1. Introduction

L'Enseignement Mathématique, t. 39 (1993), p. 317-348

ZEROS OF POLYNOMIALS WITH 0, 1 COEFFICIENTS

by A.M. ODLYZKO and B. POONEN [1]

ABSTRACT. Zeros of polynomials with 0, 1 coefficients exhibit many interesting features, including fractal appearance. This paper obtains bounds for such zeros. It shows that zeros with a sufficiently large negative real part are real. It also proves that the closure of the set of these zeros is path connected.

1. INTRODUCTION

Zeros of polynomials with random coefficients occur in many scientific and engineering problems. A general overview of the subject and references can be found in the book of Bharucha-Reid and Sambandham [4], which is the basic reference on this topic. There is a wealth of information about distribution of zeros in the complex plane and on the real line. Almost all of the results are for coefficients chosen independently from a common distribution that is continuous, and usually Gaussian.

In this paper we consider zeros of polynomials with 0, 1 coefficients. These zeros have some features that distinguish them from those of the commonly considered families of random polynomials. Let

$$(1.1)\qquad P = \left\{ f(z) : f(z) = 1 + \sum_{j=1}^{d} a_j z^j \,, \quad a_j = 0 \text{ or } 1 \text{ for all } j \right\} .$$

(We exclude polynomials with constant term 0, as their zeros, other than 0, are those of polynomials of lower degree with coefficients 0, 1.) Define

$$(1.2)\qquad W = \{ z \in \mathbf{C} : f(z) = 0 \text{ for some } f \in P \} .$$

[1] Current address: Dept. of Mathematics, Univ. of California, Berkeley, California 94720.

For each degree d, there are 2^{d-1} polynomials $f(z) \in P$ of degree d, and so W is a countable set.

There are few published results about W. In [8] it was shown that $\mathrm{Re}(z) < 3/2$ for all $z \in W$. This was used to prove that if $f(2)$ is a prime for some $f(z) \in P$, then $f(z)$ is irreducible over the rationals. (For further results relating zeros to irreducibility, see [12]. It is conjectured that almost all $f(z) \in P$ are irreducible, but this is still open. This is in contrast to the case of fixed degree polynomials when the range over which the coefficients are allowed to run increases. There it is known that almost all polynomials are not only irreducible, but also have S_n as their Galois group. For latest results and references on this topic, see [14].)

Our results are best illustrated by pictures of zeros. Figure 1 shows all zeros of the polynomials with coefficients 0, 1 of degrees $\leqslant 16$, and with constant term 1, except for the negative real zeros that are < -1.5. We show that W lies between the curves

$$C_1 = \left\{ z : |z| \leqslant 1, \frac{|z|}{1-|z|} = \left|\frac{2-z}{1-z}\right| \right\} \tag{1.3}$$

and

$$C_2 = \left\{ z : |z| \geqslant 1, \frac{1}{|z|-1} = \left|\frac{2z-1}{1-z}\right| \right\}. \tag{1.4}$$

The curve C_1 is mapped to C_2 by $z \to 1/z$. This mapping takes W to itself, since if $z \in W$, and z is a root of $f(z) \in P$ and $\deg f(z) = d$, then $1/z$ is a root of $z^d f(1/z) \in P$. We show that all $z \in W$ are enclosed strictly between C_1 and C_2. From this it follows that for all $z \in W$,

$$\frac{1}{\varphi} < |z| < \varphi , \tag{1.5}$$

where

$$\varphi = \frac{1 + 5^{1/2}}{2} \tag{1.6}$$

is the "golden ratio." (The bound (1.5) has been proved independently in different contexts by Flatto, Lagarias, and Poonen [13] and by Solomyak [16].) We also show that the line segment $[-\varphi, -\varphi^{-1}] \in \overline{W}$. However, $-\varphi \notin W$ and $-\varphi^{-1} \notin W$. Further, there is a constant $\delta > 0$ such that if

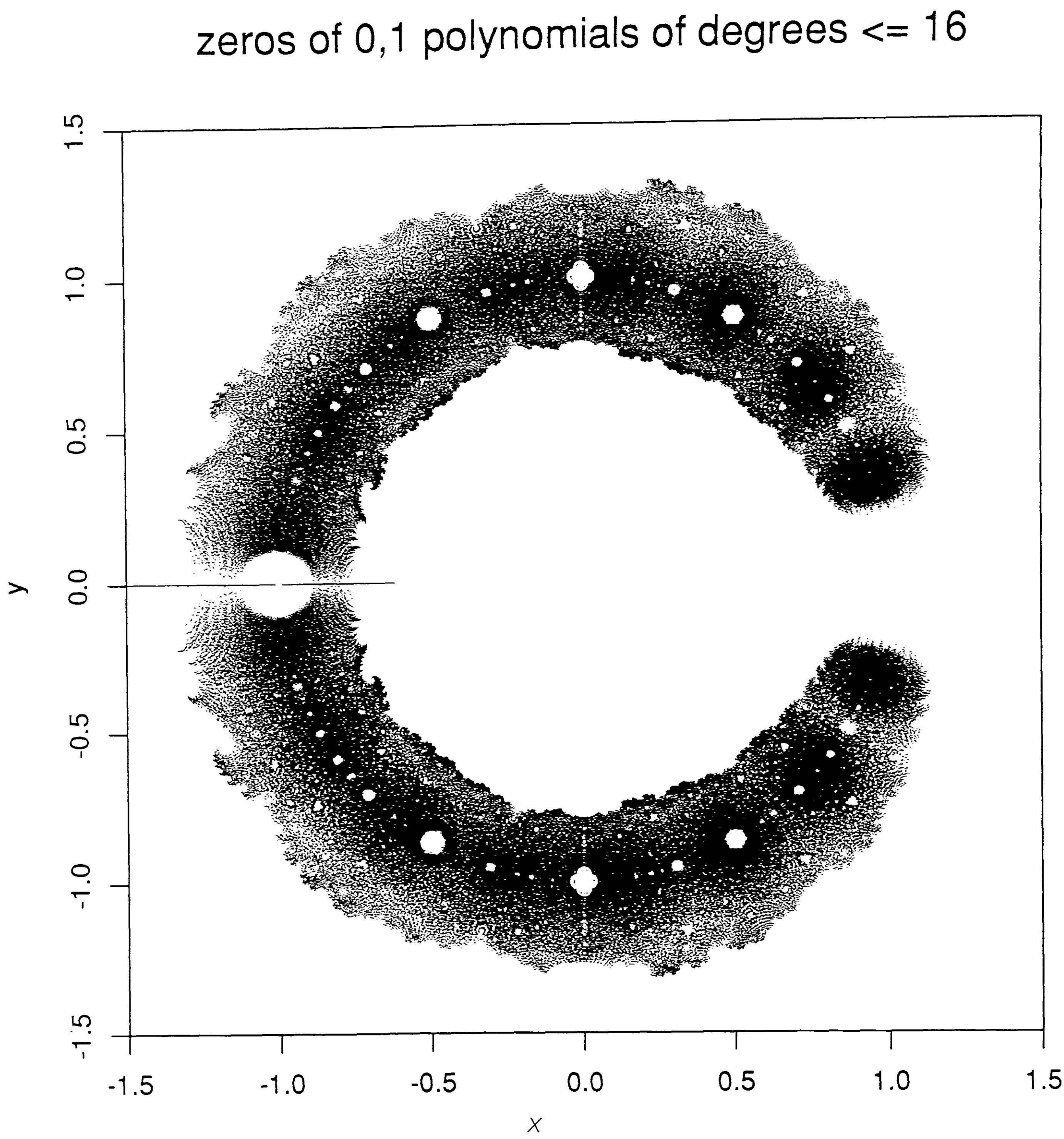

FIGURE 1

Scatterplot of zeros $z = x + iy$ of polynomials of degrees $\leqslant 16$
with constant term 1 and coefficients 0 and 1.
The "segments" along the negative real axis are created by negative real zeros.
Negative real zeros < -1.5 are not shown.

$z \in W$, $|z| \leqslant \varphi^{-1} + \delta$, then $z \in R$. Thus the "spike" along the negative real axis that is visible in Figure 1, connecting curves C_1 and C_2 with the exception of a small gap at -1, is due to zeros.

Since polynomials in P have nonnegative coefficients, $1 \notin W$. However, since $\zeta \in W$ for every root of unity $\zeta \neq 1$, $1 \in \overline{W}$, where $\overline{W}$ denotes

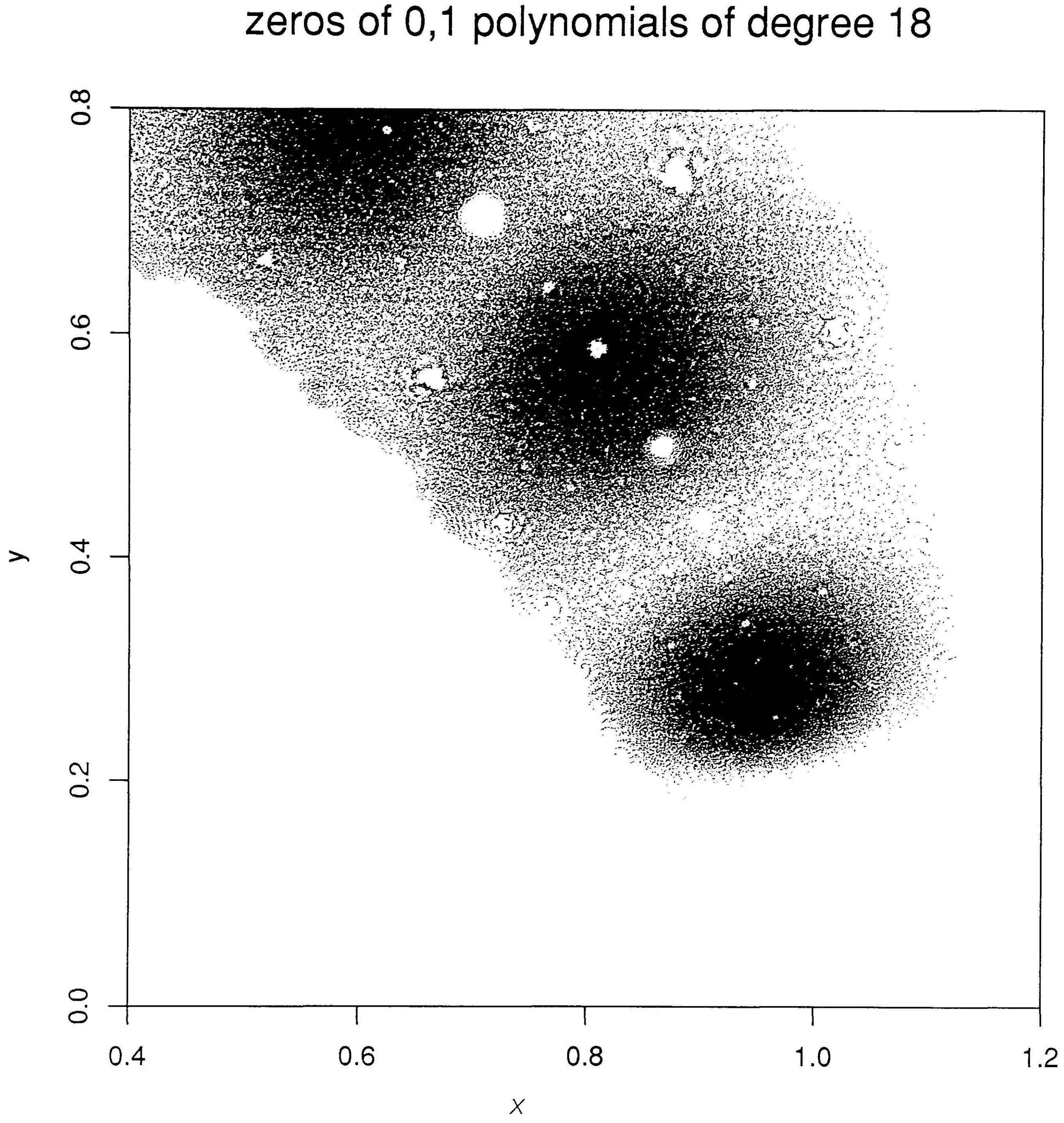

FIGURE 2

Scatterplot of zeros of polynomials of degree 18
with constant term 1 and coeffficients 0 and 1 that are near to $z = 1$.

the closure of W. We answer a question posed by J.H. Conway and Richard Parker about the behaviour of W near 1 by proving there exist points $z = x + iy \in W$ such that $0 < x - 1y = o(|y|)$, so that these points come in tangent to the x-axis.

Figure 2 shows the zeros of polynomials $f(x) \in P$ of degree 18 that are close to $z = 1$. Figure 3 shows zeros of polynomials $f(x) \in P$ of all degrees $\leqslant 32$ that fall in a certain small region of the complex plane. Figures 4 , 5

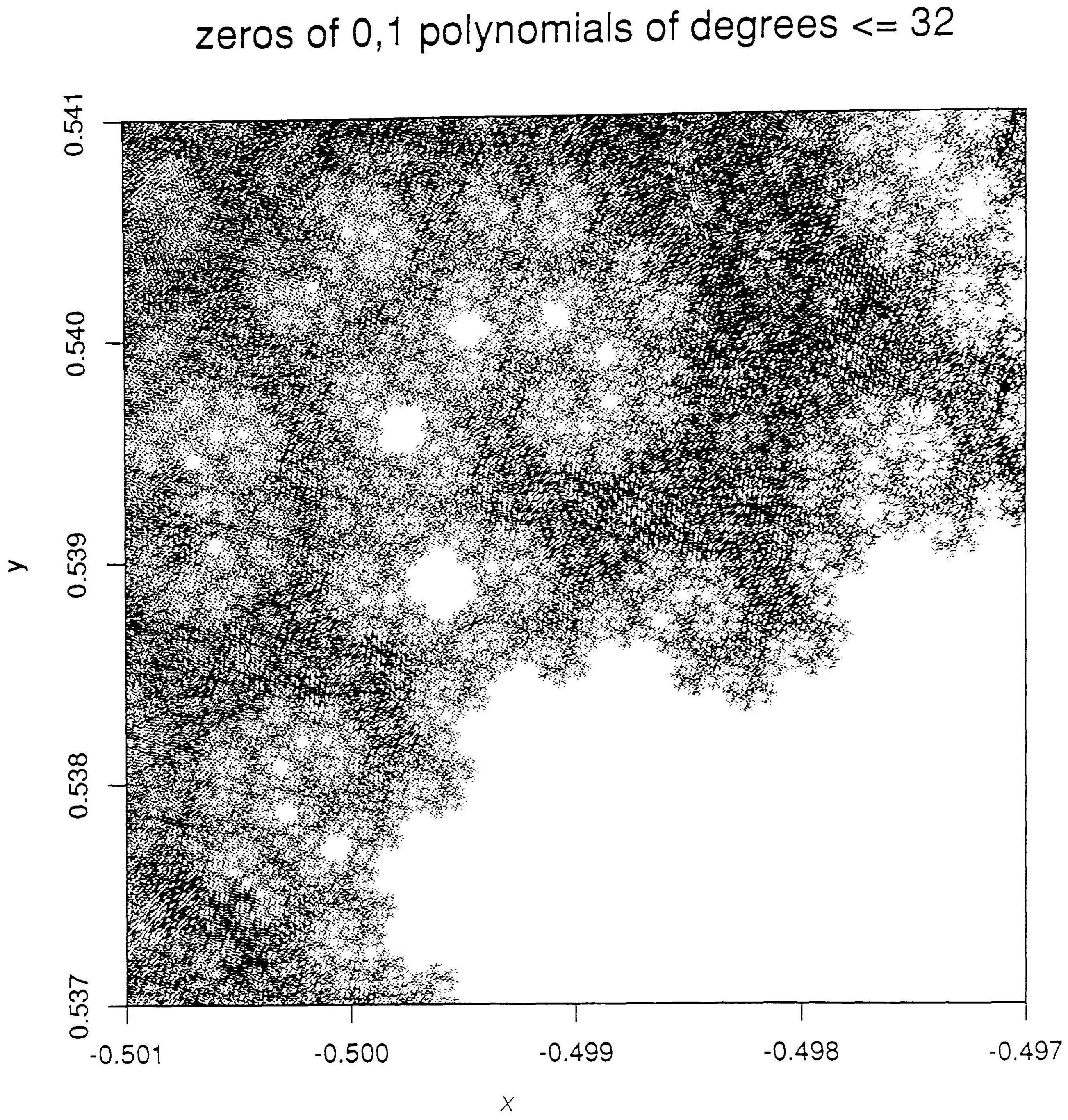

FIGURE 3

Scatterplot of zeros of polynomials of degrees $\leqslant 32$ with coefficients 0 and 1.

and 6 show pictures of parts of $\overline{W}$. The region depicted in Figure 4 is the same as that of Figure 3. Section 6 explains how these pictures were created.

Theorem 2.1 of Section 2, which says that W is contained between C_1 and C_2, is not best possible. The only points of $\overline{W}$ that are in $C_1 \cup C_2$ are $1, -\varphi, -\varphi^{-1}$. In Section 6 we will show how to obtain more precise bounds for W. However, because of the fractal nature of W, there is no simple description of its shape.

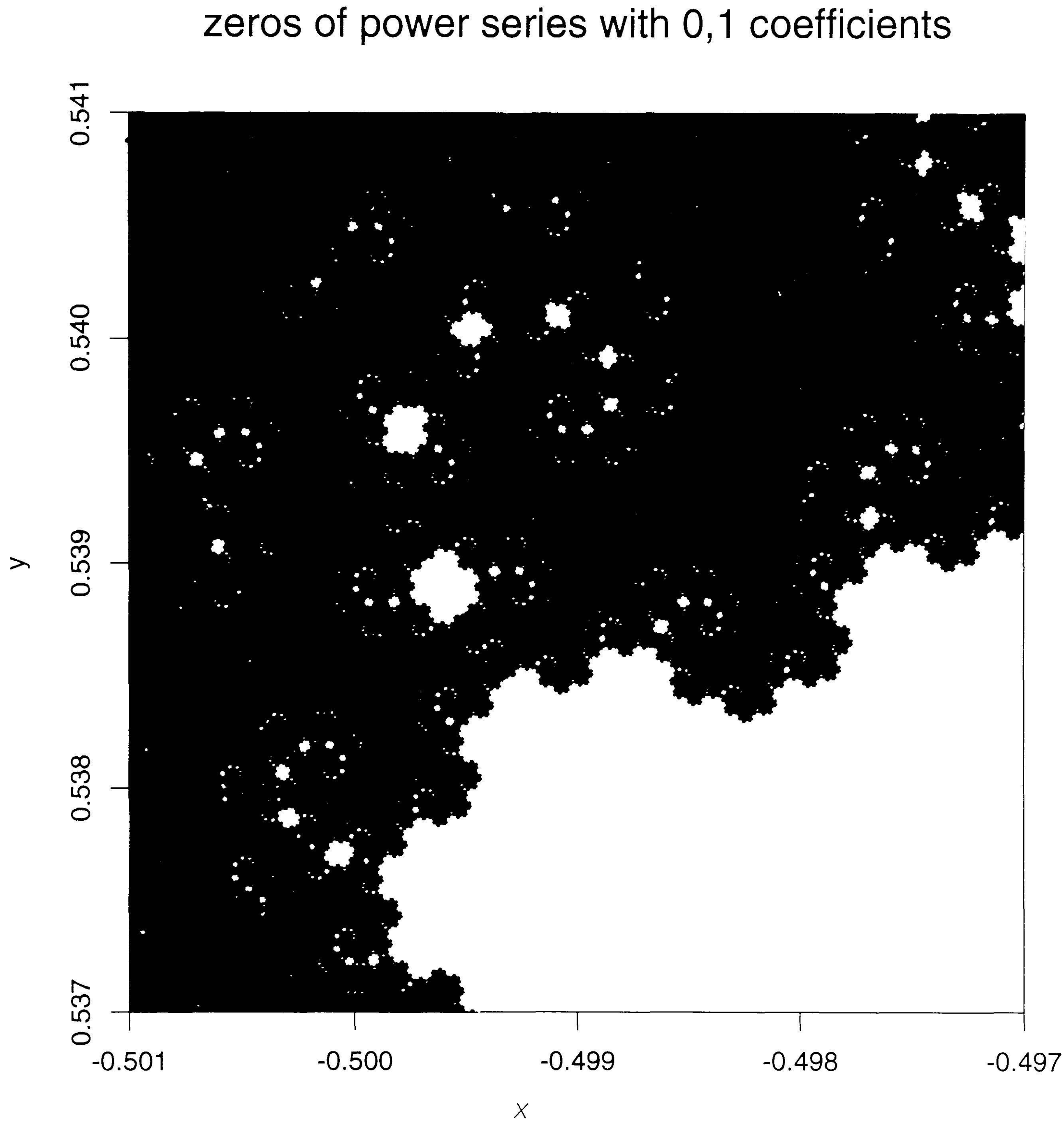

FIGURE 4

Section of $\bar{W}$. The same region, with points from W displayed, is shown in Figure 3. Black denotes points $z \in \bar{W}$.

Many features visible in the graphs can be explained (at least heuristically, and often rigorously) by using known results or methods. When one graphs zeros of any single polynomial with coefficients 0 and 1, most of them are close to the unit circle $|z| = 1$ and they are equidistributed in angles, so that the first quadrant, for example, has close to 1/4 of the total. This phenomenon is true for all polynomials whose coefficients do not vary much, as follows from results originating with Erdös and Turán [11]. For statements and references to general results, see [4].

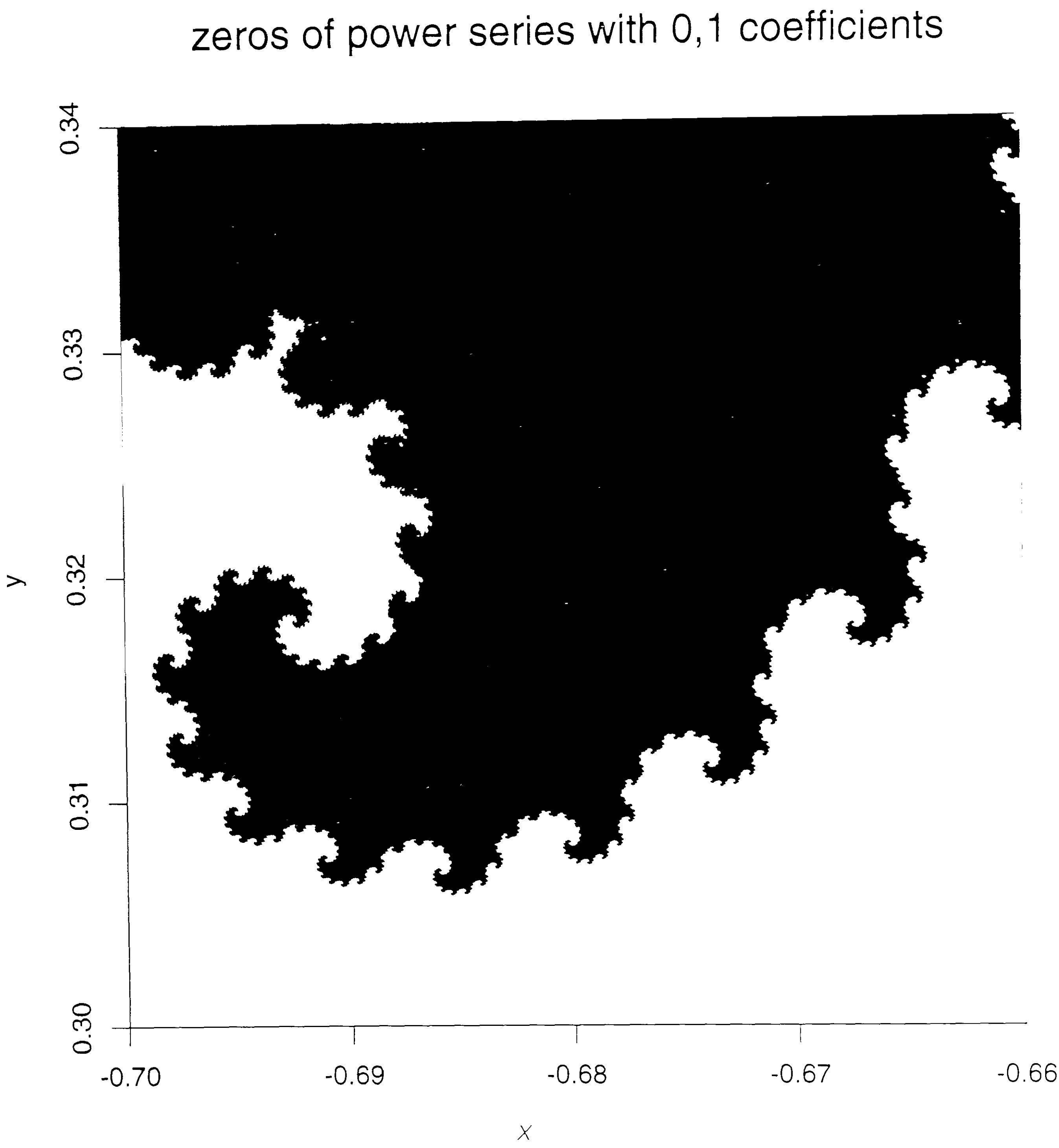

FIGURE 5

Section of $\overline{W}$, the set of zeros of power series with 0, 1 coefficients with black denoting $z \in \overline{W}$.

The expected number of real roots of a random polynomial (which have to be negative for $f(z) \in P$) grows logarithmically with n, as was first noted by Kac and Rice (see [4]). Furthermore, the variance is small.

In Figures 1 and 2, there is a perceptible clustering of zeros. This is a reflection of the "averaging phenomenon" for roots of random polynomials [4, 15], and again is not special to 0, 1 coefficients. The "average" of the polynomials of degree n that are in P is

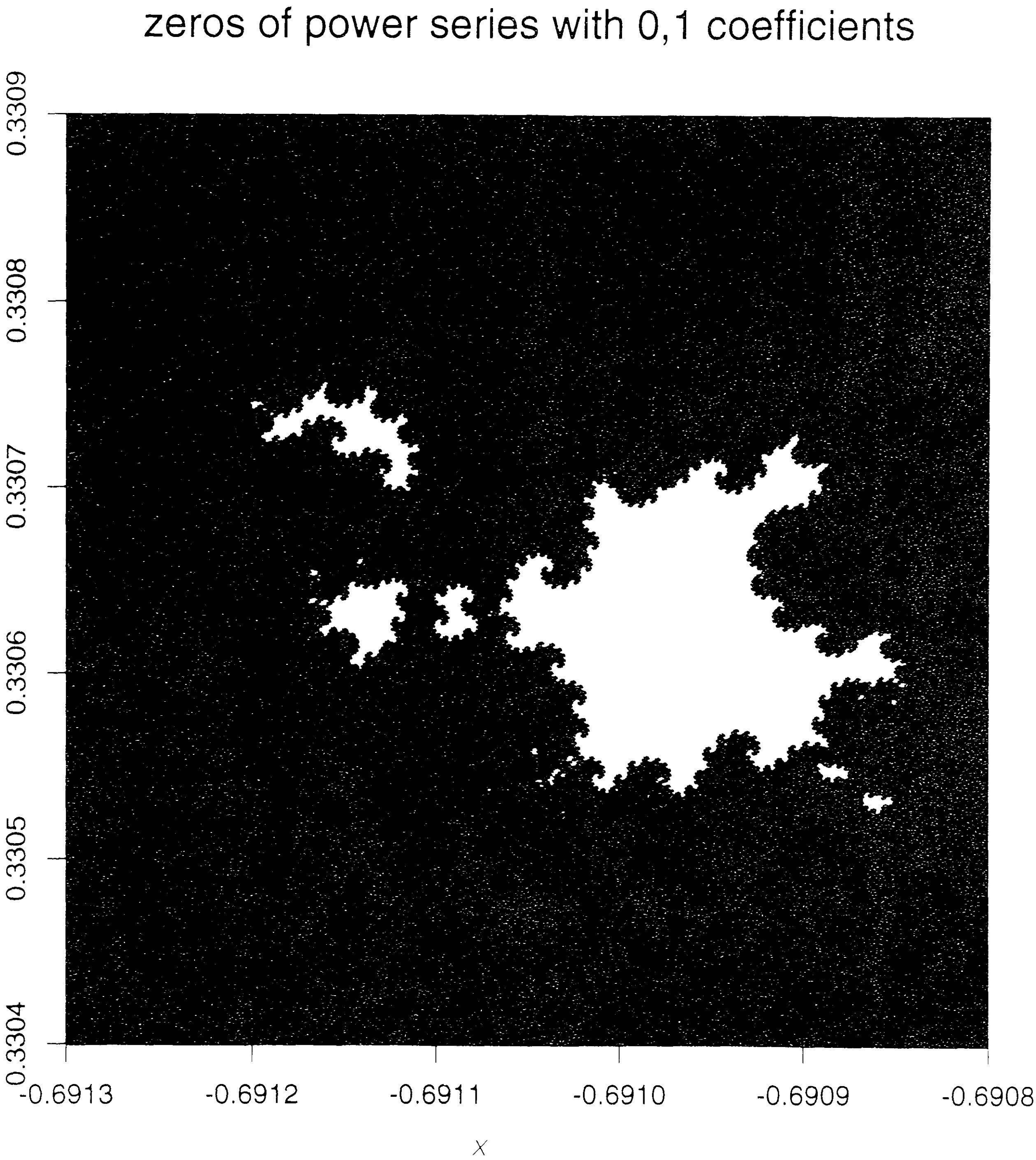

FIGURE 6

Section of $\overline{W}$. This is an enlargement by a factor of 80 of a section of Figure 5, showing some of the holes contained in $\overline{W}$.

$$g(z) = z^n + 1 + \frac{1}{2}\sum_{k=1}^{n-1} z^k = \frac{(1-2z)z^n - z + 2}{2(1-z)}, \tag{1.7}$$

and on average the zeros of $f(z) \in P$ tend to cluster near the zeros of $g(z)$.

Figures 1 and 2 show several large "holes," which contain either just one or no zeros. These holes are usually centered at algebraic integers α of low degree and small height (i.e., algebraic integers α that satisfy polynomial equations with small integral coefficients). The most prominent of the holes are at the roots of unity, such as -1 and i. As one computes zeros of polynomials $f(z) \in P$ of increasing degrees, the large holes in Figures 1 and 2 fill up. However, there are other holes, such as those visible in Figures 3-6, that are free of zeros even when the degree increases.

We show in Section 3 that there is an open neighborhood of $\{z : |z| = 1, z \neq 1\}$ that is in $\bar{W}$. In Section 4 we prove that $\bar{W}$ is connected. The more involved argument in Section 5 proves that $\bar{W}$ is path connected. Since the unit circle is contained in $\bar{W}$, but $0 \notin \bar{W}$, $\bar{W}$ is clearly not simply connected. Numerical experiments suggest that $\bar{W}$ has "holes" in it besides the big hole containing 0. (That is, $\mathbf{C} \setminus \bar{W}$ has more than 2 connected components.) In particular, the disk of radius 10^{-5} centered at $-0.69098 + 0.33062i$ appears to be part of such a hole. This hole and some neighboring ones are pictured in Figures 5 and 6. Other, even larger holes, can be seen in Figures 3 and 4.

$\bar{W}$ has a fractal appearance that is reminiscent of some of the Julia sets [1, 10]. In Section 6 we sketch arguments that explain how this arises. However, we do not have estimates for such interesting parameters as the Hausdorff dimension of the boundary of $\bar{W}$.

In contrast to our result that $\bar{W}$ is path connected, the Mandelbrot set is only known to be connected, although it is conjectured to be path connected [1, 10]. Our methods are simpler than those used to study the connectedness of the Mandelbrot set. They are similar to the techniques developed for investigating iterated function systems [1].

Results similar to those for polynomials with 0, 1 coefficients can also be obtained for other families of polynomials with a small set of possible coefficients. For example, for ± 1 coefficients, pictures of zeros are qualitatively similar to those of 0, 1 polynomials. There is symmetry about the imaginary axis as well as the real axis (corresponding to changing the variable z to $-z$). There are two "spikes" of zeros along the real axis that fill the intervals $[-2, -1/2]$ and $[1/2, 2]$, while there are no other zeros in $|z| \leqslant 1/2 + \delta$ or $|z| \geqslant 2 - \delta$ for some $\delta > 0$. For polynomials with cubic roots of unity as coefficients, there are no "spikes", but the zeros still have a fractal appearance.

The set

$$\bar{W} \cap \{z : |z| < 1\} \tag{1.8}$$

is the set of zeros of power series

$$f(z) = 1 + \sum_{k=1}^{\infty} a_k z^k, \quad a_k = 0 \text{ or } 1. \tag{1.9}$$

Since $z^{-1} \in W$ for all $z \in W$, it is sufficient to study $z \in W, |z| \leqslant 1$, and in some ways it is more natural to deal with the above power series.

Some of our methods and results are similar to those of Thierry Bousch [5, 6], whose work was brought to our attention by D. Zagier. The report [5] proves that the closure of the set of zeros of polynomials with coefficients 0, ± 1 is connected. The thesis [6] contains, along with a variety of other results, general methods for studying similar problems. In the area where our work overlaps [5, 6], we obtain a somewhat stronger result by proving path connectivity.

Boris Solomyak [16] has studied zeros of power series of the form (1.9), but with the c_k, $k \geqslant 1$, allowed to take any real values in the interval [0, 1]. He shows that the bound (2.4) holds there as well, and that there is a "spike" of real zeros along the negative real axis. However, the zeros of Solomyak's functions are substantially different from those we investigate. For example, he shows that segments of the boundary he investigates have everywhere dense sets of points where a tangent exists, as well as everywhere dense sets of points with no tangent. There are also no holes in Solomyak's set of zeros.

The paper of Brenti, Royle, and Wagner [7] discusses various properties of chromatic polynomials. While it is not directly related to our work, the numerical evidence it presents shows that zeros of chromatic polynomials may also exhibit fractal behavior. This may also be true for the partition function zeros of [3].

2. Bounds and Locations for Zeros

A polynomial $f(z) \in P$ can have multiple zeros. If $\zeta \neq 1$ is a d-th root of unity, then ζ is a zero of

$$g(z) = \sum_{j=0}^{d-1} z^j ,$$

and therefore a zero of $g(z^k)$ for any k such that $d \mid k - 1$. Hence it is a zero of multiplicity 2 for $g(z)\,g(z^k)$, a polynomial in P. Higher multiplicities can be obtained by iterating this procedure. On the other hand, we do not know whether any $z \in W$ that is not a root of unity can be a multiple root of any $f(z) \in P$. There do exist power series with coefficients 0, 1 that have double zeros z with $|z| < 1$, as will be shown in Section 3.

Inside a disk $\{z : |z| < r\}$ for $r < 1$, any polynomial $f(z) \in P$ can have only a bounded number of zeros. We prove a slightly more general result that will be used later on.