C. R. Math. Rep. Acad. Sci. Canada Vol. **42** (3) 2020, pp. 30–37

Gaussian Primes

J.B. Friedlander, FRSC

Abstract. We survey some of the many interesting questions and results that accrue to the prime numbers which are the sum of two squares.[1]

Résumé. On survole quelques-uns des plusieurs questions et résultats intéressants qui s'accroissent aux nombres premiers qui sont la somme de deux carrés.

1**. Introduction** The very simple field extension of the rational numbers $\mathbb{Q}$ by the imaginary unit i is known as the "gaussian field". The elements of its integral closure, the ring $\mathbb{Z}[i]$, are labelled "gaussian integers". The prime ideals of this ring should perhaps be called gaussian primes. However, we shall find it convenient to ignore the ramified and the inert primes and use this terminolgy to instead describe the odd rational primes which occur as norms, $p = 4a^2 + b^2$, and so which split as a product of distinct pairs of prime ideals $\mathfrak{p}$, $\overline{\mathfrak{p}}$ generated by $2a \pm bi$. The primes p are, by a famous theorem due to Fermat and Euler, just the primes of the form $4q + 1$. By restricting to positive a and b we obtain a one-to-one correspondence between p and its "coordinates" $(2a, b)$ in the first quadrant of the complex plane.

Convention: Throughout the paper, when we mention the counting of prime numbers (or ideals) we shall normally be referring to a count of these corresponding points. Thus, each corresponds to a single value of p; it corresponds to only one of the two prime ideals (we'll denote this one by $\mathfrak{p}$) into which p splits, to only one of the four points in the full plane corresponding to $\mathfrak{p}$ and to none of the four points in the plane corresponding to the other ideal $\overline{\mathfrak{p}}$.

Beyond the Fermat-Euler result, there are quite a number of beautiful and significant theorems and conjectures about these gaussian primes and many of them are described most naturally in terms of their coordinates.

Most, if not all, of the questions we discuss, and of the results referenced here, have natural generalizations from the gaussian field to other quadratic fields or, alternatively, from the quadratic form $4a^2 + b^2$ to other binary quadratic forms, but, apart from mentioning a few of the more general recent results, we shall confine ourselves to this special, though fundamental, case.

Received by the editors on September 30, 2020.
Supported in part by NSERC grant A5123
AMS Subject Classification: 11N05,11N32.

[1]Based closely on the expository part of a one hour lecture delivered at the 60th Birthday Conference in Honour of W. D. Duke, ETH Zurich, June 2019.

2. **Quantitative Distribution** The first question one might ask is to wonder how many of them there are. A little elementary number theory, dating at least to Euler, shows that, for given m, the integer $4m^2+1$ can only have prime factors congruent to one modulo four. It follows from a very slight modification of Euclid's famous proof of the infinitude of primes that the same is true of the set of primes of the form $4q+1$. (Suppose there are only finitely many, let m denote their product and then ask about the prime factors of $4m^2+1$.)

By the Prime Number Theorem one knows that $\pi(x)$, the number of prime numbers up to x, satisfies the asymptotic formula

$$\pi(x) \sim \int_2^x dt/\log t \sim x/\log x \ ,$$

hence one might reasonably guess that about half of them are gaussian and

$$\pi(x;4,1) = \sum_{\substack{p\leqslant x \\ p\equiv 1 (\mathrm{mod}\, 4)}} 1 \sim x/2\log x.$$

This is indeed the case, by a very simple modification of the proof of the Prime Number Theorem. Although there are some very substantial problems sparked by the "prime number race" between $\pi(x;4,1)$ and $\pi(x;4,3)$, there seems little new in the more detailed quantitative measure of either quantity beyond what are already the crucial problems for $\pi(x)$ itself.[2]

3. **Geographic Distribution** Having determined, at least as far as asymptotics, the frequency of the primes we are considering, it is natural to study questions about their spatial positioning in the plane.

3.1. Uniformity in reduced residue classes Perhaps the first question one might ask is that which is analogous to the problem of primes in arithmetic progressions, where now by an arithmetic progression we are intending the translate by an eligible integer vector of a sublattice of $2\mathbb{Z}\times\mathbb{Z}$. Here too, the relevant set of (gaussian) integers can be singled out by the orthogonality property of residue class characters χ and so we obtain the expected result:

$$\sum_{\substack{N\mathfrak{p}\leqslant x \\ \mathfrak{p}\equiv\mathfrak{m}(\mathrm{mod}\,\mathfrak{f})}} 1 \sim \frac{1}{\varphi(\mathfrak{f})}\pi(x;4,1) \ .$$

3.2. Uniformity in sectors A question, novel to this planar setting and solved by E. Hecke [He] with the introduction of his Grössencharaktere $\chi(\mathfrak{a})\mathfrak{a}/|\mathfrak{a}|^m$ for $m\in\mathbb{Z}$, was the demonstration of the uniform distribution of primes in sectors of the plane:

[2] On the other hand, the quantitative study of primes in arithmetic progressions to large moduli brings with it new questions of fundamental character.

Let $0 \leqslant \alpha < \beta \leqslant \pi/2$. Then

$$\sum_{\substack{N\mathfrak{p}\leqslant x\\ \alpha<\arg\mathfrak{p}\leqslant\beta}} 1 \sim \frac{\beta-\alpha}{\pi}\frac{x}{\log x}.$$

3.3. Uniformity in smooth domains The previous result leads also in a well-known fashion to a demonstration of the uniformity of distribution of such prime points in any nice region. Specifically, given a domain $\mathcal{D}$ of area $|\mathcal{D}|$ in the first quadrant, what is required is only that one can one approximate it well by a disjoint union over small (but not too small) polar boxes of the type

$$\mathcal{B} = \{z; c_1 x < |z| \leqslant c_2 x,\ \alpha < \arg z \leqslant \beta\}$$

and then apply the sector estimate to each small box. We obtain

$$\#\{\mathfrak{p}; \mathfrak{p}/\sqrt{x} \in \mathcal{D}\} \sim \frac{|\mathcal{D}|}{2\pi}\frac{x}{\log x}\ .$$

3.4. Primes of small height Here, we must hasten to say that, by "height" here, we are only referring to the size of the second coordinate b. Here one has the result:

For suitable $\alpha < 1$ there are infinitely many primes $\mathfrak{p} = 2a+bi$ with $1 \leqslant b \leqslant p^\alpha$ where as usual $p = N\mathfrak{p}$.

The first result of this nature was due to M. Coleman [Co] in 1993. The current record still seems to be $\alpha = .119$, due to G. Harman and P. Lewis [HL] and dating to 2001.

Although quite nice, these results are quite far from the expected truth, namely $b = 1$, which just rephrases a famous E. Landau problem:

Conjecture: There are infinitely many prime numbers of the form $n^2 + 1$.

3.5. Small gaps A fundamental problem in analytic number theory is that of showing that every interval of the type $(x - x^\theta, x)$, with x sufficiently large, contains a prime number and where one would like to show that this holds for every $\theta > 0$. The case of $\theta \geqslant 1$ follows at once from the Prime Number Theorem but the known error terms in that theorem are still much too weak to quickly prove it for any $\theta < 1$. Nevertheless, that breakthrough was achieved by Hoheisel more than eighty years ago and, after various improvements over the years, the current record of $\theta = 0.525$ is due to R. Baker, G. Harman and J. Pintz [BHP].

This suggests an analogous question for our situation and, in the case of this problem, we shall state the result in a more general form due to G. Harman, A. Kumchev and P. Lewis [HKL] in 2004.

Theorem: *Let $Q(x, y)$ be a primitive positive definite binary quadratic form. For all $(s, t) \in \mathbb{R}^2$ with $Q(s, t)$ sufficiently large, there exists $(m, n) \in \mathbb{Z}^2$ such that $Q(m, n)$ is prime and*

$$|Q(m, n)) - Q(s, t)| < Q(s, t)^{0.53}\ .$$

Here, the exponent is slightly worse than that in the one dimensional case, due to the lack of applicability here of a deep and powerful tool due to N. Watt [Wa].

3.6. Bounded gaps The previous subsection dealt with the quantification of gaps in the plane wherein every point was approximated by a prime point. We now wish to mention a somewhat similar sounding problem wherein we ask that such gaps occur between primes (that is, prime pairs) on the same vertical line, Now such gaps can only be required to happen infinitely often. Rather than Hoheisel's theorem, the motivational source here is the famous twin prime problem and the recent breakthrough approaches thereto by D. Goldston, J. Pintz and C. Yildirim [GPY], by Y. Zhang [Zh] and by J. Maynard [Ma1].

By adapting those ideas to the situation for gaussian primes, A. Vatwani [Va] proved the following result:

Theorem: *There exist infinitely many prime pairs* $\mathfrak{p} = 2a + bi$, $\mathfrak{p}' = 2a + b'i$ *with* $|b - b'| \leqslant 246$.

We note that the two primes have the same first coordinate.

In obtaining the constant 246, Vatwani thus matched the then best known value in the rational case.

The previous result reminds one of an attractive problem, sometimes attributed to P. Erdös and which has no natural analogue in the rational case.

Problem: Show (or disprove) that there is an infinite sequence of distinct gaussian primes $\mathfrak{p}_n$ and a constant C such that $|\mathfrak{p}_n - \mathfrak{p}_{n-1}| \leqslant C$. A colourful way of phrasing this: "Can you march to infinity, taking steps of bounded length and stepping only on the gaussian primes?"

4. Arithmetic Distribution

4.1. Early results An interesting and quite different class of problems arises when we consider the size of subsets of the gaussian primes whose cooordinates are distinguished by having interesting arithmetic properties. As in the previous situations, the problem is most interesting when the set of primes in question is a sparse one.

The first results of this nature date from the period 1995–1998 and, of these, the very first was the theorem of E. Fouvry and H. Iwaniec [Fo-I] which, in its motivating special case, considered one of the coordinates to be prime. Namely:

Theorem: *We have*

$$\sum_{\substack{p_1 \leqslant x \\ p_1 = 4a^2 + p_2^2}} \sum \log p_1 \log p_2 \sim c_1 x$$

where

$$c_1 = \frac{\pi}{4} \prod_p \left(1 - \frac{\chi_4(p)}{p-1}\right) ,$$

with χ_4 being the non-principal character of modulus four.

There followed the theorem of J. Friedlander and H. Iwaniec [FI1,2,3], which showed how to count the frequency of primes which can be written as the sum of a square and a fourth power, giving the first demonstrated example of infinitely many primes in a polynomial sequence of density a power smaller than one.

Theorem: *We have*

$$\sum_{\substack{p\leqslant x\\ p=4a^2+b^4}}\sum \log p \sim c_2 x^{\frac{3}{4}}\ , \qquad c_2 = 2\pi^{-1}\int_0^1 \sqrt{1-t^4}\,dt\ .$$

4.2. Recent results It took about twenty years for these ideas to become more completely assimilated but fairly recently, in the period 2015–2018, several new developments have now taken place.

In the first of these, D.R. Heath-Brown and X. Li [HbL] succeeded in combining the thrust of the two previous theorems to show the following.

Theorem: *We have*

$$\sum_{\substack{p_1\leqslant x\\ p_1=4a^2+p_2^4}}\sum \log p_1 \log p_2 \sim c_3 x^{\frac{3}{4}}\ ,$$

for a suitable postive constant c_3.

Another more recent step, taken by P.C-H. Lam, D. Schindler and S.Y. Xiao [LSX], was to generalize the result of Fouvry-Iwaniec so that it no longer applies solely to the sums of two squares but now to the integers represented by any arbitrary, primitive, positive-definite, binary quadratic form.

It would be nice to see a proof of the analogous generalization of the Friedlander-Iwaniec theorem.

One way to produce a sparse set of primes, long expected to be the case but only recently shown so by J. Maynard [Ma2], is to require that there be previously specified digits which are missing from their decimal (or other digital) expansion. A nice theorem of this type for gaussian primes has recently been given by K. Pratt [Pr].

As Pratt's main theorem is rather technical to state, we are going to give it in a less general form. We consider a set $\mathcal{A}$ of positive integers which is constructed as follows. We fix a set of one, two or three forbidden decimal digits and ask that $\mathcal{A}$ consist of those integers having no forbidden digit in their decimal expansion. We are going to count the gaussian primes up to x which have a coordinate in $\mathcal{A}$. For technical reasons in the proof it is easier to do this if we count only those coordinates in $\mathcal{A}$ with no small prime factor, say less than $P = \exp\big((\log x)^{\frac{1}{3}}\big)$. Then,

Theorem: *With the above notation we have*

$$\sum_{\substack{p=4a^2+b^2\leqslant x\\ b\in\mathcal{A},(b,\prod_{p'\leqslant P}p')=1}}\sum \log p \sim \frac{c_4}{\log P}\sum_{\substack{4a^2+b^2\leqslant x\\ b\in\mathcal{A}}}\sum 1\ ,$$

where the positive constant c_4 depends only on the choice of digits which are to be forbidden.

5. **Both Coordinates Specialized** In all of the results of the previous section, one is considering gaussian primes for which one of the coordinates has some special arithmetic property, a prime, a square, an integer without prohibited digits. In this section we describe recent results of the author, joint with H. Iwaniec [FI4], in which we require that both coordinates be special.

On the one hand, there are certain examples of such results which could be expected to not be too difficult. One could for example ask for both coordinates to be squarefree, and even to not have too many prime factors. But it would seem more like progress if one kept the requirement of one coordinate being special in one of the previous ways and still asked something of the other coordinate.[3]

An ideal goal would be to succeed in treating gaussian primes $2a+bi$ with a and b both being prime. Such a result seems well beyond current reach though one can certainly conjecture an expected result.

Conjecture: We have

$$\sum_{\substack{p_1\leqslant x\\ p_1=4p_2^2+p_3^2}}\sum \log p_1\log p_2\log p_3 \sim c_5 x$$

where

$$c_5 = \prod_{p\equiv 1(\mathrm{mod}\,4)}\Big(1-\frac{3}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-3}\prod_{p\equiv 3(\mathrm{mod}\,4)}\Big(1-\frac{1}{p^2}\Big)^{-1}\ .$$

This would certainly be a profound development of the original Fouvry-Iwaniec result.

In our joint work with Iwaniec we took a step in this direction by studying gaussian primes in which we still require one coordinate to be prime and the other to not have too many prime factors.

In our case, after a good deal of work, we were able to succeed with "not too many" meaning not more than seven prime factors. The proof is difficult and in some places bears resemblance (though not quite that difficult) to the arguments in our earlier work [FI2].

[3]Such a situation would have occurred if, in the result of Heath-Brown and Li, it were the square rather than the fourth power that one requested to be a prime power. Perhaps such a result is too much to ask of the current state of the art.

The main statement is rather complicated, requiring a number of preparatory functions and definitions. The argument is based on sieve methods and, though it produces an asymptotic formula, it is not for precisely what one might have wanted to count. Hence, we shall state only a corollary of the main theorem, one that gives, in a more easily digested fashion, the flavour of the work. In the transition to the corollary, the asymptotic formula is lost but the correct order of magnitude of the counting is maintained.

Theorem: *Let $\beta_a = 1$ if a has no more than seven prime factors, each larger than $a^{1/49}$ and equal to zero otherwise. Then*

$$\sum_{\substack{p_1 \leqslant x \\ p_1 = 4a^2 + p_2^2}} \sum \beta_a \log p_1 \log p_2 \asymp x \ (\log x)^{-1}.$$

Here, as usual, the notation means that the sum on the left has both upper and lower bounds of the same order of magnitude as the quantity on the right. We remark that, as quite normally happens with sieve arguments, the upper bound is much easier to prove than the lower bound.

Remarks: The problem studied in this section targets the original Fouvry-Iwaniec result. It can be transformed into an analogous question or questions about each of the other results mentioned in the previous Section 4. One could also consider problems in which the sparsity of the set of gaussian primes is achieved by hybridizing an arithmetic constraint on one or both of the cooordinates with a spatial constraint of one of the types described in Section 3. It would seem worthwhile to know which of these problems will lead to interesting, yet accessible, results.

REFERENCES

BHP. R.C. Baker, G. Harman and J. Pintz, The difference between consecutive primes II, *Proc. London Math. Soc.* **83** (2001), 532–562.

Co. M.D. Coleman, The Rosser-Iwaniec sieve in number fields, with an application, *Acta Arith.* **65** (1993), 53–83.

FoI. E. Fouvry and H. Iwaniec, Gaussian primes, *Acta Arith.* **79** (1997), 249–287.

FI1. J.B. Friedlander and H. Iwaniec, Using a parity-sensitive sieve to count prime values of a polynomial, *Proc. Nat. Acad. Sci. U.S.A.* **94** (1997), 10541058.

FI2. J.B. Friedlander and H. Iwaniec, The polynomial $X^2 + Y^4$ captures its primes, *Ann. of Math.* **148** (1998), 945–1040.

FI3. J.B. Friedlander and H. Iwaniec, Asymptotic sieve for primes, *Ann. of Math.* **148** (1998), 1041–1065.

FI4. J.B. Friedlander and H. Iwaniec, Coordinate distribution of Gaussian primes, *J. European Math. Soc.*, to appear.

GPY. D.A. Goldston, J. Pintz and C.Y. Yildirim, Primes in tuples I, *Ann. of Math.* **170** (2009), 819–862.

HKL. G. Harman, A. Kumchev and P. Lewis, The distribution of prime ideals of imaginary quadratic fields, *Trans. Amer. Math. Soc* **356** (2004), 599–620.

HL. G. Harman and P. Lewis, Gaussian primes in narrow sectors, *Mathematika* **48** (2001), 119–135.

HbL. D.R. Heath-Brown and X. Li, Prime values of $a^2 + p^4$, *Invent. Math.* **208** (2017), 441499.

He. E. Hecke, Eine neue Art von Zetafunktionen II, *Math. Z.* **6** (1920), 11–51.
LSX. P.C-H. Lam, D. Schindler and S.Y. Xiao, On prime values of binary quadratic forms with a thin variable, *J. London Math. Soc.* (2020) https://doi.org/10.1112/jlms.12336
Ma1. J. Maynard, Small gaps between primes, *Ann. of Math.* **181** (2015), 383–413.
Ma2. J. Maynard, Primes with restricted digits, *Invent. Math.* **217** (2019), 127–218.
Pr. K. Pratt, Primes from sums of two squares and missing digits, *Proc. London Math. Soc.* **120** (2020), 770–830.
Va. A. Vatwani, Bounded gaps between Gaussian primes, *J. Number Theory* **171** (2017), 449–473.
Wa. N. Watt, Kloosterman sums and a mean value for Dirichlet polynomials, *J. Number Theory* **53** (1995), 179–210.
Zh. Y. Zhang, Bounded gaps between primes, *Ann. of Math.* **179** (2014), 1121–1174.

Department of Mathematics, University of Toronto, Toronto, Canada M5S 2E4
e-mail: frdlndr@math.toronto.edu