Prime Percolation

Ilan Vardi

CONTENTS

This paper examines the question of whether there is an unbounded walk of bounded step size along Gaussian primes. Percolation theory predicts that for a low enough density of random Gaussian integers no walk exists, which suggests that no such walk exists along prime numbers, since they have arbitrarily small density over large enough regions. In analogy with the Cramér conjecture, I construct a random model of Gaussian primes and show that an unbounded walk of step size $k\sqrt{\log|z|}$ at z exists with probability 1 if $k > \sqrt{2\pi\lambda_c}$, and does not exist with probability 1 if $k < \sqrt{2\pi\lambda_c}$, where $\lambda_c \approx 0.35$ is a constant in continuum percolation, and so conjecture that the critical step size for Gaussian primes is also $\sqrt{2\pi\lambda_c \log|z|}$.

1. INTRODUCTION

A problem that has attracted some recent attention [Gethner 1996; Gethner and Stark 1997; Gethner et al. 1998; Guy 1994, A16] is whether there is a walk to infinity of bounded step size along the Gaussian primes. This was first posed by Basil Gordon in 1962, and it was subsequently shown in [Jordan and Rabung 1970] that there is no walk of step length $\leq \sqrt{10}$ starting from the origin. In [Gethner et al. 1998], this has been improved to include walks of step length at most $\sqrt{26}$. Note that these results do not rule out a walk to infinity of such step sizes from some point far away from the origin.

I will examine the relationship between this problem and the theory of percolation. This theory deals with the similar problem of walking to infinity on a lattice, but where the sites are now given a fixed probability ρ of being included (see [Durrett 1988; Èfros 1982; Grimmett 1989; Stauffer and Aharony 1994] for an introduction). The main feature of percolation is that it exhibits phase transi-

1058-6458/1998 \$0.50 per page

tion, i.e., there is a *critical probability* $0 < \rho_c < 1$ for which an infinite path exists with probability 1 if $\rho > \rho_c$, and no infinite path exists with probability 1 if $\rho < \rho_c$.

The theory of percolation gives a heuristic reason for believing that there is no walk to infinity along Gaussian primes since the prime number theorem implies that the density of Gaussian primes is about $2/(\pi \log x)$ in a disk of radius x, so that the "probability" of a lattice point being prime becomes smaller than any ρ. This suggests the following conjecture:

Conjecture 1.1. *For any k, there is no infinite component of Gaussian primes connected by step size at most k.*

Note that the random model predicts that even if there is no infinite component, there are still arbitrarily large components (though their expected size is finite [Grimmett 1989]). It seems that Gaussian primes do not exhibit this behavior since they are restricted to certain congruence classes.

Conjecture 1.2. *For any k, there is a bound on the largest component of Gaussian primes connected by step size at most k.*

J. H. Jordan and J. R. Rabung [1976] proved Conjecture 1.2 for $k = \sqrt{2}$, while E. Gethner and H. Stark [1997] recently proved it for $k = 2$; see Section 7.

The percolation problem most relevant to this paper is the so-called "Poisson blob model" of continuum percolation: Consider a Poisson process of intensity λ on the plane, i.e., points $X_1, X_2, \dots$, are uniformly distributed in the plane with density λ, such that the probability of having exactly n points in area a is $e^{-a\lambda}(a\lambda)^n/n!$. Around each point draw a disk of radius one. What is the probability of there being an unbounded connected set of disks? It was shown in [Zuev and Sidorenko 1985] (see [Grimmett 1989, Section 10.5]) that there is a constant λ_c such that this will occur with probability one for $\lambda > \lambda_c$, and with probability zero for $\lambda < \lambda_c$. It is believed that λ_c is approximately 0.35 [Gawlinski and Stanley 1981; Hahn and Zwanzwig 1977; Domb 1972].

The first point is that this model approximates percolation on the integer lattice with step size k. To see this, note that walk of step size at most k asks for a connected set of "disks", i.e., translations of the set of lattice points distance at most k from the origin, and two such disks are connected when the center of each disk is contained in the other. This is the same condition as having "disks" of radius $k/2$ with the simpler condition that they merely be overlapping. Zuev and Sidorenko showed that the constants $\rho_c(k)$ for this problem satisfy $\lim_{l\to\infty}(k/2)^2\rho_c(k) = \lambda_c$. If primes are in some sense "random" with density as above, then this would imply that the first "moat" around the origin, i.e., the outside part of the boundary of the connected component starting near the origin (i.e., that does not lie inside the connected component) for step size k should lie at distance about $e^{k^2/(2\pi\lambda_c)}$. This seems consistent with the data of [Gethner et al. 1998]. In fact, scaling theory for percolation (see [de Gennes 1976; Grimmett 1989; Stauffer and Aharony 1994]) predicts that the phase transition $p \downarrow p_c$ exhibits scaling properties or "power laws". By analogy, the connected component of Gaussian primes around the origin should have holes whose radius grow in size according to a power law as the distance from the origin approaches $e^{k^2/(2\pi\lambda_c)}$, and that near this value the connected component around the origin should exhibit self-similarity, i.e., the "moat" should have fractal properties (pictures from [Gethner et al. 1998] indicate that the moat does have a complicated shape for $k = \sqrt{18}$). For distances greater than $e^{k^2/(2\pi\lambda_c)}$ there should only be small isolated clusters.

Since bounded step size does not seem to produce an unbounded walk, it is interesting to see what step size is necessary. As in Cramér's conjecture on the largest difference between two primes, one constructs a random model of the Gaussian primes. This is done by independently assigning the Gaussian integer $z = a + bi$, with $|z| > 2$, the probability $2/(\pi \log |z|)$ of being open. More precisely, if $z_1, \dots, z_n, z'_1, \dots, z'_m$, with $|z_j|, |z'_k| > 2$, are distinct Gaussian integers, then the probability of that all $z_1, \dots, z_n$ are open and all $z'_1, \dots, z'_m$ are closed is

$$\left(\prod_{j=1}^{n} \frac{2}{\pi \log |z_j|}\right)\left(\prod_{k=1}^{m}\left(1-\frac{2}{\pi \log |z'_k|}\right)\right).$$

Theorem 1.1. *Consider the Gaussian integers with the above probability model and consider walks of step size at most $k\sqrt{\log |z|}$ at z, where k is a constant. Then for $k < \sqrt{2\pi\lambda_c}$, with probability one, there is no unbounded open component, and for $k > \sqrt{2\pi\lambda_c}$, with probability one, there is an unbounded open component.*

This suggests another conjecture:

Conjecture 1.3. *Consider walks along the Gaussian primes with step size at most $k\sqrt{\log |z|}$ at the prime $z = a + bi$. For any $k < \sqrt{2\pi\lambda_c}$, there is no unbounded walk, and for any $k > \sqrt{2\pi\lambda_c}$, there is an unbounded walk.*

The problem of distribution of Gaussian primes has been studied by Hecke [1918; 1920], who showed that Gaussian primes are evenly distributed among sectors $\theta_1 < \arg z < \theta_2$. However, current methods are very far from approaching the questions posed here. For example, the existence of infinite walks as in Conjecture 1.3 implies that the difference between consecutive rational primes p_n and p_{n+1} is $O(\sqrt{p_n \log p_n})$, which is better than known bounds, even assuming the Riemann Hypothesis, which only gives $O(\sqrt{p_n}\log^2 p_n)$.

Some aspects of this problem not examined in this paper consist of the following:

(a) Sieve methods [Halberstam and Richert 1974; Rademacher 1923] can be employed to give upper bounds on the density of an infinite component connected by step size $\leq k$. In particular, these methods would show that the number of Gaussian integers that lie inside a circle of radius x and belong to an infinite component connected by step size $\leq k$ is $O(x^2/\log^A x)$ for any fixed A. Note that a bound of the form at least $O(x)$ is required to prove Conjecture 1.1, but it seems that sieve methods are unable to approach this.

(b) Clearly, this problem can be generalized to general representations of integers by quadratic forms, in particular, analogous results should hold for representations of integers as a sum of four squares, i.e., walks along quaternionic primes.

(c) Similar questions can be raised in other number theoretic contexts [Vardi $\geq$ 1998]. For example, consider the set of pairs of relatively prime integers in the plane connected if they are distance one apart: Does this set have a limiting density and if so, is it nonzero?

2. GAUSSIAN INTEGERS

Recall that the Gaussian integers are complex numbers $a + bi$, where a and b are integers. Elementary properties of the Gaussian integers are given in [Hardy and Wright 1979], and the one relevant here is that Gaussian integers have a unique prime factorization analogous to ordinary integers. The difference is that primes of the form $4k+3$ remain prime, while 2 and primes of the form $4k+1$ split into two prime factors: for example, $2 = (1+i) \times (1-i)$, $5 = (2+i)(2-i)$. In general, a Gaussian integer $a+bi$, with $a, b \neq 0$, is prime if and only if a^2+b^2 is a rational prime (a prime in the ordinary sense).

Now, let $\pi_1(x)$ be the number of ordinary primes of the form $4k+1$ that are less than or equal to x, and let $\pi_3(x)$ be the corresponding number for primes of the form $4k+3$. In a circle of radius x, there are exactly $8\pi_1(x^2)$ Gaussian primes corresponding to factorizations of ordinary primes of the form $4k+1$, since each $a+bi$ contributes $\pm a \pm bi$ and $\pm b \pm ai$, while there are $4\pi_3(x)$ Gaussian primes corresponding to rational primes of the form $4k+3$, since each rational prime p of this form also gives the primes $-p$ and $\pm ip$. The prime number theorem and the Chebotarev density theorem imply

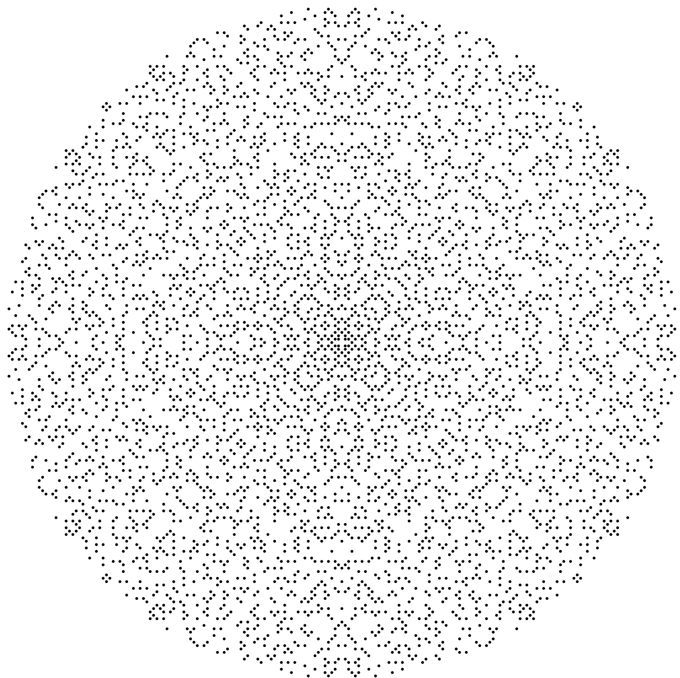

FIGURE 1. Gaussian primes at distance at most 100 from the origin.

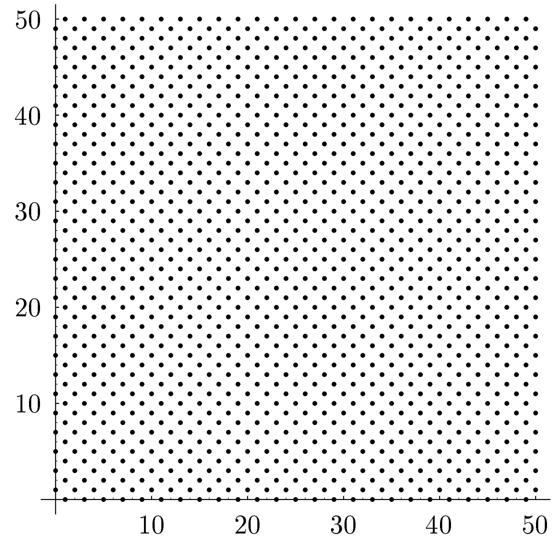

FIGURE 2. Odd Gaussian integers.

that $\pi_1(x) \sim \pi_3(x) \sim x/(2\log x)$, so the number of Gaussian primes in a circle of radius x is

$$4 + 8\pi_1(x^2) + 4\pi_3(x) \sim \frac{2x^2}{\log x} + \frac{2x}{\log x} \sim \frac{2x^2}{\log x}.$$

Since the number of Gaussian integers inside a circle of radius x is asymptotic to πx^2, this says that the density of Gaussian primes is asymptotic to $2/(\pi \log x)$.

Since all Gaussian primes except $\pm 1 \pm i$ must be of the form $a + bi$, where a and b have different parity, only Gaussian integers of this form, known as *odd* Gaussian integers, will be considered. Note that these form a lattice that is the same as the ordinary plane integer lattice when considered sideways (i.e., at a 45° angle) so that step size k along the ordinary lattice corresponds to step size $k/\sqrt{2}$ along odd Gaussian integers.

Next, it will be useful to know what the percentage sieved out is. Now if a prime q is of the form $4k + 3$, then q remains prime, so of the q^2 numbers $a + bi$, $0 \le a, b < q$, which I will identify with $[0, q-1] \times [0, q-1]$, only 0 is divisible by q. It follows that the numbers prime to q have density $1 - 1/q^2$. On the other hand, if p is of the form $4k + 1$, then there are $2p - 1$ numbers in $[0, p-1] \times [0, p-1]$ that have a common factor with p, namely $(0, 0)$, and $(a, \pm\iota a) \bmod p$, where $\iota^2 \equiv -1 \pmod p$ and $a = 1, 2, \ldots, p-1$. This means that the density of numbers prime to such a p is $1-(2p-1)/p^2 = (1-1/p)^2$. Another way to see this is to consider the factorization $p = a^2 + b^2 = (a+bi)(a-bi)$, where $a+bi$ and $a-bi$ are distinct Gaussian primes each generating an ideal of norm p, so there are $p-1$ numbers relatively prime to p in a fundamental domain mod $a \pm bi$ containing p Gaussian integers.

In general, given a rational integer $N = 2M$, where M is odd, the density of odd Gaussian integers relatively prime to N will be given by

$$\begin{aligned}\delta(M) &= \prod_{p \mid M} \delta(p) \\ &= \prod_{\substack{p \mid M \\ p \equiv 1 \ (\mathrm{mod}\, 4)}} \left(1 - \frac{1}{p}\right)^2 \prod_{\substack{q \mid M \\ q \equiv 3 \ (\mathrm{mod}\, 4)}} \left(1 - \frac{1}{q^2}\right).\end{aligned}$$

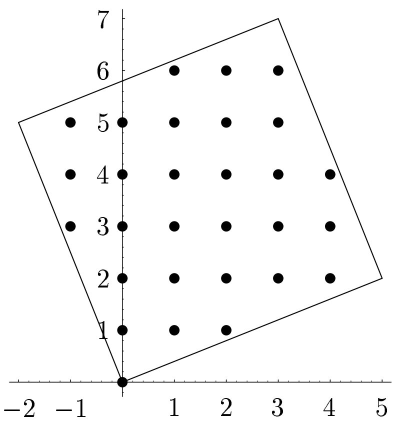

FIGURE 3. Fundamental domain for $5 + 2i$ of norm 29.

As a consequence of this, one sees that sieving by primes of the form $4k + 3$ is not very efficient, and in fact, sieving by *all* such primes only removes a finite fraction since

$$\alpha = \prod_{q \equiv 3 \pmod 4} \left(1 - \frac{1}{q^2}\right)$$
$$= 0.85610898172189347690603300614806117348 11 \ldots .$$

This value can be computed by relating it to $\beta = 1/\sqrt{2\alpha}$, the constant appearing in a closely related formula of Landau and Ramanujan which says that the number of integers $\leq x$ that can be written as a sum of two squares is asymptotic to $\beta x/\sqrt{\log x}$. In [Flajolet and Vardi 1996] it is shown that

$$\beta = \frac{1}{\sqrt{2}} \prod_{n=1}^{\infty} \left(\left(1 - \frac{1}{2^{2^n}}\right) \frac{\zeta(2^n)}{L(2^n)} \right)^{1/2^{n+1}}$$
$$= 0.76422365358922066299069873125009232811 68 \ldots ,$$

where $L(s) = 1 - 3^{-s} + 5^{-s} - 7^{-s} + \cdots$. This gives a fast algorithm for computing β due to the lacunary character of the product.

For the purposes of the next section, it will be necessary to compute the minimal density among integers $\leq N$, which is given by the following result:

Proposition 2.1. *The minimum value of $\delta(n)$, for n odd, $n \leq N$, is asymptotic to $8/(e^\gamma \pi \log\log N)$ and is taken on by*

$$M = \prod_{\substack{p \leq x \\ p \equiv 1 \bmod 4}} p \prod_{\substack{q \leq \sqrt{x/2} \\ q \equiv 3 \bmod 4}} q,$$

where $x \sim 2 \log N$.

Proof. Clearly the minimal density will occur for a product of consecutive primes $p \leq x$, $p \equiv 1 \pmod 4$ and consecutive primes $q \leq y$, $q \equiv 3 \pmod 4$. Solving for $(1-1/q^2) \leq (1-1/p)^2$, one sees that taking $y = \sqrt{x/2}$ is optimal.

Next, one notes that

$$\prod_{\substack{p \leq x \\ p \equiv 1 \bmod 4}} \left(1 - \frac{1}{p}\right)^2 \sim \frac{8}{e^\gamma \pi \alpha \log x}.$$

To see why this holds, let

$$A = \prod_{\substack{p \leq x \\ p \equiv 1 \bmod 4}} \left(1 - \frac{1}{p}\right), \qquad B = \prod_{\substack{q \leq x \\ q \equiv 3 \bmod 4}} \left(1 - \frac{1}{q}\right).$$

Merten's theorem [Hardy and Wright 1979] says that $AB \sim 2e^{-\gamma}/\log x$, while

$$\frac{A}{B} = \frac{\displaystyle\prod_{\substack{p \leq x \\ p \equiv 1 \bmod 4}} \left(1 - \frac{1}{p}\right) \prod_{\substack{q \leq x \\ q \equiv 3 \bmod 4}} \left(1 + \frac{1}{q}\right)}{\displaystyle\prod_{\substack{q \leq x \\ q \equiv 3 \bmod 4}} \left(1 - \frac{1}{q^2}\right)} \longrightarrow \frac{4}{\pi\alpha},$$

since the product in the numerator tends to $1/L(1)$, which equals $4/\pi$. Solving for A^2 gives the result.

Finally, as noted above, the prime number theorem (see [Davenport 1980]) and the Chebotarev density theorem imply that

$$\sum_{\substack{p \leq x \\ p \equiv 1 \bmod 4}} \log p \sim \frac{x}{2},$$

which shows that M is of order $e^{x/2}$. The proposition follows. $\square$

3. PERCOLATION

In this section I will examine several examples of percolation on the two dimensional lattice. The theory of percolation was introduced first developed mathematically by Broadbent and Hammersley [1957]. They modeled the probability that a fluid will seep through a solid substance that has wide and narrow channels with a given probability distribution — hence the name *percolation* — and proved the existence of a critical probability at which this occurs. This type of problem is best described by *bond percolation*, where edges of the square lattice are randomly included or excluded; this is the focus of [Grimmett 1989]. Note that what is being examined here is *site percolation.* Site percolation is more general in the sense that any bond problem can be restated as a site problem. In particular, I will consider the lattices $L(k)$ where the sites are pairs of integers (a, b) and two sites are considered connected if they are distance at most k apart. The simplest case is the ordinary integer lattice $L(1)$. Unlike the bond percolation problem, where the percolation constant for the two-dimensional integer lattice has been proved by Kesten to be $\frac{1}{2}$ (see [Grimmett 1989] for a proof of this difficult result), the exact value of the site percolation constant $\rho_c(1)$ for the plane lattice $L(1)$ is not known exactly but is believed to be approximately 0.5927 [Ziff and Sapoval 1986]. The best rigorous bounds are $\rho_c(1) > 0.556$ [van den Berg and Ermakov 1996] and $\rho_c(1) < 0.679492$ [Wierman 1995].

A key concept in the theory of site percolation is the *matching lattice* which can be described in terms of the game of Go. Recall that a set of Go stones form a live group if they are a connected set under the integer lattice $L(1)$ and that a group of Go stones is captured if every point on the boundary of the group is occupied by the opponent.

The outside set of capturing white stones can be thought of as being connected by the lattice $L(\sqrt{2})$, so this is exactly the matching lattice of the ordinary nearest neighbor lattice. Conversely, if the

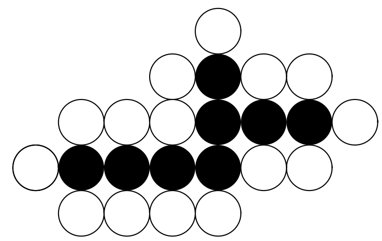

FIGURE 4. A connected set of black Go stones captured by white.

white stones are considered connected by this lattice, then they must be captured by a set of black stones that are connected by the ordinary lattice. Thus, the matching lattice of $L(\sqrt{2})$ is $L(1)$. It is seen that boundaries of connected sets in one lattice are described by connected sets in the matching lattice. Now Kesten [1982] has shown that if ρ_c and ρ_c^* are the percolation constants of a lattice and its matching lattice, then under certain symmetry conditions one has $\rho_c + \rho_c^* = 1$ (this was first proved for the site problem on the square lattice by Russo [1981]). It follows that $\rho_c(\sqrt{2}) = 1 - \rho_c(1) \approx 0.41$.

The general problem of walking along Gaussian integers with bounded step size corresponds to looking at the integer lattice where two sites are connected if they are at distance less than or equal to a fixed k. This problem was first examined by Gilbert [1961] and as $k \to \infty$, it corresponds to *continuum percolation,* and in particular to the so called "The Poisson blob model," because the disk centers can be thought of as a Poisson process of intensity λ, where λ represents the average density of a circle center per unit area. In particular, this means that a sequence of circles of radius 1 are to be placed in the plane at centers $X_1, X_2, \ldots$, with a Poisson distribution, i.e., the probability of n centers appearing in area a is $e^{-a\lambda}(a\lambda)^n/n!$. This implies that the percentage of area covered by disks is $1 - e^{-\pi\lambda}$; see [Hall 1988].

The percolation problem for the Poisson blob model was solved in [Zuev and Sidorenko 1985] (see [Grimmett 1989, Section 10.5; Meester and Roy 1996]) by considering it as a limit of the discrete

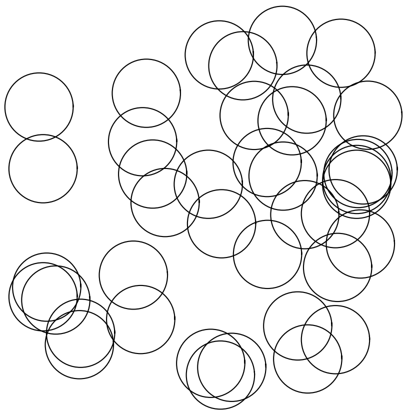

FIGURE 5. Poisson blob model of intensity $\lambda = 0.4$.

site percolation problem $L(k)$ since this problem is the same as drawing circles of radius $l = k/2$ around each lattice point and asking for an unbounded walk along overlapping circles. In [Zuev and Sidorenko 1985] it was shown that the continuous problem has percolation at intensity λ_c, where $(1 - \rho_c(k))^{(k/2)^2} \to e^{-\lambda_c}$. It follows that $\lambda_c = \lim(k/2)^2\rho_c(k)$. The exact value of λ_c is unknown, though it is believed to be about 0.35 (see [Hahn and Zwanzwig 1977; Domb 1972]). The best provable bounds are $0.174 < \lambda_c < 0.843$ [Hall 1985].

Continuum percolation is believed to satisfy the same properties as the ordinary lattice problem [Grimmett 1989; Meester and Roy 1996]. For example, it is known that if $\lambda > \lambda_c$, then with probability one, there is a unique infinite connected component. Furthermore, in [Zuev and Sidorenko 1985] it was shown that if λ_b is the smallest value for which the expected size of clusters is infinite for $\lambda > \lambda_b$, then $\lambda_b = \lambda_c$. Penrose [Penrose 1991] has also shown that for large intensities, i.e., $\lambda \to \infty$, the probability that a disk is isolated, given that it lies in a finite cluster, approaches 1. Penrose has also shown that the cluster density or "free energy" defined by

$$E(\lambda) = \lambda \sum_{n=1}^{\infty} \frac{\Pr\left(|C(0)| = n\right)}{n},$$

where $C(0)$ is the cluster containing zero is a continuous function of λ. It is also believed that continuum probability should satisfy the same conjectural properties as lattice problems, e.g., that there should not be an infinite cluster when $\lambda = \lambda_c$ (i.e., the percolation probability is continuous), and that as $\lambda \to \lambda_c$, power laws should hold (see [Grimmett 1989] [Stauffer and Aharony 1994]).

4. A RANDOM MODEL OF GAUSSIAN PRIMES

In analogy to Cramér's model of prime numbers ([Cramér 1937]; see also [Riesel 1985]), one can construct a model of Gaussian primes by considering each lattice point $z \in \mathbb{Z}^2$ with $|z| > 1$ to be "open" with probability $2/(\pi \log |z|)$ (so a nonopen lattice point will be "closed"). Using the exact same argument as Cramér, one can find the asymptotics of the largest expected "gap" of Gaussian primes:

Proposition 4.1. *Let $d(x)$ be the radius of the largest closed disk at distance at most x from the origin. Then, with probability one,* $\limsup d(x)/\log x = 1$.

Proof. (a) $\limsup d(x)/\log x \le 1$. Let $c > 1$ be a constant, and consider about each each Gaussian integer z where $|z| > 1$, a disk of radius $c \log |z|$. For each $\varepsilon > 0$ and sufficiently large z, each such disk contains at least $(1-\varepsilon)\pi c^2 \log^2 |z|$ Gaussian integers, so the probability that such a disk has no open points is at most

$$\begin{aligned}\left(1 - \frac{2}{\pi \log \left|z + c\log|z|\right|}\right)^{(1-\varepsilon)\,\pi c^2 \log^2 |z|} \\ \ll \left(1 - \frac{2}{\pi \log |z|}\right)^{(1-\varepsilon)\pi c^2 \log^2 |z|} \\ \ll |z|^{-2(1-\varepsilon)\,c^2}.\end{aligned}$$

Now summing over all such Gaussian integers gives an upper bound

$$\sum_{|z|>1} |z|^{-2(1-\varepsilon)\,c^2},$$

which converges for any $c > 1$, upon choosing small enough ε. The Borel–Cantelli Lemma shows that the probability of infinitely many such disks being closed is zero.

(b) $\limsup d(x)/\log x \geq 1$. Let $c < 1$, and for each Gaussian integer z with $|z| > 1$, consider a disk of radius $c \log|z|$ centered at $z \log^2|z|$. As in the previous part, it follows that for any $\varepsilon > 0$ and sufficiently large $|z|$, each disk has a probability at least $A\,|z|^{-2(1+\varepsilon)\,c^2}$ of being closed, where A is an absolute constant. Moreover, all of these events are independent, so the fact that the series

$$\sum_{|z|>1} |z|^{-2(1+\varepsilon)\,c^2} \log^b |z|$$

diverges for sufficiently small ε and any fixed b proves the result. □

Remark. Clearly, Proposition 4.1 can be generalized to other (noncircular) domains.

The general philosophy of this paper leads to the following conjecture:

Conjecture 4.1. *Let $d_p(x)$ be the radius of the largest prime free disk of distance at most x from the origin. Then* $\limsup d_p(x)/\log x = 1$.

The best result in this direction is due to Coleman [1990], who has shown that $d_p(x) = O_\varepsilon(x^{7/12+\varepsilon})$ for any $\varepsilon > 0$. Lower bounds of Rankin type on the size of prime free regions should be obtainable using methods for rational primes; see [Guy 1994, A8].

As in the final remarks of Section 3, one could conjecture that walks of size $r\sqrt{\log|z|}$ along Gaussian primes should satisfy all the properties (including conjectural ones) for continuum percolation. In some cases this requires normalization; for example, one defines the normalized cluster density, or "free energy", of the step size at most $k\sqrt{\log|z|}$ as

$$E_p(k) = \lim_{x\to\infty} \frac{\begin{pmatrix}\text{Number of clusters of Gaussian}\\ \text{primes in a circle of radius } x\end{pmatrix}}{2x/\big(k\sqrt{\log x}\big)}.$$

Conjecturally, this should equal the cluster density of the the Poisson blob of intensity $\lambda = k^2/(2\pi)$, i.e., $E_p(k) = E(k^2/(2\pi))$, where $E(\lambda)$ was defined in Section 3.

Note that the random model is taken to be a first order approximation of the behavior of Gaussian primes. A. Hildebrand and H. Maier [1989], and A. Granville [1995a; 1995b], have shown that in some cases rational primes exhibit behavior which is different from the one predicted by the random model.

5. PROOF OF THEOREM 1.1

The idea of the proof is that the percolation problem for the random model of Gaussian primes approximates the Poisson blob model. To see this, consider the map

$$f_s(z) = z/\big(s\sqrt{\log|z|}\,\big),$$

for $|z| > 1$. This map normalizes the Gaussian prime distribution in the sense that, under this map, the lattice points that appear with probability $2/(\pi \log|z|)$, approach a Poisson process of intensity $\lambda = 2s^2/\pi$. Furthermore, this map sends a disk with center z and radius $s\sqrt{\log|z|}$ to a disk with center $z/\big(s\sqrt{\log|z|}\,\big)$ and radius 1 in the sense that

$$f_s(z + s\sqrt{\log|z|}\,e^{i\theta}) = \frac{z}{s\sqrt{\log|z|}} + e^{i\theta} - \frac{z\,\cos(\theta - \arg z)}{2|z|\log|z|} + O\bigg(\frac{1}{|z|\sqrt{\log|z|}}\bigg).$$

Since continuum percolation occurs at $\lambda = \lambda_c$, this gives $2s_c^2/\pi = \lambda_c$ or $s_c = \sqrt{\pi\lambda_c/2}$. This predicts that the random prime model should have an infinite connected component of disks for $l = k/2 > \sqrt{\pi\lambda_c/2}$. Since the overlapping disk problem corresponds to walks of length the diameter of the disks, this implies that for $k > \sqrt{2\pi\lambda_c}$, there should be an unbounded walk of step size at most $k\sqrt{\log|z|}$.

Equivalently, the map $g_s(z) = s\sqrt{\log|z|}\,z$ is the "inverse" of $f_s(z)$ and takes a circle with center z

and radius 1 to a circle with center $s\sqrt{\log|z|}\,z$ and radius $s\sqrt{\log|z|}$. The map $g_s(z)$ increases areas by a factor of $s^2\log|z|$ as can be seen from the formula

$$f_s(z+\varepsilon) = sz\sqrt{\log|z|}+s\varepsilon\sqrt{\log|z|}+O\left(\frac{1}{\sqrt{\log|z|}}\right),$$

where ε is a complex number of absolute value ≤ 1. To the Poisson process of intensity λ will be associated a probability distribution on the Gaussian integers through the map $z \mapsto \lfloor sz\sqrt{\log|z|}\rfloor$, where $\lfloor x+iy\rfloor = \lfloor x\rfloor + i\lfloor y\rfloor$ ($\lfloor x\rfloor$ is the greatest integer less than or equal x) and a Gaussian integer $a+bi$ is considered "open" if there is a point of the Poisson process that maps to it. From the above, it is seen that the probability that the Gaussian integer z is open is

$$\frac{\lambda}{s^2\log|z|} + O\left(\frac{1}{\log^{3/2}|z|}\right).$$

To prove the result, one has to show that for $s > \sqrt{\pi\lambda_c/2}$, with probability one, there is an infinite connected component by step size $s\sqrt{\log|z|}$ in the random model of Gaussian primes, and that for $s < \sqrt{\pi\lambda_c/2}$, with probability one, there is no such infinite component.

(a) Assume $s > \sqrt{\pi\lambda_c/2}$. Let $\lambda = 2s^2/\pi$, then $\lambda > \lambda_c$ so there is a λ_1 such that $\lambda > \lambda_1 > \lambda_c$ and consider the Poisson blob model of intensity λ_1. Since $\lambda_1 > \lambda_c$ there exists an infinite component of disks of radius one. Let $s_1 = \sqrt{\pi\lambda_1/2}$, and consider the map f_{s_1} on the disks and the probability distribution on the Gaussian integers generated by $z \mapsto \lfloor s_1 z\sqrt{\log|z|}\rfloor$. By the above discussion it is seen that the disks of the Poisson model will have radius asymptotic to $s_1\sqrt{\log|z|}$ at center $\lfloor s_1\sqrt{\log|z|}\rfloor$, so that if two disks intersect in the Poisson model, then they are at distance at most $\sqrt{2}+o(1)$ when mapped to the random prime model. Therefore replacing each such disk with radius $s\sqrt{\log|z|}$ at $\lfloor s_1 z\sqrt{\log|z|}\rfloor$ will produce disks which definitely overlap if they overlapped in the Poisson blob model (for large $|z|$). Also, by the above, the probability of z being open is asymptotic to

$$\lambda/(s_1^2\log|z|) \sim s^2/(s_1^2\log|z|) < 2/(\pi\log|z|)$$

for large enough $|z|$, so this implies that, with probability one, there is an infinite connected component in the Gaussian prime model.

(b) Assume $s < \sqrt{\pi\lambda_c/2}$, and assume that with probability one there is an infinite connected component in the Gaussian prime model for this s. Let $s < s_1 < \sqrt{\pi\lambda_c/2}$. Consider the Poisson blob model of intensity $\lambda_1 = 2s_1^2/\pi < \lambda_c$. This does not exhibit percolation, but the Gaussian prime model does for step size $s_1\sqrt{\log|z|}$ since $s_1 > s$. The map $f_s(z)$ takes circles of radius radius $s_1\sqrt{\log|z|}$ into circles of radius $s_1/s > 1$ so that these intersect if and only if the circles of radius $s_1\sqrt{\log|z|}$ intersect. Since the probability distribution on the Gaussian prime model resulting from the map f_{s_1} is asymptotic to $2s_1^2/(s^2\pi\log|z|) > 2/(\pi\log|z|)$ for sufficiently large $|z|$, this gives a contradiction, proving the theorem. □

6. WALKS RELATIVELY PRIME TO AN INTEGER

The walks considered here are of Gaussian integers relatively prime to a given N, which reduces the question of infinite components to examining the finite set of Gaussian integers modulo N. One can think of integers modulo N as being in a fundamental domain $[0, N-1] \times [0, N-1]$, i.e., a big square. The integers relatively prime to N in this square have reflection symmetries generated by $(a,b) \mapsto (-a,b)$ and $(a,b) \mapsto (b,a)$, corresponding to reflection the vertical line $x = N/2$ and reflection about the 45 degree line $x = y$. I will also assume that N is even so that there is an extra reflection about the line $x + y = N/2$ given by the map $(a,b) \mapsto (N/2-b, N/2-a)$. These generate 16 reflections that break up the fundamental domain into 16 triangles each of which is a reflection of its neighbors along an adjacent side. A fundamental domain for these reflections can be given by the set

$$F(N) = \{(a, b) : a \geq b,\ a \leq N/2,\ a + b \leq N/2\}.$$

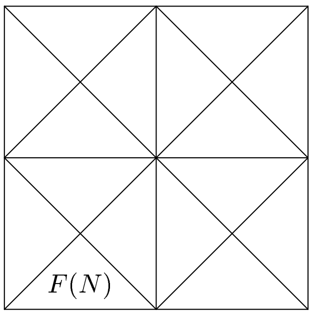

FIGURE 6. Reflection symmetries modulo N.

Since the fundamental square tiles the plane by translations by $(N, 0)$ and $(0, N)$, the following results are clear:

Proposition 6.1. *There is a walk to infinity along Gaussian integers relatively prime to N if and only if there is a path inside the triangle $F(N)$ that touches all 3 edges of the triangle.*

Proposition 6.2. *If there is no walk to infinity along Gaussian integers relatively prime to N, then there is an upper bound on the largest connected component, so Conjecture 1.2 holds.*

Proposition 6.3. *There is at most one infinite connected component of Gaussian integers relatively prime to N.*

The next question is how well different $2M$, for M odd, model percolation. In general, one should expect this to hold:

Conjecture 6.1. (i) *Set $D(M) = \big|\{M < x : \delta(M) < \rho_c(k/\sqrt{2})$ and there is a walk to infinity relatively prime to $2M\}\big|$. Then*

$$\lim_{x\to\infty} \frac{D(M)}{\big|\{M < x : \delta(M) < \rho_c(k/\sqrt{2})\}\big|} = 0.$$

(ii) *Set $G(M) = \big|\{M < x : \delta(M) > \rho_c(k/\sqrt{2})$ and there is no walk to infinity relatively prime to $2M\}\big|$. Then*

$$\lim_{x\to\infty} \frac{G(M)}{\big|\{N < x : \delta(M) > \rho_c(k/\sqrt{2})\}\big|} = 0.$$

(Note that step size k corresponds to step size $k/\sqrt{2}$ along odd Gaussian integers.)

This approach reduces a question of infinite walks to a finite problem, and in particular, only a single N needs to be found to prove Conjecture 1.2, and so Conjecture 1.1 (so Conjecture 6.1 also implies Conjecture 1.2).

This gives a method for checking Conjecture 1.1: Show that there are no unbounded walks of step size k by showing that there is no walk modulo some large N of low density.

According to Conjecture 6.1, the smallest value of N which will work is the smallest $N = 2M$ for which $\delta(M) < \rho_c(k/\sqrt{2})$. By Theorem 2.1,

$$8/(e^{\gamma}\pi \log\log N) \approx \rho_c(k/\sqrt{2}) \sim 8\lambda_c/k^2,$$

so the value of N is of order $\exp\big(\exp(k^2/(e^{\gamma}\pi\lambda_c))\big)$. In other words, the growth of N is doubly exponential, and therefore so is the above algorithm. It follows that this method will not be computationally feasible except for very small values of k.

7. WALKS OF STEP SIZE $\sqrt{2}$

I will illustrate the ideas of the previous section in the case of walks of length $\sqrt{2}$. As noted above, such walks can be considered as walks on the ordinary 2–dimensional integer lattice, which has percolation at $\rho_c(1) \approx 0.59$ (in fact, with probability one, there is no walk at $\rho_c(1)$ either [Russo 1981]). One should therefore expect that percolation for numbers relatively prime to $2M$, where M is odd and squarefree, should occur when $\delta(M)$ is approximately $\rho_c(1)$.

In fact, let $N = 2M = 130 = 2 \cdot 5 \cdot 13$, then $\delta(M) = 0.545325\ldots$, which is less than 0.59. By the above, checking to see if there is an unbounded path reduces to checking to see if there is a path in the triangular region $F(65)$ that touches all three sides. To see this graphically, one draws a filled disk of radius $1/\sqrt{2}$ around every Gaussian integer relatively prime to $130 = 2\cdot5\cdot13$, and then visually checks to see if there is a connected set of disks touching all three sides.

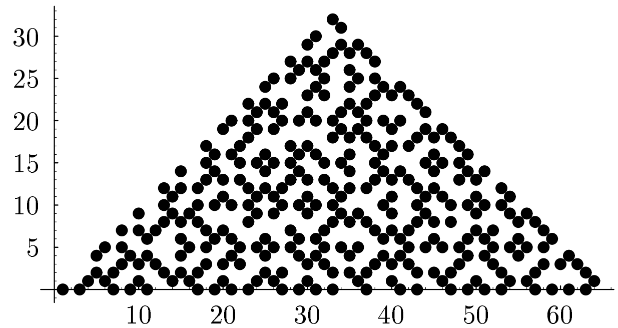

FIGURE 7. There is no unbounded walk prime to $2 \cdot 5 \cdot 13$ of step size $\sqrt{2}$.

Since no such path exists, this proves the following result (which actually follows from a stronger result of Jordan and Rabung [1976]; see below).

Theorem 7.1 [Gethner and Stark 1997]. *There is no unbounded walk of step length* $\sqrt{2}$.

As noted above, this shows that Conjecture 1.2 is true for step size $\sqrt{2}$, so there is an upper bound on the size of connected components of Gaussian primes with step size $\sqrt{2}$ and one can ask the question: "What is the largest connected component of primes that occurs infinitely often?" In the ordinary case, Hardy and Littlewood [Hardy and Littlewood 1922] made the k–tuples prime conjecture that any *admissible* k–tuple of numbers, i.e., tuples that do not have a congruence obstruction (e.g., n, $n+2$, $n+4$ has an obstruction mod 3), should occur infinitely often (see [Riesel 1985]). So analogously, one can look for admissible connected sets of Gaussian integers, and the conjecture would be that, infinitely often, each of these sets consists only of Gaussian primes.

This same question was examined by Jordan and Rabung [1976] (see also [Holben and Jordan 1968]), who showed that the largest admissible connected component has size 48 (note that this implies Theorem 7.1).

Theorem 7.2 [Jordan and Rabung 1976]. *The largest admissible* $\sqrt{2}$*-connected component has size* 48. *It is the result of removing one A and one B in Figure* 8 *in such a way that the component remains connected.*

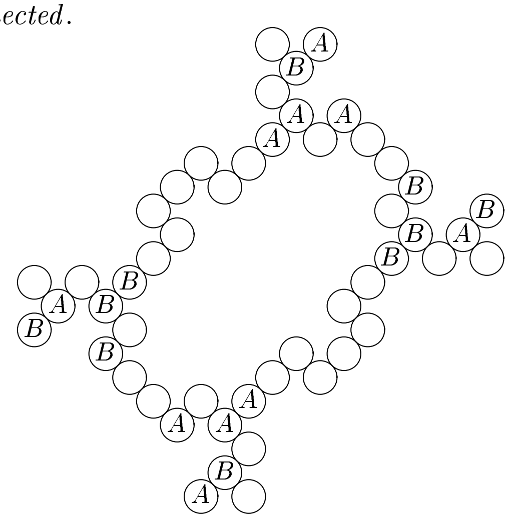

FIGURE 8. The largest admissible connected components.

I will give a brief description of the proof of this result, since this is not provided in [Jordan and Rabung 1976]: A necessary and sufficient condition for a set to be admissible is that for every Gaussian prime p there is a residue class not represented by the set. This means that one only has to check this for complex primes of norm less than or equal to the size of the set, and for primes of the form $4k+3$ less than the square root of the size of the set.

An upper bound for the number of elements in this set is provided by the proof of the Theorem 7.1. Inspection of the diagram shows that the largest connected component has 580 elements. One can do better by looking at numbers prime to $390 = 2\cdot3\cdot5\cdot13$, and this gives a largest connected component of size 71.

Figure 9 gives rise to 506 connected components of sizes 1 to 21 inclusive, and 25, 27, 28, 29, 32, 37,

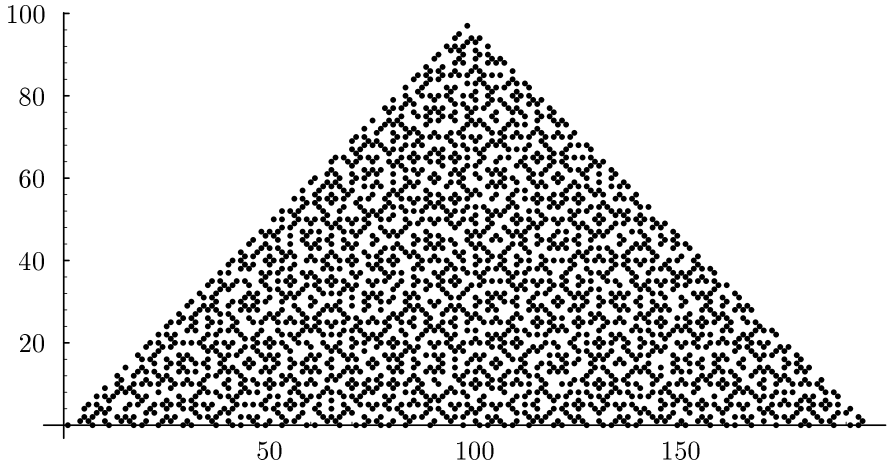

FIGURE 9. Walks relatively prime to $2 \cdot 3 \cdot 5 \cdot 13$.

50, 51, 71. The only ones that need to be checked are those of size $50, 51, 71$.

The component of size 50 is given in Figure 8 and contains $95+88i$ and $100+107i$. It only has problems, i.e., repeated values, modulo 29. As noted above, it can be made into a maximal admissible set of size 48 by removing a circle labeled A and a circle labeled B in such a way as to keep the picture connected. This results in 6 different maximal admissible sets, up to symmetry.

The next case is the component of size 51 containing $70+23i$ and $97+18i$, shown in Figure 10. It passes all tests except modulo 17. To analyze this case, two graphs are generated, one for $4+i$ and $4-i$, respectively, with the numbers $1, \dots, 17$, representing the 17 possible values modulo $4 \pm i$. One has to remove elements to leave one residue class free, and it is seen that the maximal connected subset that works is to remove the elements marked "14" in the top graph and the elements marked "5" in the bottom graph, leaving a maximal connected component of size 43. Removing "14" and "17" in the bottom graph yield maximal components of size 42 and and 41, respectively. One can also get maximal components of size 39.

The final case is the component of length 71 containing $35 \pm 22i$; see Figure 11. This component also passes every test except modulo 17. As in the previous case, the graph is given modulo $4+i$ and $4-i$, respectively, each number representing a distinct residue class modulo $4 \pm i$. A computer search shows that the largest admissible connected component has size 45, which is achieved by deleting points "2" in the $4+i$ case and "4" in the $4-i$ case.

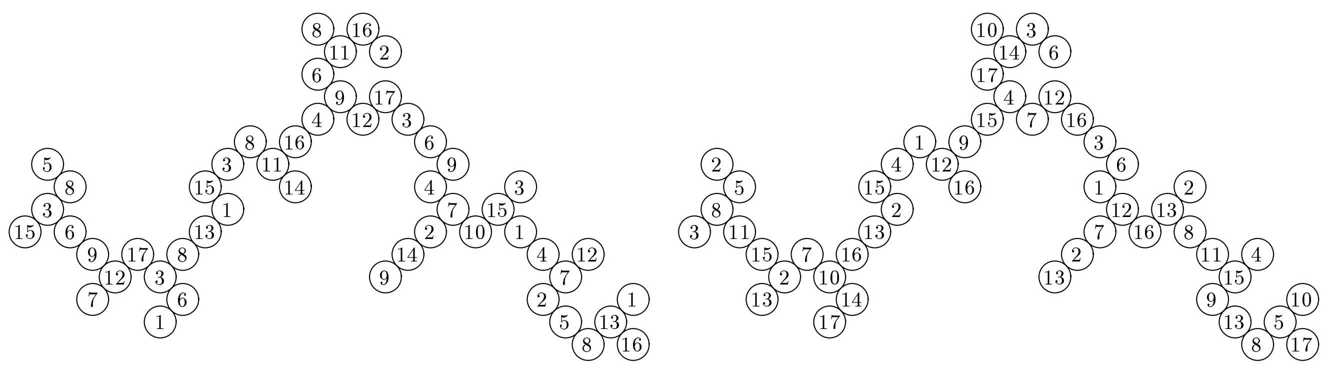

FIGURE 10. Left: Component of size 51 modulo $4+i$. Right: Component of size 51 modulo $4-i$.

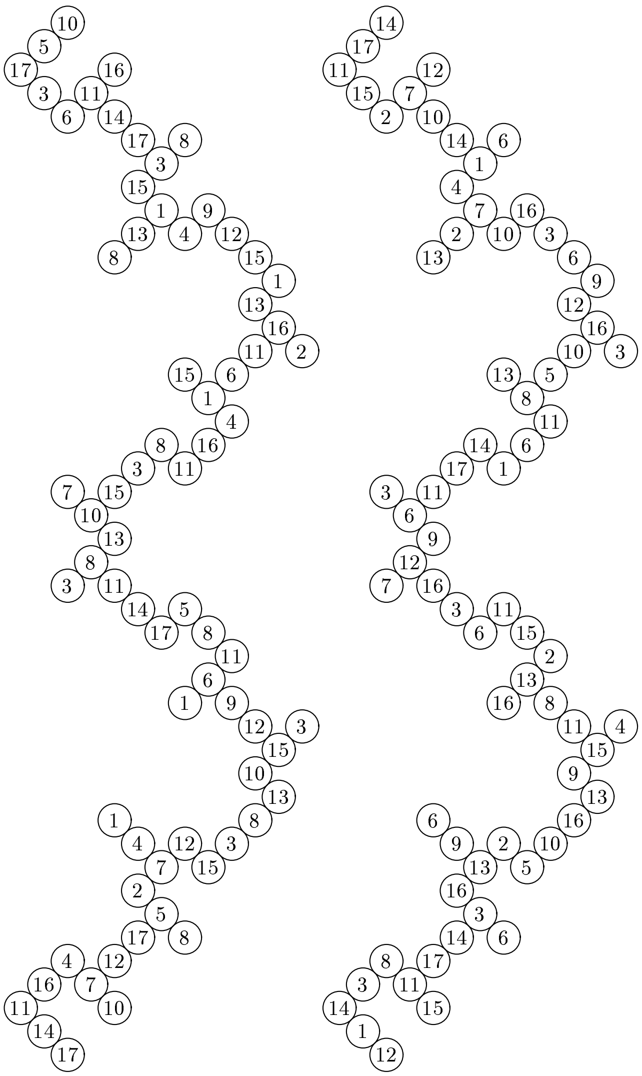

FIGURE 11. Left: Component of size 71 modulo $4+i$. Right: Component of size 71 modulo $4-i$.

Remark. Jordan and Rabung conjectured that such connected sets of Gaussian primes occur infinitely often. Following Hardy and Littlewood [Hardy and Littlewood 1922], one can further conjecture that for each such connected set, there is an asymptotic formula for the number of times it appears in a disk of radius R, as $R \to \infty$.

Remark. Gethner and Stark also showed that there is no walk of step size 2 (and also proving Conjecture 1.2 in this case). Note that with respect to odd Gaussian integers this is step size $\leq \sqrt{2}$ (king moves) and corresponds to the lattice $L(\sqrt{2})$. Since $L(\sqrt{2})$ is the matching lattice for $L(1)$, one has $\rho_c(\sqrt{2}) = 1 - \rho_c(1) \approx 0.41$ in this case.

REFERENCES

[van den Berg and Ermakov 1996] J. van den Berg and A. Ermakov, "A new lower bound for the critical probability of site percolation on the square lattice", *Random Structures Algorithms* **8**:3 (1996), 199–212.

[Broadbent and Hammersley 1957] S. R. Broadbent and J. M. Hammersley, "Percolation processes. I. Crystals and mazes", *Proc. Cambridge Philos. Soc.* **53** (1957), 629–641.

[Coleman 1990] M. D. Coleman, "The distribution of points at which binary quadratic forms are prime", *Proc. London Math. Soc.* (3) **61**:3 (1990), 433–456.

[Cramér 1937] H. Cramér, "On the order of magnitude of the difference between consecutive prime numbers", *Acta Arithmetica* **2** (1937), 23–46.

[Davenport 1980] H. Davenport, *Multiplicative number theory*, 2nd ed., Graduate Texts in Mathematics **74**, Springer, New York, 1980. Revised by Hugh L. Montgomery.

[Domb 1972] C. Domb, "A note on the series expansion method for clustering problems", *Biometrika* **59** (1972), 209–211.

[Durrett 1988] R. Durrett, "Crabgrass, measles, and gypsy moths: An introduction to modern probability", *Bull. Amer. Math. Soc.* **18** (1988), 117–143.

[Èfros 1982] A. L. Èfros, Физика и геометрия беспорядка, Nauka, Moscow, 1982. Translated as *Physics and geometry of disorder: percolation theory*, Mir, Moscow, 1986.

[Flajolet and Vardi 1996] P. Flajolet and I. Vardi, "Zeta expansions of classical constants", preprint, 1996. See www.cco.caltech.edu/~ilan/papers/zeta˙constants.ps.

[Gawlinski and Stanley 1981] E. T. Gawlinski and H. E. Stanley, "Continuum percolation in two dimensions: Monte Carlo tests of scaling and universality for noninteracting discs", *J. Phys. A* **14**:8 (1981), L291–L299.

[de Gennes 1976] P. G. de Gennes, “La percolation: un concept unificateur”, *La Recherche* **7** (1976), 919–927.

[Gethner 1996] E. Gethner, “In prime territory”, *Math. Horizons* (April 1996), 8–13.

[Gethner and Stark 1997] E. Gethner and H. M. Stark, “Periodic Gaussian moats”, *Experiment. Math.* **6**:4 (1997), 289–292.

[Gethner et al. 1998] E. Gethner, S. Wagon, and B. Wick, “A stroll through the Gaussian primes”, *Amer. Math. Monthly* **105**:4 (1998), 327–337.

[Gilbert 1961] E. N. Gilbert, “Random plane networks”, *J. Soc. Indust. Appl. Math.* **9** (1961), 533–543.

[Granville 1995a] A. Granville, “Harald Cramer and the distribution of prime numbers”, *Scand. Actuar. J.* (1995), 12–28.

[Granville 1995b] A. Granville, “Unexpected irregularites in the distribution of prime numbers”, pp. 388–399 in *Proceedings of the International Congress of Mathematicians* (Zurich, 1994), edited by S. D. Chatterji, Birkhäuser, Basel, 1995.

[Grimmett 1989] G. Grimmett, *Percolation*, Springer, New York, 1989.

[Guy 1994] R. K. Guy, *Unsolved problems in number theory*, 2nd ed., Problem Books in Mathematics, Springer, New York, 1994.

[Hahn and Zwanzwig 1977] S. W. Hahn and R. Zwanzwig, “Series expansions in a continuum percolation problem”, *J. Phys. A* **10** (1977), 1547–1555.

[Halberstam and Richert 1974] H. Halberstam and H.-E. Richert, *Sieve methods*, London Mathematical Society Monographs **4**, Academic Press, London and New York, 1974.

[Hall 1985] P. Hall, “On continuum percolation”, *Ann. Probab.* **13**:4 (1985), 1250–1266.

[Hall 1988] P. Hall, *Introduction to the theory of coverage processes*, Wiley, New York, 1988.

[Hardy and Littlewood 1922] G. H. Hardy and J. E. Littlewood, “Some problems of ‘Partitio Numerorum’ III: On the expression of a number as a sum of primes”, *Acta Math.* **44** (1922), 1–70. Reprinted as pp. 561–630 of *Collected Papers of G. H. Hardy*, vol. 1, Oxford, Clarendon Press.

[Hardy and Wright 1979] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford, Clarendon Press, 1979.

[Hecke 1918] E. Hecke, “Eine neue Art von Zetafunktionen und ihre Bezeihungen zur Verteilung der Primzahlen I”, *Math. Zeitschrift* **1** (1918), 357–376.

[Hecke 1920] E. Hecke, “Eine neue Art von Zetafunktionen und ihre Bezeihungen zur Verteilung der Primzahlen II”, *Math. Zeitschrift* **6** (1920), 11–51.

[Hildebrand and Maier 1989] A. Hildebrand and H. Maier, “Irregularities in the distribution of primes in short intervals”, *J. Reine Angew. Math.* **397** (1989), 162–193.

[Holben and Jordan 1968] C. A. Holben and J. H. Jordan, “The twin prime problem and Goldbach’s conjecture in the Gaussian integers”, *Fibonacci Quart.* **6**:5 (1968), 81–85, 92.

[Jordan and Rabung 1970] J. H. Jordan and J. R. Rabung, “A conjecture of Paul Erdős concerning Gaussian primes”, *Math. Comp.* **24** (1970), 221–223.

[Jordan and Rabung 1976] J. H. Jordan and J. R. Rabung, “Local distribution of Gaussian primes”, *J. Number Theory* **8**:1 (1976), 43–51.

[Kesten 1982] H. Kesten, *Percolation theory for mathematicians*, Progress in Probability and Statistics **2**, Birkhäuser, Boston, 1982.

[Meester and Roy 1996] R. Meester and R. Roy, *Continuum percolation*, Cambridge Tracts in Mathematics **119**, Cambridge University Press, Cambridge, 1996.

[Penrose 1991] M. D. Penrose, “On a continuum percolation model”, *Adv. in Appl. Probab.* **23**:3 (1991), 536–556.

[Rademacher 1923] H. Rademacher, “Über die Anwendung der Viggo Brunschen Methode auf die Theorie der algebraischen Zahlkörper”, *Sitzungsber. Preuss. Akad. Wiss.* **24** (1923), 211–218.

[Riesel 1985] H. Riesel, *Prime numbers and computer methods for factorization*, Progress in Mathematics **57**, Birkhäuser, Boston, 1985.

[Russo 1981] L. Russo, "On the critical percolation probabilities", *Z. Wahrsch. Verw. Gebiete* **56**:2 (1981), 229–237.

[Stauffer and Aharony 1994] D. Stauffer and A. Aharony, *Introduction to percolation theory*, 2nd ed., Taylor and Francis Ltd., London, 1994.

[Vardi $\geq$ 1998] I. Vardi, "Number theoretic percolation", in preparation.

[Wierman 1995] J. C. Wierman, "Substitution method critical probability bounds for the square lattice site percolation model", *Combin. Probab. Comput.* **4**:2 (1995), 181–188.

[Ziff and Sapoval 1986] R. M. Ziff and B. Sapoval, "The efficient determination of the percolation threshold by a frontier-generating walk in a gradient", *J. Phys A* **19** (1986), L1169–L1172.

[Zuev and Sidorenko 1985] S. A. Zuev and A. F. Sidorenko, "Continuous models of percolation theory, I", *Teoret. Mat. Fiz.* **62**:1 (1985), 76–86. In Russian; translation in *Theor. Math. Phys.* **62**, 51–59 (1985).

Ilan Vardi, Mathematics Department, Occidental College, Los Angeles, CA 90041 (ilan@cco.caltech.edu)

Received March 20, 1997; accepted in revised form February 13, 1998