Proc. Indian Acad. Sci. (Math. Sci.), Vol. 97, Nos 1–3, December 1987, pp. 111–116.

The area within a curve

M N HUXLEY

Department of Pure Mathematics, University College, Cardiff, PO Box 78, Cardiff CF1 1XL, UK

Abstract. The area of a simple closed convex curve can be estimated in terms of the number of points of a square lattice that lie within the curve. We obtain the usual error bound without integration using a form of the Hardy–Littlewood–Ramanujan circle method, and also present simple estimates for the mean square error.

Keywords. Estimation of area; lattice points.

1. Introduction

This note treats an old and familiar problem: to estimate the area A within a simple closed convex curve C, drawn on graph paper which carries a square lattice of sides $1/R$. To clear fractions we take the side of the lattice squares as the new unit of length; C has area AR^2 in these units. There are three usual methods of estimating the area. The trapezium rule adds up the distances intercepted by the curve along lattice lines parallel to the y-axis. The lattice squares rule counts the number of lattice squares whose centres lie within the curve. The lattice points rule counts the number of lattice points (crossings of the lattice lines) within the curve. To be more precise, if the curve is given by two equations $y = F(x)$, $y = f(x)$ for $a \leqslant x \leqslant b$, where $f(a) = F(a)$, $f(b) = F(b)$, the trapezium rule estimate is

$$T = \sum_{a \leqslant n \leqslant b} (F(n) - f(n)),$$

and the lattice squares and lattice points estimates are special cases of the integer part sum (in which square brackets denote the integer part)

$$N(u, v) = \sum_{a-u \leqslant n \leqslant b-u} ([F(n+u) - v] - [f(n+u) - v])$$

with $(u, v) = (\frac{1}{2}, \frac{1}{2})$ and $(0, 0)$ respectively.

We write $D(u, v)$ for the discrepancy

$$D(u, v) = N(u, v) - AR^2,$$

and similarly

$$E(u) = T(u) - AR^2 = \sum_{a-u \leqslant n \leqslant b-u} (F(n+u) - f(n+u)).$$

If the curve C is contained in a rectangle with sides parallel to the coordinate axes, lengths B and H respectively, then it is easy to see geometrically that

$$|E(u)| \leqslant 2H,$$

$$|D(u, v)| \leqslant 2B + 2H + 4.$$

The error bounds on the right are both order of magnitude $O(R)$ as R tends to infinity. If the curve C is itself a rectangle, these bounds cannot be improved. We use the following order of magnitude notation: $g(R) \ll G(R)$ as R tends to infinity means that

$$\limsup_{R \to \infty} \frac{|g(R)|}{G(R)} < \infty.$$

Any term $g(R)$ in an estimate may be replaced by $O(G(R))$, read as 'an error term of order of magnitude $G(R)$'.

From now on we suppose that C is a smooth curve with continuous non-zero radius of curvature. Estimates of the form

$$D(u, v) \ll R^{(2/3)+\varepsilon} \tag{1}$$

(for any $\varepsilon > 0$) were given by Voronoi, Sierpiński, van der Corput, Landau, Jarnik and Vinogradov, sometimes as special cases only. Jarnik also showed that the exponent 2/3 cannot be reduced without further smoothness conditions. The problem is treated at length in Landau [7] and briefly in Vinogradov [10]. A remark apparently missing from Landau is that the bounds

$$E(u) \ll R^{1/2}, \tag{2}$$

$$\int_0^1 |E(u)|^2 \, du \leqslant \int_0^1 \int_0^1 |D(u, v)|^2 \, du \, dv \ll R \tag{3}$$

and, if the curve C is three times differentiable,

$$\int_0^1 |E(u)|^2 \, du \gg R \tag{4}$$

are easy to establish using Fourier series. Thus the error term $E(u)$ in the trapezium rule for a closed curve is well understood. Moreover (3) suggests that if C is sufficiently smooth, then the exponent in (1) should fall, perhaps to $\frac{1}{2} + \varepsilon$.

A related problem, considered by Swinnerton-Dyer [9], is to bound P, the number of lattice points on the curve C. He found the result

$$P \ll R^{3/5+\varepsilon}.$$

The ambiguous status of lattice points on the curve means that

$$D(0, 0) \geqslant P/2.$$

Swinnerton-Dyer has suggested that the geometrical ideas in [9] could be applied to the lattice point problem.

Recently Iwaniec and Mozzochi [5], building on previous work by Bombieri and

Iwaniec [1], obtained

$$D(0,0) \ll R^{7/11+\varepsilon}$$

for the circle. Their method can be generalized [3, 4]. In this note we give a strictly elementary proof of (1), taking ideas from [5] and [10] (and thus ultimately from Sierpiński [8]). We use the continued fraction algorithm, which tradition ascribes to Brahmagupta, and the circle method, which first appears in Hardy and Ramanujan's great paper on the partition function [2]. For completeness we also prove (2), (3) and (4) using exponential sums. We hope to shed light on the Iwaniec–Mozzochi memoir, which uses the circle method, exponential sums and the large sieve, and on Jutila's work on Voronoi summation (see for example [6]), which rests on the transformation formula for an Eisenstein series just as Hardy and Ramanujan rely on the transformation formula for the discriminant function. Applying the circle method to exponential sums is a new and exciting development, which may go as far as the circle method for exponential integrals in the hands of Hardy, Littlewood and Vinogradov.

2. The circle method

The circle method in its usual form divides the circle into a large finite number of arcs, each containing a point whose polar angle is $2\pi p/q$ where p and q are small integers. We divide the curve into a large number K of arcs, each of length $\ll L = R/K$, and also of length $\gg L$, and we pick a point on each arc where the tangent has a rational gradient p/q, with p and q numerically as small as possible. We draw the tangent at these points, forming a K-sided polygon circumscribed to the curve. We also draw a polygon with parallel sides inscribed to the curve. Since by assumption the radius of curvature lies between multiples of R, the distance between corresponding sides of the two polygons is $\ll \delta = R/K^2$ and the difference in area of the two polygons is $\ll \delta K L$.

The discrepancy of a polygon is treated in Vinogradov [10], but only through worked examples. There is a contribution from each side and from each vertex: the bound is

$$\ll \sum_k S(\alpha_k, L_k),$$

where the sum is over sides of the polygon, L_k being the side length and α_k its gradient. The sum $S(\alpha, l)$ is defined as follows. Take the continued fraction

$$|\alpha| = a_0 + 1/a_1 + 1/a_2 + \cdots$$

and truncate it at an index r with $q_r \leqslant l$. Then

$$S(\alpha, l) = a_0 + a_1 + \cdots + a_r + l/q_r.$$

We have $L_k \ll L$ for each side k. Since $q_r \ll L$, we have $r \ll \log L$. Fixing $a_0, \ldots, a_{i-1}$ in a continued fraction determines an interval of length Δ say, on which $a_i = n$ for a subinterval of length $\ll \Delta/n^2$. For $|\alpha| \leqslant 1$ the arc on the curve corresponds to an interval of length $\gg 1/K$ within which we must choose a rational number. (For $|\alpha| > 1$

the same argument applies to $1/|\alpha|$). The expectation of a_i is thus

$$\ll \sum_{n \ll L} n \cdot \frac{1}{n^2} \ll \log L$$

independently of $a_0, \ldots, a_{i-1}$. Finally for fixed p_r/q_r there are

$$\ll 1 + K/q_r L$$

tangents for which the truncation of the continued fraction for the gradient p/q gives p_r/q_r. Thus

$$\sum_k S(\alpha_k, L_k) \ll K \log^2 L + \sum_{q \ll L} \left(1 + \frac{K}{qL}\right) \frac{L}{q} \cdot q \ll K \log^2 L + L^2.$$

Since the area of the curve lies between the areas of the circumscribed and inscribed polygons we have

$$D(0,0) \ll \delta K L + K \log^2 L + L^2.$$

Taking

$$K = R^{2/3} (\log R)^{-2/3} + \mathrm{O}(1)$$

gives

$$N(0,0) = AR^2 + \mathrm{O}(R^{2/3} (\log R)^{4/3}),$$

and the same estimate holds for $N(u, v)$ on translating the curve by the vector (u, v).

3. Fourier analysis of the discrepancy

In this section we use $\sim$ to indicate 'possesses the Fourier series', $\approx$ to indicate 'is asymptotically equal to', and write $e(t)$ for $\exp 2\pi i t$. Since $N(u, v)$ is periodic in u and in v, it has a Fourier series

$$N(u, v) \sim \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{hk}\, e(hu + kv)$$

with $c_{00} = A$. Writing the 'row of teeth' function as

$$\rho(t) = [t] - t + \tfrac{1}{2} \sim \sum_{k \neq 0} \frac{e(kt)}{2\pi i k},$$

we have

$$N(u, v) = T(u) + \sum_{a-u \leqslant n \leqslant b-u} (\rho(F(n+u) - v)) - \rho(f(n+u) - v)).$$

The second term can be expanded again as

$$\sim \sum_{k \neq 0} \frac{e(-kv)}{2\pi ik} \sum_{a \quad u \leq n \leq b \quad u} (e(kF(n+u)) - e(kf(n+u)))$$

$$\sim \sum_{k \neq 0} \frac{e(-kv)}{2\pi ik} \sum_h e(hu) \int_0^1 \sum_{a \quad x \leq n \leq b-x} (e(kF(n+x)) - e(kf(n+x)))e(-hx)\mathrm{d}x.$$

Thus for $k \neq 0$

$$2\pi ik\, c_{h,-k} = \int_a^b (e(kF(x) - e(kf(x)))e(-hx)\,\mathrm{d}x = I_1 - I_2,$$

say. Since f'' and $-F''$ are both $\gg 1/R$, van der Corput's second derivative test gives

$$I_1, I_2 \ll (R/|k|)^{1/2}, \quad c_{hk} \ll R^{1/2}/|k|^{3/2}.$$

Similarly we could consider x as a function of y on the curve and obtain

$$c_{hk} \ll R^{1/2}/|h|^{3/2}.$$

Thus

$$|c_{hk}|^2 \ll R/(h^2 + k^2)^{3/2}.$$

Now

$$E(u) \sim \sum_{h \neq 0} c_{h0} e(hu) \ll \sum_{h \neq 0} |c_{h0}| \ll R^{1/2},$$

which gives (2), and by Parseval's identity

$$\int_0^1 |E(u)|^2 \mathrm{d}u = \sum_{h \neq 0} |c_{h0}|^2 \leq \int_0^1 \int_0^1 |D(u,v)|^2 \mathrm{d}u\,\mathrm{d}v = \sum_{(h,k) \neq (0,0)} \sum |c_{hk}|^2 \ll R,$$

which is (3).

For the lower bound we assume that the curve is three times differentiable. If x_0, x_1 give the maximum of $F(x)$ and the minimum of $f(x)$ respectively, then

$$I_1 \approx \left(\frac{i}{kF''(x_0)}\right)^{\frac{1}{2}} e(kF(x_0)),$$

$$I_2 \approx \left(\frac{i}{kf''(x_1)}\right)^{\frac{1}{2}} e(kf(x_1)),$$

where the square roots of complex numbers are taken with positive real part. If these two integrals cancel, then

$$e(kF(x_0) - kf(x_1)) \approx -i.$$

This approximate equality cannot hold for $k = 1$ and for $k = 2$, so

$$|c_{01}|^2 + |c_{02}|^2 \gg R.$$

By symmetry between the x and y directions, we have

$$|c_{10}|^2 + |c_{20}|^2 \gg R,$$

which gives (4).

References

[1] Bombieri E and Iwaniec H, On the order of $(\frac{1}{2}+it)$, *Ann. Scuola Normale Superiore di Pisa cl. Sci.* (4) **13** (1986) 449–472

[2] Hardy G H and Ramanujan S, Asymptotic formulae in combinatory analysis, *Proc. London Math. Soc.* 2 **17** (1918) 75–115

[3] Huxley M N and Watt N, Exponential sums and the Riemann zeta function *Proc. London Math. Soc.* (to appear)

[4] Huxley M N, Exponential sums and lattice points (to appear)

[5] Iwaniec H and Mozzochi C J, On the circle and divisor problems (to appear)

[6] Jutila M, *Lectures on a method in the theory of exponential sums* (Bombay: Tata Institute of Fundamental Research) (1987)

[7] Landau E, *Vorlesungen über Zahlentheorie* (Leipzig: Hirzel) Band 2 (1927)

[8] Sierpin'ski W, A certain problem in the calculus of asymptotic functions, *Prace Mat.-Fiz.* **17** (1906) 77–118 and *Oeuvres Choisies* 73–108

[9] Swinnerton-Dyer H P F, The number of lattice points on a convex curve, *J. Number Theory* **6** (1974) 128–135

[10] Vinogradov I M, *An introduction to the theory of numbers* (Oxford: Pergamon) (1955)