of the $\mathfrak{f}_{z_p}$ for $z_p \in dz_p$. To evaluate (6.7), we transform from the coordinate system $(a_0, a_1, \cdots, a_n, b_0, b_1, \cdots, b_n)$ to the coordinate system $(x_1, x_2, \cdots,$

$x_r, a_r, a_{r+1}, \cdots, a_n, y_1, y_2, \cdots, y_r, b_r, b_{r+1}, \cdots, b_n)$. The relationships between these systems are

$$(6.8) \qquad\qquad g(z_p) = 0, \qquad\qquad p = 1, 2, \cdots, r.$$

Now this is the situation envisaged by theorem 5.3. The appropriate Jacobian is accordingly

$$(6.9) \quad \left| \frac{\partial(g, \cdots\ g)}{\partial(z_1, \cdots, z_r)} \bigg/ \frac{\partial(g, \cdots, g)}{\partial(c_0, \cdots, c_{r-1})} \right|^2$$

$$= \left\| \begin{matrix} g'(z_1) & 0 & \cdots & 0 \\ 0 & g'(z_2) & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & g'(z_r) \end{matrix} \right\| \bigg/ \left\| \begin{matrix} 1 & z_1 & \cdots & z_1^{r-1} \\ 1 & z_2 & \cdots & z_2^{r-1} \\ \cdot & \cdot & \cdots & \cdot \\ 1 & z_r & \cdots & z_r^{r-1} \end{matrix} \right\|^2$$

$$= \prod_{p=1}^{r} |g'(z_p)|^2 \bigg/ \prod_{1 \leq p < q \leq r} |z_p - z_q|^2.$$

In (6.9) the denominator is to be taken as unity if $r = 1$. Thus

$$(6.10) \quad P(z_1, \cdots, z_r) = \int\limits_{g(z_p) = 0;\ p = 1, \cdots, r} \frac{\displaystyle\prod_{p=1}^{r} |g'(z_p)|^2}{\displaystyle\prod_{1 \leq p < q \leq r} |z_p - z_q|^2} w(c)\, dc_r \cdots dc_n.$$

In (6.10) the conditions defining the range of integration are to be used to express $c_0, \cdots, c_{r-1}$ in terms of $c_r, \cdots, c_n$ and $z_1, \cdots, z_n$. The integral (6.10) exists because all the moments of $w(c)$ are supposed to exist.

We now consider the important special case $r = n$. The integration in (6.10) will be with respect to dc_n only. Since

$$(6.11) \qquad\qquad g(z) = c_n \prod_{j=1}^{n} (z - z_j),$$

we have

$$(6.12) \qquad\qquad g'(z_p) = c_n \prod_{j \neq p}^{n} (z_p - z_j);$$

and hence

$$(6.13) \qquad\qquad \prod_{p=1}^{n} |g'(z_p)|^2 = |c_n|^{2n} \prod_{1 \leq p < q \leq n} |z_p - z_q|^4.$$

Thus

$$(6.14) \qquad P(z_1, \cdots, z_n) = \left[\prod_{1 \leq p < q \leq n} |z_p - z_q|^2 \right] \int |c_n|^{2n} w(c)\, dc_n.$$

In (6.14) it is supposed that the argument c of $w(c)$ is expressed in the form

$$(6.15) \qquad\qquad a_{n-r} = \Re(c_n H_r), \qquad b_{n-r} = \Im(c_n H_r), \qquad r = 1, 2, \cdots, n,$$

where H_r is the sum of the products of $-z_1, -z_2, \cdots, -z_n$ taken r at a time. Since there are $n!$ permutations of $z_1, z_2, \cdots, z_n$, we have on combination of (3.3) and (6.14)

$$(6.16) \quad M(t, \mathfrak{z}) = \frac{1}{n!} \left[\prod_{j=1}^{n} \left\{ \int_{z_j \in \mathfrak{z}'} + e^t \int_{z_j \in \mathfrak{z}} \right\} \right] P(z_1, \cdots, z_n) \, dz_1 \cdots dz_n.$$

Another important special case of (6.10) is $r = 1$. We obtain

$$(6.17) \qquad\qquad P(z_1) = \int_{g(z_1)=0} |g'(z_1)|^2 w(c) \, dc_1 \cdots dc_n$$

$$= \int_{g(z_1)=0} |g'|^2 \psi(g, g') \, dg',$$

where $\psi(g, g')$ is the joint frequency function of the real and imaginary parts of $g^*(z_1)$ and $g^{*\prime}(z_1)$. However $P(z_1) dz_1$ is the expected number of roots in the small rectangle dz_1; and therefore the condensed probability set function of the zeros of $g^*(z)$ is

$$(6.18) \qquad\qquad F(\mathfrak{z}) = \frac{1}{n} \int_{\mathfrak{z}} P(z_1) \, dz_1.$$

Finally we can remove the restrictions originally placed on c^*. For we can find a sequence of random variables c_ν^*, $\nu = 1, 2, \cdots$, such that each c_ν^* has moments of all order and a continuous frequency function, and such that $\operatorname*{dlim}_{\nu \to \infty} c_\nu^* = c^*$. Corresponding to c_ν^* we can find $M_\nu(t, \mathfrak{z})$ or $F_\nu(\mathfrak{z})$ as in (6.16) or (6.18); and, when $\mathfrak{z}$ is open,

$$(6.19) \qquad\qquad M(t, \mathfrak{z}) = \lim_{\nu \to \infty} M_\nu(t, \mathfrak{z}), \qquad F(\mathfrak{z}) = \lim_{\nu \to \infty} F_\nu(\mathfrak{z})$$

which provides the required set functions corresponding to c^*, in accordance with theorems 3.1 and 3.2. In removing the restrictions on c^*, we ought to assume (in order to avail ourselves of theorems 3.2 and 4.2) that $c_n^* = 0$ has not a positive probability. But if $c_n^* = 0$ has a positive probability, there is the same positive probability of an infinite root of $g^* = 0$; and we can work with the conditional probability given that $c_n^* \neq 0$.

7. The characteristic functional

The characteristic functional

$$(7.1) \qquad\qquad C[\theta] = E\left[\exp\left\{ i \sum_{j=1}^{n} \theta(z_j^*) \right\} \right]$$

offers a more general approach than the moment-generating set function and the method of product densities considered in section 6. On the other hand it seems much less tractable, and I have been unable to derive any useful results from it; but I present it nevertheless in the hope that the reader may manage to manipulate it successfully. In (7.1), $\theta(z)$ is an arbitrary function of z, and z_j^* are the zeros of $g^*(z) = 0$. We have to express $\sum_{j=1}^{n} \theta(z_j^*)$ in terms of g^*, and then take expected values in (7.1) with respect to the distribution of g^*. The first part of this programme can be achieved, at least symbolically, when θ is an integral function by using Cauchy's theorem:

$$(7.2) \qquad C[\theta] = E\left[\lim_{R \to \infty} \exp\left\{ (2\pi)^{-1} \oint_{|z|=R} \theta(z) \, d \log g^*(z) \right\} \right].$$

The second part of the programme is however harder, since $g^*(z)$ regarded as a function of z is a deterministic random function whose value is known for all z as soon as it is known for n distinct values $\zeta_1, \cdots, \zeta_n$. Thus, by Lagrange's formula,

$$(7.3) \qquad g^*(z) = \sum_{j=1}^{n} g^*(\zeta_j) \prod_{k \neq j}^{n} \frac{z - \zeta_k}{\zeta_j - \zeta_k}$$

can be inserted in (7.2) and the expected value taken over the n-dimensional joint distribution of $g^*(\zeta_j)$ for convenient fixed ζ_j. But the manipulative difficulties seem severe.

The characteristic functional reduces to the moment-generating set function for a set $\mathfrak{z}$ if θ is taken equal to $-it$ times the indicator function of $\mathfrak{z}$. If $\mathfrak{z}$ is bounded by a Jordan curve and $\Delta_{\mathfrak{z}} \log g^*(z)$ denotes the variation of $\log g^*(z)$ in describing this curve, we may write

$$(7.4) \qquad M(t, \mathfrak{z}) = E\left[\exp \left\{ \frac{t}{2\pi i} \Delta_{\mathfrak{z}} \log g^*(z) \right\} \right],$$

into which (7.3) may be substituted as before.

8. Normally distributed complex coefficients

We shall determine the condensed distribution of the zeros of $g^*(z)$ when the coefficients c^* are normally distributed about a mean γ with nondegenerate variance-covariance matrix V. The frequency function of c^* is thus

$$(8.1) \qquad (2\pi)^{-n-1} |V|^{-1/2} \exp \left\{ -\tfrac{1}{2} (c - \gamma)^\dagger V^{-1} (c - \gamma) \right\}.$$

We require

LEMMA 8.1. *Assuming conformable partitioning and the existence of the relevant inverses, and provided A is symmetric,*

$$(8.2) \quad a^\dagger A^{-1} a = (a_1^\dagger, a_2^\dagger) \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$= a_1^\dagger A_{11}^{-1} a_1 + (a_2 - A_{21} A_{11}^{-1} a_1)^\dagger (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} (a_2 - A_{21} A_{11}^{-1} a_1).$$

Let

$$(8.3) \qquad AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = I,$$

so that

$$(8.4) \qquad \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

The upper right-hand corner of (8.4) gives

$$(8.5) \qquad B_{12} = -A_{11}^{-1} A_{12} B_{22},$$

and substitution of (8.5) into the lower right-hand corner of (8.4) gives

$$(8.6) \qquad B_{22} = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}.$$

Similarly the upper right-hand corner of (8.4) gives $A_{12} = -A_{11} B_{12} B_{22}^{-1}$, which substituted into the upper left-hand corner of (8.4) yields

$$(8.7) \qquad A_{11}^{-1} = B_{11} - B_{12} B_{22}^{-1} B_{21}.$$

Since A is symmetric, so is B; and (8.5) gives

$$(8.8) \qquad A_{21} A_{11}^{-1} = (A_{11}^{-1} A_{12})^\dagger = (-B_{12} B_{22}^{-1})^\dagger = -B_{22}^{-1} B_{21}.$$

From (8.2), (8.6), (8.7), (8.8)

$$(8.9) \quad a_1^\dagger A_{11}^{-1} a_1 + (a_2 - A_{21} A_{11}^{-1} a_1)^\dagger (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} (a_2 - A_{21} A_{11}^{-1} a_1)$$

$$= a_1^\dagger (B_{11} - B_{12} B_{22}^{-1} B_{21}) a_1 + (a_2 + B_{22}^{-1} B_{21} a_1)^\dagger B_{22} (a_2 + B_{22}^{-1} B_{21} a_1)$$

$$= a_1^\dagger B_{11} a_1 + 2 a_1^\dagger B_{12} a_2 + a_2^\dagger B_{22} a_2 = a^\dagger B a = a^\dagger A^{-1} a .$$

LEMMA 8.2. $|A| = |A_{11}| \, |A_{22} - A_{21} A_{11}^{-1} A_{12}|$.
This follows immediately from the identity

$$(8.10) \qquad \begin{pmatrix} I & 0 \\ -A_{21} A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} - A_{21} A_{11}^{-1} A_{12} \end{pmatrix}.$$

Let $z = x + iy$ be a fixed complex number, and define

$$(8.11) \qquad \begin{cases} z_0 = (1, z, z^2, \cdots, z^n) \\ z_1 = (0, 1, 2z, \cdots, n z^{n-1}) \\ Z = \begin{pmatrix} z_0 \\ z_1 \end{pmatrix}. \end{cases}$$

It is easy to verify that the column vector q defined by

$$(8.12) \qquad\qquad q = \{\Re g, \Re g', \Im g, \Im g'\} = \hat{Z} c ;$$

and hence the real and imaginary parts of $g^*(z)$ and $g^{*\prime}(z)$ are normally distributed about a mean $\hat{Z}\gamma$ with variance-covariance matrix $\hat{Z} V \hat{Z}^\dagger$. Here and elsewhere $\hat{Z}^\dagger$ denotes the transpose of $\hat{Z}$ and not the real isomorph of $Z^\dagger$. From (6.17) the condensed frequency function of the zeros of $g^*(z) = 0$ is

$$(8.13) \quad P(z) = \frac{1}{4\pi^2 n} \int_{g=0} |g'|^2 |\hat{Z} V \hat{Z}^\dagger|^{-1/2}$$

$$\exp\{-\tfrac{1}{2}(q - \hat{Z}\gamma)^\dagger (\hat{Z} V \hat{Z}^\dagger)^{-1} (q - Z\gamma)\} dg' .$$

Now identify the matrix $\hat{Z} V \hat{Z}^\dagger$ with the matrix A of lemma 8.1, taking the partitioning such that A_{11} is the variance-covariance matrix of the real and imaginary parts of g^*. Accordingly we take

$$(8.14) \qquad\qquad a_1 = -\hat{z}_0 \gamma , \qquad a_2 = g' - \hat{z}_1 \gamma ,$$

so that lemmas 8.1 and 8.2 yield

$$(8.15) \quad P(z) = \frac{1}{4\pi^2 n} |A_{11}|^{-1/2} |A_{22} - A_{21} A_{11}^{-1} A_{12}|^{-1/2} \int (a_2 + \hat{z}_1 \gamma)^\dagger (a_2 + \hat{z}_1 \gamma)$$

$$\cdot \exp\{-\tfrac{1}{2} a_1^\dagger A_{11}^{-1} a_1 - \tfrac{1}{2}(a_2 - A_{21} A_{11}^{-1} a_1)^\dagger (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}$$

$$\cdot (a_2 - A_{21} A_{11}^{-1} a_1)\} da_2 .$$

Put $b = a_2 - A_{21} A_{11}^{-1} a_1$. Then (8.6) gives, using results from Turnbull and Aitken (see pp. 175–176 in [17]),

$$(8.16) \quad P(z) = \frac{\exp\{-\tfrac{1}{2} a_1^\dagger A_{11}^{-1} a_1\}}{2\pi n |A_{11}|^{1/2}} \cdot \frac{|B_{22}|^{1/2}}{2\pi}$$

$$\cdot \int \exp\{-\tfrac{1}{2} b^\dagger B_{22} b\} (b + \hat{z}_1 \gamma + A_{21} A_{11}^{-1} a_1)^\dagger (b + \hat{z}_1 \gamma + A_{21} A_{11}^{-1} a_1) \, db$$

$$= \exp\{-\tfrac{1}{2} a_1^\dagger A_{11} a_1\} \frac{\text{trace}\{D\}}{2\pi n |A_{11}|^{1/2}},$$

where

$$(8.17) \qquad D = B_{22}^{-1} + (\hat{z}_1\gamma + A_{21}A_{11}^{-1}a_1)^{\dagger}(\hat{z}_1\gamma + A_{21}A_{11}^{-1}a_1).$$

Since

$$(8.18) \qquad A_{11} = \hat{z}_0 V \hat{z}_0^{\dagger}, \qquad A_{12} = A_{21}^{\dagger} = \hat{z}_0 V \hat{z}_1^{\dagger}, \qquad A_{22} = \hat{z}_1 V \hat{z}_1^{\dagger},$$

we obtain

$$(8.19) \quad D = \hat{z}_1 V \hat{z}_1^{\dagger} - \hat{z}_1 V \hat{z}_0^{\dagger}(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}\hat{z}_0 V \hat{z}_1^{\dagger}$$
$$+ [\hat{z}_1\gamma - \hat{z}_1 V \hat{z}_0^{\dagger}(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}\hat{z}_0\gamma][\hat{z}_1\gamma - \hat{z}_1 V \hat{z}_0^{\dagger}(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}\hat{z}_0\gamma]^{\dagger}$$
$$= [\hat{z}_1 - (\hat{z}_1 V \hat{z}_0^{\dagger})(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}\hat{z}_0][V + \gamma\gamma^{\dagger}][\hat{z}_1 - (\hat{z}_1 V \hat{z}_0^{\dagger})(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}\hat{z}_0]^{\dagger}.$$

We have thus established

THEOREM 8.1. *If the coefficients* $c^* = \{a_0^*, a_1^*, \cdots, a_n^*, b_0^*, b_1^*, \cdots, b_n^*\}$ *of the equation*

$$(8.20) \qquad g^*(z) = \sum_{j=0}^{n} (a_j^* + ib_j^*)(x + iy)^j = 0$$

are distributed normally about a vector mean γ *with nondegenerate variance-covariance matrix* V, *the condensed frequency function of the roots of* $g^*(z) = 0$ *is*

$$(8.21) \quad P(z) = P(x, y) = \frac{\exp\{-\frac{1}{2}(\hat{z}_0\gamma)^{\dagger}(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}(\hat{z}_0\gamma)\}}{2\pi n |\hat{z}_0 V \hat{z}_0^{\dagger}|^{1/2}} \text{ trace}(\zeta M \zeta^{\dagger}),$$

where

$$(8.22) \qquad z_0 = (1, z, z^2, \cdots, z^n), \qquad z_1 = (0, 1, 2z, \cdots, nz^{n-1}),$$

and

$$(8.23) \qquad \zeta = \hat{z}_1 - (\hat{z}_1 V \hat{z}_0^{\dagger})(\hat{z}_0 V \hat{z}_0^{\dagger})^{-1}\hat{z}_0,$$

and

$$(8.24) \qquad M = V + \gamma\gamma^{\dagger}$$

is the matrix of second moments of c^* *about the origin.*

In deriving this result we have assumed that V is nonsingular. However, in view of theorem 3.1, theorem 8.1 remains valid provided $\hat{z}_0 V \hat{z}_0^{\dagger}$ is nonsingular.

9. Normally distributed real coefficients

The case of normally distributed real coefficients arises when c^* is distributed normally about a mean $\gamma = \{a^{\dagger}, 0\}$ with variance-covariance matrix

$$(9.1) \qquad V = \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix}.$$

We shall assume that the distribution is nondegenerate in the sense that U is nonsingular.

As we have just remarked in section 8, the condensed distribution of the zeros is given by theorem 8.1 provided $\hat{z}_0 V \hat{z}_0^{\dagger}$ is nonsingular. Accordingly we shall prove

THEOREM 9.1. *If* U *is nonsingular and* $|y| \geq \delta > 0$, *theorem 8.1 holds.*

Let

$$(9.2) \qquad x_0 = (1, x, x^2, \cdots, x^n), \qquad x_j = \frac{d^j x_0}{j! dx^j}, \qquad j = 1, 2, \cdots, n.$$

Set $X = \{x_0, x_1, \cdots, x_n\}$. Since $|X| = 1$, $XUX^\dagger$ is positive definite. By the matrix form of Taylor's theorem

$$(9.3) \qquad z_0 = (1, iy, i^2 y^2, \cdots, i^n y^n) X .$$

Hence

$$(9.4) \qquad \hat{z}_0 = \begin{pmatrix} y_0 X & -y_1 X \\ y_1 X & y_0 X \end{pmatrix}$$

where $y_0 + iy_1 = (1, iy, i^2 y^2, \cdots, i^n y^n)$ and y_0, y_1 are real. Since $n \geqq 1$ and $|y| \geqq \delta > 0$, neither y_0 nor y_1 is null, and there is an $(n+1) \times (n+1)$ nonsingular matrix Y whose first and second rows are y_0 and y_1. Thus

$$(9.5) \qquad \hat{z}_0 \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} \hat{z}_0^\dagger = \begin{pmatrix} y_0\, XUX^\dagger\, y_0^\dagger & y_0\, XUX^\dagger\, y_1^\dagger \\ y_1\, XUX^\dagger\, y_0^\dagger & y_1\, XUX^\dagger\, y_1^\dagger \end{pmatrix}$$

is a minor on the leading diagonal of the positive-definite matrix $YXUX^\dagger Y^\dagger$ and is therefore positive definite itself.

There remains the interesting case $y = 0$. We shall show that the real axis supports a line density of the condensed distribution of the zeros of g^*. We write in place of (9.1)

$$(9.6) \qquad V = \begin{pmatrix} U & 0 \\ 0 & v^{16} I \end{pmatrix}, \qquad 0 < \delta = v^7 \leqq (\tfrac{1}{2}\pi)^{1/2}$$

and proceed to evaluate

$$(9.7) \qquad f(x) = \lim_{v \to 0+} \int_{-\delta}^{\delta} P(z)\, dy ,$$

which will then represent the required line density on the real axis, in view of theorem 3.1. Define

$$(9.8) \qquad \left. \begin{aligned} I_{jk} &= I_{kj} = x_j x_k^\dagger \\ U_{jk} &= U_{kj} = x_j U\, x_k^\dagger \end{aligned} \right\} \qquad j, k = 0, 1, \cdots, n .$$

Put

$$(9.9) \qquad y = \lambda v^8 , \qquad |y| \leqq \delta = v^7 , \qquad |\lambda| \leqq v^{-1} .$$

Then

$$(9.10) \qquad \hat{z}_0 = \begin{pmatrix} x_0 - \lambda^2 v^{16} x_2 , & -\lambda v^8 x_1 \\ \lambda v^8 x_1 , & x_0 - \lambda^2 v^{16} x_2 \end{pmatrix} + O(v^{21}) ,$$

$$(9.11) \qquad \hat{z}_1 = \begin{pmatrix} x_1 - 3\lambda^2 v^{16} x_3 , & -2\lambda v^8 x_2 \\ 2\lambda v^8 x_2 , & x_1 - 3\lambda^2 v^{16} x_3 \end{pmatrix} + O(v^{21}) .$$

Thus

$$(9.12) \qquad \hat{z}_0 V \hat{z}_0^\dagger = \begin{pmatrix} U_{00} - 2\lambda^2 v^{16} U_{02} , & \lambda v^8 U_{01} \\ \lambda v^8 U_{01} , & v^{16} I_{00} + \lambda^2 v^{16} U_{11} \end{pmatrix} + O(v^{21})$$

and

$$(9.13) \qquad |\hat{z}_0 V \hat{z}_0^\dagger| = \Delta v^{16} + O(v^{21}) ,$$

where

$$(9.14) \qquad \Delta = U_{00} I_{00} + \lambda^2 (U_{00} U_{11} - U_{01}^2) .$$

Hence

$$(9.15) \qquad (\hat{z}_0 V \hat{z}_0^\dagger)^{-1} = \Delta^{-1} \begin{pmatrix} I_{00} + \lambda^2 U_{11} , & -\lambda v^{-8} U_{01} \\ -\lambda v^{-8} U_{01} , & v^{-16} U_{00} - 2\lambda^2 U_{02} \end{pmatrix} + O(v^5) .$$

It is evident that $\Delta > 0$ since U_{00} and

(9.16)
$$\begin{pmatrix} U_{00} & U_{01} \\ U_{01} & U_{11} \end{pmatrix}$$

are minors on the principal diagonal of $XUX^\dagger$. Next write

(9.17) $A_j = x_j \mathbf{a}$, $i = 0, 1, 2, \cdots, n$

From (9.10)

(9.18) $\hat{z}_0 \gamma = \begin{pmatrix} A_0 - \lambda^2 v^{16} A_2 \\ \lambda v^8 A_1 \end{pmatrix} + O(v^{21})$.

Some straightforward algebra now gives

(9.19) $-\frac{1}{2}(\hat{z}_0 \gamma)^\dagger (\hat{z}_0 V \hat{z}_0^\dagger)^{-1}(\hat{z}_0 \gamma)$

$$= -\frac{1}{2\Delta}\{I_{00}A_0^2 + \lambda^2(U_{11}A_0^2 - 2U_{01}A_0A_1 + U_{00}A_1^2)\} + O(v^5),$$

(9.20) $\text{trace}(\boldsymbol{\zeta} M \boldsymbol{\zeta}^\dagger) = \Delta^{-1}I_{00}(U_{00}U_{11} - U_{01}^2) + \Delta^{-2}(U_{00}A_1 - U_{01}A_0)^2 + O(v^5)$.

Hence, from (8.21)

(9.21) $P(z) = \dfrac{1}{2\pi n v^8}\left\{\dfrac{I_{00}(U_{00}U_{11} - U_{01}^2)}{\Delta^{3/2}} + \dfrac{I_{00}^2(U_{00}A_1 - U_{01}A_0)^2}{\Delta^{5/2}}\right\}$

$$\cdot \exp\left\{-\frac{1}{2\Delta}[I_{00}A_0^2 + \lambda^2(U_{11}A_0^2 - 2U_{01}A_0A_1 + U_{00}A_1^2)]\right\} + O(v^{-3}).$$

The exponential term in (9.21) is never greater than unity, and the remaining terms involving λ are of the form $\Delta^{-3/2}$ and $\Delta^{-5/2}$. We have to integrate over $-\delta = -v^7 < y = \lambda v^8 < v^7 = \delta$. Hence we may multiply (9.21) by v^8 and integrate over λ from $-v^{-1}$ to v^{-1}. Allowing $v \to 0+$ we obtain from (9.7)

(9.22) $f(x) = \dfrac{1}{\pi n}\displaystyle\int_0^\infty \left\{\dfrac{I_{00}(U_{00}U_{11} - U_{01}^2)}{\Delta^{3/2}} + \dfrac{I_{00}^2(U_{00}A_1 - U_{01}A_0)^2}{\Delta^{5/2}}\right\}$

$$\cdot \exp\left\{-\frac{1}{2\Delta}[I_{00}A_0^2 + \lambda^2(U_{11}A_0^2 - 2U_{01}A_0A_1 + U_{00}A_1^2)]\right\}\, d\lambda.$$

In (9.22) make the substitutions

(9.23) $\mu = \dfrac{I_{00}U_{00}}{\lambda^2}$, $p = U_{00}U_{11} - U_{01}^2$, $q = \dfrac{(U_{00}A_1 - U_{01}A_0)^2}{U_{00}}$.

We get

(9.24) $f(x) = \dfrac{\exp(-\frac{1}{2}A_0^2/U_{00})}{2\pi n U_{00}}\displaystyle\int_0^\infty \dfrac{1}{(\mu+p)^{3/2}}\left\{p + \dfrac{q\mu}{(\mu+p)}\right\}$

$$\cdot \exp\left\{\frac{-q}{2(\mu+p)}\right\}\, d\mu.$$

If $q = 0$, (9.24) gives

(9.25) $f(x) = \dfrac{p^{1/2}\exp(-\frac{1}{2}A_0^2/U_{00})}{\pi n U_{00}}$.

If $q \neq 0$, we make the substitutions

(9.26) $r = \sqrt{\dfrac{q}{p}}$, $t = \sqrt{\dfrac{q}{\mu+p}}$,

and find

$$(9.27) \qquad f(x) = \frac{q^{1/2} \exp\left(-\frac{1}{2} A_0^2 / U_{00}\right)}{(2\pi)^{1/2} n U_{00}} [1 + K(r)],$$

where

$$(9.28) \qquad K(r) = \frac{(2/\pi)^{1/2}}{r^2} \int_r^\infty [t^2 - (1 + r^2)] \, e^{-t^2/2} dt.$$

The Hermitian integrals (see p. 590 in [11]) are defined by

$$(9.29) \qquad \mathrm{Hh}_m(r) = \int_r^\infty \frac{(t-r)^m}{m!} \, e^{-t^2/2} dt.$$

Thus

$$(9.30) \qquad K(r) = \frac{(2/\pi)^{1/2}}{r^2} \{ 2\mathrm{Hh}_2(r) + 2r \, \mathrm{Hh}_1(r) - \mathrm{Hh}_0(r) \}$$

$$= \frac{(2/\pi)^{1/2}}{r} \mathrm{Hh}_1(r),$$

in view of the recurrence relation

$$(9.31) \qquad (m+1) \mathrm{Hh}_{m+1}(r) + r \mathrm{Hh}_m(r) - \mathrm{Hh}_{m-1}(r) = 0.$$

We can now combine (9.25) and (9.27) into a single result and collect the results in

THEOREM 9.2. *If the coefficients* $a^* = (a_0^*, a_1^*, \cdots, a_n^*)$ *of the equation*

$$(9.32) \qquad g^*(z) = \sum_{j=0}^n a_j^* z^j = 0$$

are real and distributed such that the joint distribution of $g^*(x)$ *and* $dg^*(x)/dx$ *is a non-degenerate normal distribution for each fixed* x, *the condensed distribution of the roots of* $g^*(z) = 0$ *has a line-density frequency function on the real axis*

$$(9.33) \quad f(x) = \frac{(U_{00}U_{11} - U_{01}^2)^{1/2} \exp\left(-\frac{1}{2} A_0^2 / U_{00}\right)}{(2\pi)^{1/2} n U_{00}} \left[r + \left(\frac{2}{\pi}\right)^{1/2} \mathrm{Hh}_1(r) \right],$$

where

$$(9.34) \qquad A_0 = E[g^*(x)], \qquad A_1 = E\left[\frac{dg^*(x)}{dx}\right],$$

$$(9.35) \quad U_{00} = \mathrm{var}[g^*(x)], \qquad U_{01} = \mathrm{cov}\left[g^*(x), \frac{dg^*(x)}{dx}\right],$$

$$U_{11} = \mathrm{var}\left[\frac{dg^*(x)}{dx}\right],$$

and

$$(9.36) \qquad r = \frac{|U_{00}A_1 - U_{01}A_0|}{[U_{00}(U_{00}U_{11} - U_{01}^2)]^{1/2}}.$$

Using the British Association tables [1], we may calculate table I. When r is sufficiently large for $K(r)$ to be negligibly small in comparison with unity, we have from (9.27)

$$(9.37) \quad nf(x)\,dx \cong \frac{q^{1/2} \exp\left(-\frac{1}{2} A_0^2 / U_{00}\right)}{(2\pi)^{1/2} U_{00}} \, dx = \frac{1}{\sqrt{2\pi}} \, e^{-\tau^2/2} d\tau$$

where

$$(9.38) \qquad \tau = A_0 U_{00}^{-1/2}$$

is the expected value of g^* divided by its standard error. It will then be a reasonable approximation to take the percentile points[2] of the real roots of $g^*(x)$ as the real roots of the corresponding percentile points of $g^*(x)$. This fact was previously discovered by Geary [9] for the particular case $n = 1$. Also in the particular case $n = 1$, (9.37) forms the basis of Fieller's theorem [6] which is much used in bio-assay. In the usual textbooks on bio-assay (for example, see pp. 27–29 in [8]), there seems to be no allowance for the fact that (9.37) is an approximation involving the neglect of $K(r)$. Fieller and Finney use fiducial theory. I do not know on what logical foundations fiducial theory rests, but I can imagine that Fieller's theorem might be correct within a framework resting on such foundations (if they exist). The particular case when the coefficients a_j^* are independ-

<div align="center">

TABLE I

</div>

r	$K(r)$		r	$K(r)$	
0.0	∞		1.2	0.0935	0408
0.1	7.0187	0662	1.4	0.0523	8306
0.2	3.0684	4636	1.6	0.0290	5246
0.3	1.7784	0828	1.8	0.0158	6176
0.4	1.1521	9418	2.0	0.0084	9070
0.5	0.7911	8623	2.5	0.0016	0331
0.6	0.5622	4244	3.0	0.0002	5477
0.7	0.4082	2679	3.5	0.0000	3342
0.8	0.3005	1809	4.0	0.0000	0357
0.9	0.2231	8031	4.5	0.0000	0031
1.0	0.1666	3094	5.0	0.0000	0002

ently and normally distributed about zero mean with unit variance is of interest. We have $U = I$ and $a = 0$, so $A_0 = A_1 = 0$ and (9.25) gives

$$(9.39) \qquad f(x) = \frac{(I_{00}I_{11} - I_{01}^2)^{1/2}}{\pi n I_{00}} = \frac{1}{2\pi n} \sqrt{\frac{1}{x}\frac{d}{dx}\left[x\frac{d}{dx}\log\sum_{j=0}^{n} x^{2j}\right]}$$

$$= \frac{(1 - h_{n+1}^2)^{1/2}}{\pi n (1 - x^2)},$$

where h_n is defined by (1. 1). Kac's result [12] follows at once on multiplication by n and integration over $-\infty \leqq x \leqq \infty$. Actually (9.39) contains more information than Kac's formula, since it not only affords the average number of real roots but it also shows how they are distributed along the real axis.

<div align="center">

REFERENCES

</div>

[1] *British Association Mathematical Tables*, Vol. 1, 1st ed., Edinburgh, Neill and Co., Ltd., 1931. (*Note:* Important information is omitted from later editions.)

[2] J. L. COOLIDGE, "The continuity of the roots of an algebraic equation," *Annals of Math.*, 2d Ser., Vol. 9 (1907–1908), pp. 116–118.

[3] H. CRAMÉR, *Mathematical Methods of Statistics*, 2d ed., Princeton, Princeton University Press, 1946.

[4] M. A. CREASY, "Limits for the ratio of means," *Jour. Roy. Stat. Soc.*, Ser. B, Vol. 16 (1954), pp. 186–194.

[5] P. ERDÖS and P. TURÁN, "On the distribution of roots of polynomials," *Annals of Math.*, Vol. 51 (1950), pp. 105–119.

[2] For $n \geqq 2$, this is a heuristic statement; I have not attempted to define what is meant by the percentile points of a condensed distribution.

[6] E. C. FIELLER, "The biological standardization of insulin," *Jour. Roy. Stat. Soc.*, Ser. B, Vol. 7 (1940), pp. 1–64.

[7] ——, "Some problems in interval estimation," *Jour. Roy. Stat. Soc.*, Ser. B, Vol. 16 (1954), pp. 175–185.

[8] D. J. FINNEY, *Statistical Method in Biological Assay*, 1st ed., London, Griffin and Co., 1952.

[9] R. C. GEARY, "The frequency distribution of the quotient of two normal variables," *Jour. Roy. Stat. Soc.*, Ser. A, Vol. 93 (1930), pp. 442–446.

[10] J. M. HAMMERSLEY, "An extension of the Slutzky-Fréchet theorem," *Acta Math.* (Uppsala), Vol. 87 (1952), pp. 243–257.

[11] H. JEFFREYS and B. S. JEFFREYS, *Methods of Mathematical Physics*, Cambridge, The University Press, 1946.

[12] M. KAC, "On the average number of real roots of a random algebraic equation," *Bull. Amer. Math. Soc.*, Vol. 49 (1943), pp. 314–320 and p. 938.

[13] ——, "On the average number of real roots of a random algebraic equation," *Proc. Lond. Math. Soc.*, Ser. 2, Vol. 50 (1948), pp. 390–408.

[14] J. E. LITTLEWOOD and A. C. OFFORD, "On the number of real roots of a random algebraic equation," *Jour. Lond. Math. Soc.*, Vol. 13 (1938), pp. 288–295.

[15] ——, "On the number of real roots of a random algebraic equation. II," *Proc. Camb. Phil. Soc.*, Vol. 35 (1939), pp. 133–148.

[16] A. RAMAKRISHNAN, "Stochastic processes relating to particles distributed in a continuous infinity of states," *Proc. Camb. Phil. Soc.*, Vol. 46 (1950), pp. 595–602.

[17] H. W. TURNBULL and A. C. AITKEN, *An Introduction to the Theory of Canonical Matrices*, 3d ed., London and Glasgow, Blackie and Son, 1948.

[18] J. H. M. WEDDERBURN, *Lectures on Matrices*, American Mathematical Society Colloquium Publ., Vol. 17, 1934.

[19] S. S. WILKS, *Mathematical Statistics*, 2d ed., Princeton, Princeton University Press, 1946.

[20] P. ERDÖS and A. C. OFFORD, "On the number of real roots of a random algebraic equation," *Proc. Lond. Math. Soc.*, Ser. 3, Vol. 6 (1956), pp. 139–160.