

# Optimal stretching for lattice points and eigenvalues

Richard S. Laugesen and Shiya Liu

**Abstract.** We aim to maximize the number of first-quadrant lattice points in a convex domain with respect to reciprocal stretching in the coordinate directions. The optimal domain is shown to be asymptotically balanced, meaning that the stretch factor approaches 1 as the “radius” approaches infinity. In particular, the result implies that among all  $p$ -ellipses (or Lamé curves), the  $p$ -circle encloses the most first-quadrant lattice points as the radius approaches infinity, for  $1 < p < \infty$ .

The case  $p=2$  corresponds to minimization of high eigenvalues of the Dirichlet Laplacian on rectangles, and so our work generalizes a result of Antunes and Freitas. Similarly, we generalize a Neumann eigenvalue maximization result of van den Berg, Bucur and Gittins. Further, Ariturk and Laugesen recently handled  $0 < p < 1$  by building on our results here.

The case  $p=1$  remains open, and is closely related to minimizing energy levels of harmonic oscillators: which right triangles in the first quadrant with two sides along the axes will enclose the most lattice points, as the area tends to infinity?

## 1. Introduction

Among ellipses of given area centered at the origin and symmetric about both axes, which one encloses the most integer lattice points in the open first quadrant? One might guess the optimal ellipse would be circular, but a non-circular ellipse can enclose more lattice points, as shown in Figure 1. Nonetheless, optimal ellipses must become more and more circular as the area increases to infinity, by a striking result of Antunes and Freitas [2].

---

*Key words and phrases:* lattice points, planar convex domain,  $p$ -ellipse, Lamé curve, spectral optimization, Laplacian, Dirichlet eigenvalues, Neumann eigenvalues.

*2010 Mathematics Subject Classification:* primary 35P15; secondary 11P21, 52C05.

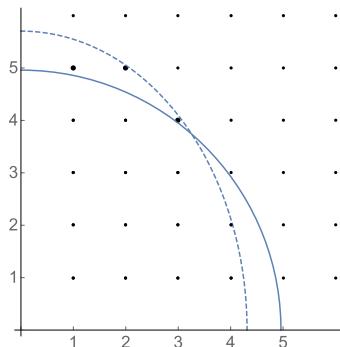


Figure 1. Circle  $s=1$  and ellipse  $s=1.15$ , for radius  $r=4.96$ . The ellipse encloses three more points than the circle, as shown in bold.

To formulate the problem more precisely, consider the number of positive-integer lattice points lying in the elliptical region

$$\left(\frac{x}{s-1}\right)^2 + \left(\frac{y}{s}\right)^2 \leq r^2,$$

where the ellipse has “radius”  $r > 0$  and semiaxes proportional to  $s^{-1}$  and  $s$ . Notice that the area  $\pi r^2$  of the ellipse is independent of the “stretch factor”  $s$ . Denote by  $s=s(r)$  a value (not necessarily unique) of the stretch factor that maximizes the lattice point count. The theorem of Antunes and Freitas says  $s(r) \rightarrow 1$  as  $r \rightarrow \infty$ , as illustrated in Figure 2. In other words, optimal ellipses become circular in the infinite limit. (Their theorem was stated differently, in terms of minimizing the  $n$ -th eigenvalue of the Dirichlet Laplacian on rectangles, with the square being asymptotically minimal. Section 10 explains the connection.) The analogous result for optimal ellipsoids becoming asymptotically spherical was proved recently in three dimensions by van den Berg and Gittins [6] and in higher dimensions by Gittins and Larson [11], once again in the eigenvalue formulation.

This paper extends the result of Antunes and Freitas from circles to essentially arbitrary concave curves in the first quadrant that decrease between the intercept points  $(0, 1)$  and  $(1, 0)$ . The “ellipses” in this situation are the images of the concave curve under rescaling by  $s^{-1}$  and  $s$  in the horizontal and vertical directions, respectively. Theorem 2 says the maximizing  $s(r)$  is bounded. Theorem 3 shows under a mild monotonicity hypothesis on the second derivative of the curve that  $s(r) \rightarrow 1$  as  $r \rightarrow \infty$ . Thus the most “balanced” curve in the family will enclose the most lattice points in the limit.

Marshall [21] recently extended this result to higher dimensions by somewhat different methods. We have generalized also to translated lattices in 2-dimensions [19].

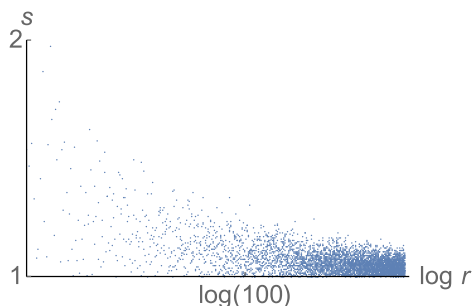


Figure 2. Optimal  $s$ -values for maximizing the number of lattice points in the 2-ellipse. The graph plots the largest value  $s(r)$  versus  $\log r$ . The plotted  $r$ -values are multiples of  $\sqrt{3}/10$ , an irrational number chosen in the hope of exhibiting generic behavior. The horizontal axis is at height  $s=1$ .

Theorem 4 allows the curvature to blow up or degenerate at the intercept points, which permits us to treat the family of  $p$ -ellipses for  $1 < p < \infty$ . In each case the  $p$ -circle is asymptotically optimal for the lattice counting problem in the first quadrant. The case  $p=1$  is open problem. Our numerical evidence in Section 9 suggests that the first-quadrant right triangle enclosing the most lattice points does *not* necessarily approach a 45–45–90 degree triangle as  $r \rightarrow \infty$ . Instead one seems to get an infinite limit set of optimal triangles. See Conjecture 15, where we describe the recent proof by Marshall and Steinerberger [22].

If one counts lattice points in the *closed* first quadrant, that is, counting points on the axes as well, then the results reverse direction from maximization to minimization of the lattice count. Theorem 7 shows that the value  $s=s(r)$  minimizing the number of enclosed lattice points will tend to 1 as  $r \rightarrow \infty$ . In the case of circles and ellipses, this result was obtained recently by van den Berg, Bucur and Gittins [5] (and in higher dimensions by Gittins and Larson [11], generalized by Marshall [21]). As explained in Section 10, they showed that the maximizing rectangle for the  $n$ -th eigenvalue of the Neumann Laplacian must approach a square as  $n \rightarrow \infty$ .

This paper builds on the framework of Antunes and Freitas for ellipses, with new ingredients introduced to handle general concave curves. First we develop a new non-sharp bound on the counting function (Proposition 10) in order to control the stretch factor  $s(r)$ . Then we prove more precise lattice counting estimates (Proposition 11) of Krätzel type, relying on a theorem of van der Corput (Appendix A).

Convex decreasing curves in the first quadrant, such as  $p$ -ellipses with  $0 < p < 1$ , have been treated by Ariturk and Laugesen [4] by building on this paper's results.

## Spectral motivations and results

This paper is inspired by recent efforts to understand the behavior of high eigenvalues of the Laplacian. Write  $\lambda_n$  for the  $n$ -th eigenvalue of the Dirichlet Laplacian on a bounded domain  $\Omega$  of area 1 in the plane. (We restrict to 2 dimensions for simplicity.) Denote the minimum value of this eigenvalue over all such domains by  $\lambda_n^*$ , and suppose it is achieved on a domain  $\Omega_n^*$ . What can one say about the shape of this minimizing domain?

For the first eigenvalue, the minimizing domain  $\Omega_1^*$  is a disk, by the Faber–Krahn inequality. For the second eigenvalue,  $\Omega_2^*$  is a union of two disjoint equal-area disks, as Krahn and later P. Szego showed. A long-standing conjecture says  $\Omega_3^*$  should be a disk and  $\Omega_4^*$  should be a union of disjoint non-equal-area disks. For higher eigenvalues ( $n \geq 5$ ), minimizing domains found numerically do not have recognizable shapes; see [1] and [23] and references therein. Antunes and Freitas remark, though, that the “most natural guess” is  $\Omega_n^*$  approaches a disk as  $n \rightarrow \infty$ , which is known to occur if the area normalization is strengthened to a perimeter normalization [3] and [7]. This conjecture would imply the famous Pólya conjecture  $\lambda_n \geq 4\pi n/|\Omega|$ , as Colbois and El Soufi [8, Corollary 2.2] showed using subadditivity of  $n \mapsto \lambda_n^*$ .

A partial result of Freitas [10] succeeds in determining the leading order asymptotic as  $n \rightarrow \infty$  of the minimum value of the eigenvalue sum  $\lambda_1 + \dots + \lambda_n$  (rather than of  $\lambda_n$  itself). This result provides no information on the shapes of the minimizing domains. Larson [17] shows among convex domains that the disk asymptotically maximizes the Riesz means of the Laplace eigenvalues, for Riesz exponents  $\geq 3/2$ . If this Riesz exponent could be lowered to 0, giving asymptotic maximality of the disk for the eigenvalue counting function, then one would obtain the desired conjecture about the eigenvalue minimizing domain  $\Omega_n^*$ .

A complete resolution for rectangular domains was found by Antunes and Freitas [2], using lattice counting methods as explained in Section 10. They proved that the minimizing domain for  $\lambda_n$  among rectangles approaches a square as  $n \rightarrow \infty$ . Similarly, the cube is asymptotically minimal in 3 and higher dimensions [6] and [11].

## Open problem for the harmonic oscillator

Asymptotic optimality of the square for minimizing Dirichlet eigenvalues of the Laplacian on rectangles suggests an analogous open problem for harmonic oscillators. Consider the Schrödinger operator in 2 dimensions with parabolic potential  $(sx)^2 + (y/s)^2$ , where  $s > 0$  is a parameter. Write  $s(n)$  for a parameter value that

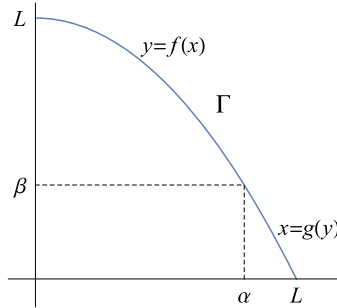


Figure 3. A concave decreasing curve  $\Gamma$  in the first quadrant, with intercepts at  $L$ . The point  $(\alpha, \beta)$  on the curve is relevant to Theorem 3.

minimizes the  $n$ -th eigenvalue of this operator. What is the limiting behavior of  $s(n)$  as  $n \rightarrow \infty$ ?

The results on rectangular domains (which may be regarded as infinite potential wells) might suggest  $s(n) \rightarrow 1$ , but we think that it is not the case. Instead we believe  $s(n)$  might cluster around infinitely many values as  $n \rightarrow \infty$ . Indeed, after rescaling, the Schrödinger operator has eigenvalues of the form  $s(j-1/2) + (k-1/2)/s$ , which leads to a lattice point counting problem inside right triangles, like in Section 9 for  $p=1$ , except now the lattices are shifted by  $1/2$  to the left and downwards. For the unshifted lattice, numerical work in Section 9 suggests that the optimal stretching parameter  $s$  does not converge to 1, and instead has many cluster points as  $r \rightarrow \infty$ . Recent investigations [19, Section 10] suggest that this clustering phenomenon persists for shifted lattices and hence for the harmonic oscillator.

## 2. Assumptions and definitions

The first quadrant is the *open* set  $\{(x, y) : x, y > 0\}$ .

Assume throughout the paper that  $\Gamma$  is a concave, strictly decreasing curve in the first quadrant. Our theorems assume the  $x$ - and  $y$ -intercepts of the curve are equal, occurring at  $x=L$  and  $y=L$  respectively, as shown in Figure 3. Write  $\text{Area}(\Gamma)$  for the area enclosed by the curve  $\Gamma$  and the  $x$ - and  $y$ -axes.

We represent the curve  $\Gamma$  by  $y=f(x)$  for  $x \in [0, L]$ , so that  $f$  is a concave strictly decreasing function, and of course  $f$  is continuous. In particular

$$L = f(0) > f(x) > f(L) = 0$$

whenever  $x \in (0, L)$ . Denote the inverse function of  $f(x)$  by  $g(y)$  for  $y \in [0, L]$ , so that  $g$  also is concave and strictly decreasing.

We define a rescaling of the curve by parameter  $r > 0$ :

$$\begin{aligned} r\Gamma &= \text{image of } \Gamma \text{ under the radial scaling } \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \\ &= \text{graph of } rf(x/r), \end{aligned}$$

and define an area-preserving stretch of the curve by:

$$\begin{aligned} \Gamma(s) &= \text{image of } \Gamma \text{ under the diagonal scaling } \begin{pmatrix} s^{-1} & 0 \\ 0 & s \end{pmatrix} \\ &= \text{graph of } sf(sx), \end{aligned}$$

where  $s > 0$ . In other words,  $\Gamma(s)$  is obtained from  $\Gamma$  after compressing the  $x$ -direction by  $s$  and stretching the  $y$ -direction by  $s$ . Define the counting function for  $r\Gamma(s)$  by

$$\begin{aligned} N(r, s) &= \text{number of positive-integer lattice points lying inside or on } r\Gamma(s) \\ &= \#\{(j, k) \in \mathbb{N} \times \mathbb{N} : k \leq rsf(js/r)\}. \end{aligned}$$

For each  $r > 0$ , consider the set

$$S(r) = \operatorname{argmax}_{s > 0} N(r, s)$$

consisting of the  $s$ -values that maximize the number of first-quadrant lattice points enclosed by the curve  $r\Gamma(s)$ . The set  $S(r)$  is well-defined because the maximum is indeed attained, as the following argument shows. The curve  $r\Gamma(s)$  has  $x$ -intercept at  $rs^{-1}L$ , which is less than 1 if  $s > rL$  and so in that case the curve encloses no positive-integer lattice points. Similarly if  $s < (rL)^{-1}$ , then  $r\Gamma(s)$  has height less than 1 and contains no lattice points in the first quadrant. Thus for each fixed  $r > 0$ , if  $s$  is sufficiently small or sufficiently large then the counting function  $N(r, s)$  equals zero, while obviously for intermediate values of  $s$  the integer-valued function  $s \mapsto N(r, s)$  is bounded. Hence  $N(r, s)$  attains its maximum at some  $s > 0$ .

For later reference, we write down this bound on optimal  $s$ -values.

**Lemma 1.** ( *$r$ -dependent bound on optimal stretch factors*) *If  $\Gamma$  is a concave, strictly decreasing curve in the first quadrant with equal intercepts (as in Figure 3), then*

$$S(r) \subset [(rL)^{-1}, rL] \quad \text{whenever } r \geq 2/L.$$

*Proof.* The curve  $r\Gamma(1)$  has horizontal and vertical intercepts at  $rL \geq 2$ . Hence by concavity,  $r\Gamma(1)$  encloses the point  $(1, 1)$ , and so the counting function  $s \mapsto N(r, s)$  is greater than zero when  $s=1$ . On the other hand when  $s < (rL)^{-1}$  or  $s > rL$ , we know  $N(r, s)=0$  by the paragraph before the lemma. Thus the maximum can only be attained when  $s$  lies in the interval  $[(rL)^{-1}, rL]$ .  $\square$

### 3. Main results

The curve  $\Gamma$  has  $x$ - and  $y$ -intercepts at  $L$ , in the theorems in that follow; see Figure 3. We start by improving Lemma 1 to show the maximizing set  $S(r)$  is bounded, and the bounds can be evaluated explicitly in the limit as  $r \rightarrow \infty$ .

**Theorem 2.** (Uniform bound on optimal stretch factors) *If  $\Gamma$  is a concave, strictly decreasing curve in the first quadrant then*

$$S(r) \subset [s_1, s_2] \quad \text{for all } r \geq 2/L,$$

for some constants  $s_1, s_2 > 0$ . Furthermore, given  $\varepsilon > 0$ ,

$$S(r) \subset \left[ \frac{1}{4+\varepsilon}, 4+\varepsilon \right] \quad \text{for all large } r.$$

The proof appears in Section 4.

If the concave decreasing curve is smooth with monotonic second derivative, then in addition to being bounded above and below the maximizing set  $S(r)$  converges to  $\{1\}$ , as the next theorem shows. Recall that  $g$  is the inverse function of  $f$ .

**Theorem 3.** (Optimal concave curve is asymptotically balanced) *Assume  $(\alpha, \beta) \in \Gamma$  is a point in the first quadrant such that  $f \in C^2[0, \alpha]$  with  $f' < 0$  on  $(0, \alpha]$  and  $f'' < 0$  on  $[0, \alpha]$ , and similarly  $g \in C^2[0, \beta]$  with  $g' < 0$  on  $(0, \beta]$  and  $g'' < 0$  on  $[0, \beta]$ . Further suppose  $f''$  is monotonic on  $[0, \alpha]$  and  $g''$  is monotonic on  $[0, \beta]$ .*

*Then the optimal stretch factor for maximizing  $N(r, s)$  approaches 1 as  $r$  tends to infinity, with*

$$S(r) \subset [1 - O(r^{-1/6}), 1 + O(r^{-1/6})],$$

and the maximal lattice count has asymptotic formula

$$\max_{s>0} N(r, s) = r^2 \text{Area}(\Gamma) - rL + O(r^{2/3}).$$

The theorem is proved in Section 7. Slight improvements to the decay rate  $O(r^{-1/6})$  and the error term  $O(r^{2/3})$  are possible, as explained after Proposition 11.

#### More general curves for lattice counting

We want to weaken the smoothness and monotonicity assumptions in Theorem 3. We start with a definition of piecewise smoothness.

*Definition ( $PC^2$ )*

(i) We say a function  $f$  is piecewise  $C^2$ -smooth on a half-open interval  $(0, \alpha]$  if  $f$  is continuous and a partition  $0 = \alpha_0 < \alpha_1 < \dots < \alpha_l = \alpha$  exists such that  $f \in C^2(0, \alpha_1]$  and  $f \in C^2[\alpha_{i-1}, \alpha_i]$  for  $i = 2, \dots, l$ . Write  $PC^2(0, \alpha]$  for the class of such functions.

(ii) Write  $f' < 0$  to mean that  $f'$  is negative on the subintervals  $(0, \alpha_1]$  and  $[\alpha_{i-1}, \alpha_i]$  for  $i=2, \dots, l$ , with the derivative being taken in the one-sided senses at the partition points  $\alpha_1, \dots, \alpha_l$ . The meaning of  $f'' < 0$  is analogous.

(iii) We label partition points using the same letter as for the right endpoint. In particular, the partition for  $g \in PC^2(0, \beta]$  is  $0 = \beta_0 < \dots < \beta_\ell = \beta$ .

For the next theorem, take a point  $(\alpha, \beta) \in \Gamma$  lying in the first quadrant and suppose we have numbers  $a_1, a_2, b_1, b_2 > 0$  and positive valued functions  $\delta(r)$  and  $\varepsilon(r)$  such that as  $r \rightarrow \infty$ :

$$(1) \quad \delta(r) = O(r^{-2a_1}), \quad f''(\delta(r))^{-1} = O(r^{1-4a_2}),$$

$$(2) \quad \varepsilon(r) = O(r^{-2b_1}), \quad g''(\varepsilon(r))^{-1} = O(r^{1-4b_2}).$$

(The second condition in (1) says that  $f''(x)$  cannot be too small as  $x \rightarrow 0$ .) Let

$$e = \min\{\frac{1}{6}, a_1, a_2, b_1, b_2\}.$$

Now we extend Theorem 3 to a larger class of concave decreasing curves.

**Theorem 4.** (Optimal concave curve is asymptotically balanced) *Assume  $f \in PC^2(0, \alpha]$  with  $f' < 0$  and  $f'' < 0$ , and  $f''$  is monotonic on each subinterval of the partition. Similarly assume  $g \in PC^2(0, \beta]$  with  $g' < 0$  and  $g'' < 0$ , and  $g''$  is monotonic on each subinterval of the partition. Suppose the positive functions  $\delta(r)$  and  $\varepsilon(r)$  satisfy conditions (1) and (2).*

*Then the optimal stretch factor for maximizing  $N(r, s)$  approaches 1 as  $r$  tends to infinity, with*

$$S(r) \subset [1 - O(r^{-e}), 1 + O(r^{-e})],$$

*and the maximal lattice count has asymptotic formula*

$$\max_{s>0} N(r, s) = r^2 \text{Area}(\Gamma) - rL + O(r^{1-2e}).$$

The proof is presented in Section 7.

*Example 5.* (Optimal  $p$ -ellipses for lattice point counting) Fix  $1 < p < \infty$ , and consider the  $p$ -circle

$$\Gamma : |x|^p + |y|^p = 1,$$

which has intercept  $L=1$ . That is, the  $p$ -circle is the unit circle for the  $\ell^p$ -norm on the plane. Then the  $p$ -ellipse

$$r\Gamma(s) : |sx|^p + |s^{-1}y|^p \leq r^p$$



has first-quadrant counting function

$$N(r, s) = \#\{(j, k) \in \mathbb{N} \times \mathbb{N} : (js)^p + (ks^{-1})^p \leq r^p\}.$$

We will show that the  $p$ -ellipse containing the maximum number of positive-integer lattice points must approach a  $p$ -circle in the limit as  $r \rightarrow \infty$ , with

$$S(r) \subset [1 - O(r^{-e}), 1 + O(r^{-e})]$$

where  $e = \min\{\frac{1}{6}, \frac{1}{2p}\}$ .

Theorem 3 fails to apply to  $p$ -ellipses when  $1 < p < 2$ , because the second derivative of the curve is not monotonic (see  $f''(x)$  below), and the theorem fails to apply when  $2 < p < \infty$  because  $f''(0) = 0$  in that case. Instead we will apply Theorem 4.

To verify that the  $p$ -circle satisfies the hypotheses of Theorem 4, we let  $\alpha = \beta = 2^{-1/p}$  and choose

$$\delta(r) = r^{-1/p}, \quad \varepsilon(r) = r^{-1/p},$$

for all large  $r$ . Then  $\delta(r) = r^{-2a_1}$  with  $a_1 = 1/2p$ . Next,

$$\begin{aligned} f(x) &= (1 - x^p)^{1/p}, \\ f'(x) &= -x^{p-1}(1 - x^p)^{-1+1/p}, \\ f''(x) &= -(p-1)x^{p-2}(1 - x^p)^{-2+1/p}, \end{aligned}$$

so that

$$|f''(\delta(r))|^{-1} \leq (\text{const.})r^{1-2/p},$$

and hence  $a_2 = 1/2p$  in (1). Thus  $f$  satisfies hypothesis (1).

Further, the interval  $(0, \alpha)$  can be partitioned into subintervals on which  $f''$  is monotonic, because the third derivative

$$f'''(x) = -(p-1)x^{p-3}(1 - x^p)^{-3+1/p}((1+p)x^p + p - 2)$$

vanishes at most once in the unit interval.

The calculations are the same for  $g$ , and so the desired conclusion for  $p$ -ellipses follows from Theorem 4 when  $1 < p < \infty$ .

For  $p = \infty$ , the  $\infty$ -circle is a Euclidean square and the  $\infty$ -ellipse is a rectangle. Many different rectangles of given area can contain the same number of lattice points. For example, a  $4 \times 1$  rectangle and  $2 \times 2$  square each contain 4 lattice points in the first quadrant. All such matters can be handled by the explicit formula  $N(r, s) = \lfloor rs^{-1} \rfloor \lfloor rs \rfloor$  for the counting function when  $p = \infty$ .

The case  $p = 1$  is an open problem, as discussed in Section 9. The case  $0 < p < 1$  has been handled by Ariturk and Laugesen [4] using results in this paper.

Incidentally, an explicit estimate on the number of lattice points in the full  $p$ -ellipse in all four quadrants was obtained by Krätzel [16, Theorem 2] for  $p \geq 2$ . See the informative survey by Ivić *et al.* [14, §3.1].

### Lattice points in the closed first quadrant, and Neumann eigenvalues

Our results have analogues for lattice point counting in the closed (rather than open) first quadrant, as we now explain. When counting nonnegative-integer lattice points, which means we include lattice points on the axes, the counting function for  $r\Gamma(s)$  is

$$\mathcal{N}(r, s) = \#\{(j, k) \in \mathbb{Z}_+ \times \mathbb{Z}_+ : k \leq r s f(j s / r)\},$$

where  $\mathbb{Z}_+ = \{0, 1, 2, 3, \dots\}$ . Define

$$\mathcal{S}(r) = \operatorname{argmin}_{s > 0} \mathcal{N}(r, s).$$

In other words, the set  $\mathcal{S}(r)$  consists of the  $s$ -values that minimize the number of lattice points inside the curve  $r\Gamma(s)$  in the *closed* first quadrant. Notice we employ the calligraphic letters  $\mathcal{N}$  and  $\mathcal{S}$  when working with nonnegative-integer lattice points.

**Theorem 6.** (Uniform bound on optimal stretch factors) *If  $\Gamma$  is a concave, strictly decreasing curve in the first quadrant then*

$$\mathcal{S}(r) \subset [s_1, s_2] \quad \text{for all } r \geq 2/L,$$

for some constants  $s_1, s_2 > 0$ .

**Theorem 7.** (Optimal concave curve is asymptotically balanced) *Under the assumptions of Theorem 3, the optimal stretch factor for minimizing  $\mathcal{N}(r, s)$  approaches 1 as  $r$  tends to infinity:*

$$\begin{aligned} \mathcal{S}(r) &\subset [1 - O(r^{-1/6}), 1 + O(r^{-1/6})], \\ \min_{s > 0} \mathcal{N}(r, s) &= r^2 \operatorname{Area}(\Gamma) + rL + O(r^{2/3}), \end{aligned}$$

and under the assumptions of Theorem 4 we have similarly that:

$$\begin{aligned} \mathcal{S}(r) &\subset [1 - O(r^{-e}), 1 + O(r^{-e})], \\ \min_{s > 0} \mathcal{N}(r, s) &= r^2 \operatorname{Area}(\Gamma) + rL + O(r^{1-2e}). \end{aligned}$$

The proofs are in Section 8.

Consequently, we reprove a recent theorem of van den Berg, Bucur and Gittins [5] saying that the optimal rectangle of area 1 for maximizing the  $n$ -th Neumann eigenvalue of the Laplacian approaches a square as  $n \rightarrow \infty$ . See Section 10 for discussion, and an explanation of why the approach in this paper is simpler.

#### 4. Proof of Theorem 2

To control the stretch factors and hence prove Theorem 2, we will first derive a rough lower bound on the counting function, and then a more sophisticated upper bound. The leading order term in these bounds is simply the area inside the rescaled curve and thus is best possible, while the second term scales like the length of the curve and so at least has the correct order of magnitude.

Assume  $\Gamma$  is concave and decreasing in the first quadrant, with  $x$ - and  $y$ -intercepts at  $L$  and  $M$  respectively. The intercepts need not be equal, in the lemmas and proposition below. Recall that  $N(r, s)$  counts the positive-integer lattice points under the curve  $\Gamma$ , while  $\mathcal{N}(r, s)$  counts nonnegative-integer lattice points.

**Lemma 8.** (Relation between counting functions) *For each  $r, s > 0$ ,*

$$\mathcal{N}(r, s) = N(r, s) + r(s^{-1}L + sM) + \rho(r, s)$$

for some number  $\rho(r, s) \in [-1, 1]$ .

*Proof.* The difference between the two counting functions is simply the number of lattice points lying on the coordinate axes inside the intercepts of  $r\Gamma(s)$ . There are

$$\lfloor rs^{-1}L \rfloor + \lfloor rsM \rfloor + 1$$

such lattice points, and so the lemma follows immediately.  $\square$

**Lemma 9.** (Rough lower bound) *The number  $N(r, s)$  of positive-integer lattice points lying inside  $r\Gamma(s)$  in the first quadrant satisfies*

$$N(r, s) \geq r^2 \text{Area}(\Gamma) - r(s^{-1}L + sM) - 1, \quad r, s > 0.$$

*Proof.* Notice  $\mathcal{N}(r, s)$  equals the total area of the squares of sidelength 1 having lower left vertices at nonnegative-integer lattice points inside the curve  $r\Gamma(s)$ . The union of these squares contains  $r\Gamma(s)$ , since the curve is decreasing. Hence  $\mathcal{N}(r, s) \geq r^2 \text{Area}(\Gamma)$ , and so

$$\begin{aligned} N(r, s) &\geq \mathcal{N}(r, s) - r(s^{-1}L + sM) - 1 \quad \text{by Lemma 8} \\ &\geq r^2 \text{Area}(\Gamma) - r(s^{-1}L + sM) - 1. \quad \square \end{aligned}$$

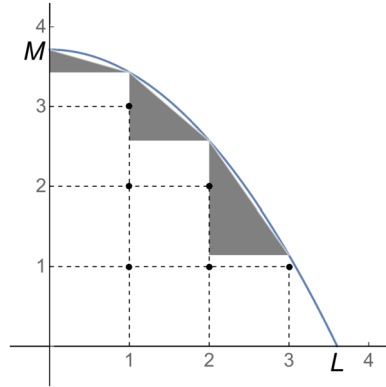


Figure 4. Positive integer lattice count satisfies  $N \leq \text{Area}(\Gamma) - \text{Area}(\text{triangles})$ , in proof of Proposition 10(a).

For the upper bound in the next proposition, remember  $\Gamma$  is the graph of  $y = f(x)$ , where  $f$  is concave and decreasing on  $[0, L]$ , with  $f(0) = M, f(L) = 0$ . We do not assume  $f$  is differentiable in the next result, although in order to guarantee the constant  $C$  in the proposition is positive, we assume  $f$  is *strictly* decreasing.

**Proposition 10.** (Two-term upper bound on counting function) *Let  $C = M - f(L/2)$ .*

(a) *The number  $N$  of positive-integer lattice points lying inside  $\Gamma$  in the first quadrant satisfies*

$$(3) \quad N \leq \text{Area}(\Gamma) - \frac{1}{2}C$$

*provided  $L \geq 1$ .*

(b) *The number  $N(r, s)$  of positive-integer lattice points lying inside  $r\Gamma(s)$  in the first quadrant satisfies*

$$N(r, s) \leq r^2 \text{Area}(\Gamma) - \frac{1}{2}Crs$$

*whenever  $r \geq s/L$ .*

*Proof.* Part (a). Clearly  $N$  equals the total area of the squares of sidelength 1 having upper right vertices at positive-integer lattice points inside the curve  $\Gamma$ . Consider also the right triangles of width 1 formed by secant lines on  $\Gamma$  (see Figure 4), that is, the triangles with vertices  $(i - 1, f(i - 1)), (i, f(i)), (i - 1, f(i))$ , where  $i = 1, \dots, \lfloor L \rfloor$ . These triangles lie above the squares by construction, and lie below  $\Gamma$  by concavity. Hence

$$(4) \quad N + \text{Area}(\text{triangles}) \leq \text{Area}(\Gamma).$$

Since  $f$  is decreasing, we find

$$\begin{aligned} \text{Area}(\text{triangles}) &= \sum_{i=1}^{\lfloor L \rfloor} \frac{1}{2} (f(i-1) - f(i)) \\ (5) \qquad \qquad \qquad &= \frac{1}{2} (f(0) - f(\lfloor L \rfloor)) \end{aligned}$$

$$(6) \qquad \qquad \qquad \geq \frac{1}{2} (M - f(L/2)) = \frac{1}{2} C,$$

because  $\lfloor L \rfloor \geq L/2$  when  $L \geq 1$ . Combining (4) and (6) proves Part (a).

Part (b). Simply replace  $\Gamma$  in Part (a) with the curve  $r\Gamma(s)$ , meaning we replace  $L, M, f(x)$  with  $rs^{-1}L, rsM, rsf(sx/r)$  respectively.  $\square$

**Proof of Theorem 2**

Recall the intercepts are assumed equal ( $L=M$ ) in this theorem. Let  $r \geq 2/L$  and suppose  $s \in S(r)$ . Then  $r \geq s/L$  by Lemma 1, and so the upper bound in Proposition 10(b) gives

$$N(r, s) \leq r^2 \text{Area}(\Gamma) - \frac{1}{2} Crs.$$

The lower bound in Lemma 9 with “ $s=1$ ” says

$$(7) \qquad \qquad \qquad N(r, 1) \geq r^2 \text{Area}(\Gamma) - 2rL - 1.$$

The value  $s \in S(r)$  is a maximizing value, and so  $N(r, 1) \leq N(r, s)$ . The preceding inequalities therefore imply

$$\frac{1}{2} Crs \leq 2rL + 1 \leq \frac{5}{2} rL.$$

Hence  $s \leq 5L/C \equiv s_2$ , and so the set  $S(r)$  is bounded above.

Interchanging the roles of the horizontal and vertical axes, we similarly find  $s^{-1} \leq 5L/\tilde{C} \equiv s_1^{-1}$ , so that the set  $S(r)$  is bounded below away from 0, completing the first part of the proof.

The fact that  $S(r)$  is bounded will help imply an improved bound in the limit as  $r \rightarrow \infty$ . Going back to the proof of Proposition 10(a), we see from (4) and (5) that

$$N + \frac{1}{2} (f(0) - f(\lfloor L \rfloor)) \leq \text{Area}(\Gamma).$$

Rescaling the curve from  $\Gamma$  to  $r\Gamma(s)$ , so that  $N$  and  $f(x)$  become  $N(r, s)$  and  $rsf(\frac{s}{r}x)$ , respectively, and the  $x$ -intercept  $L$  becomes  $rL/s$ , we see the last inequality becomes

$$N(r, s) + \frac{1}{2}rs(f(0) - f(\frac{s}{r}[\frac{rL}{s}])) \leq r^2 \text{Area}(\Gamma).$$

Hence

$$N(r, s) \leq r^2 \text{Area}(\Gamma) - \frac{1}{2}rsL + o(r) \quad \text{as } r \rightarrow \infty,$$

where to get the error term  $o(r)$  we used that  $s \in S(r)$  is bounded above and below ( $s_1 \leq s \leq s_2$ ) and  $f(L) = 0$ . Since  $s$  is a maximizing value we have  $N(r, 1) \leq N(r, s)$ , and so (7) and the above inequality imply

$$\frac{1}{2}rsL + o(r) \leq 2rL + 1,$$

which implies  $\limsup_{r \rightarrow \infty} s \leq 4$ . Similarly  $\limsup_{r \rightarrow \infty} s^{-1} \leq 4$ , by interchanging the axes.

## 5. Two-term counting estimates with explicit remainder

We start with a result for  $C^2$ -smooth curves. What matters in the following proposition is that the right side of estimate (8) below has the form  $O(r^\theta)$  for some  $\theta < 1$ , and that the  $s$ -dependence in the estimate can be seen explicitly. The detailed dependence on the functions  $f$  and  $g$  will not be important for our purposes.

The horizontal and vertical intercepts  $L$  and  $M$  need not be equal, in this section.

**Proposition 11.** (Two-term counting estimate) *Take a point  $(\alpha, \beta) \in \Gamma$  lying in the first quadrant, and assume that  $f \in C^2[0, \alpha]$  with  $f' < 0$  on  $(0, \alpha]$  and  $f'' < 0$  on  $[0, \alpha]$ , and similarly  $g \in C^2[0, \beta]$  with  $g' < 0$  on  $(0, \beta]$  and  $g'' < 0$  on  $[0, \beta]$ . Further suppose  $f''$  is monotonic on  $[0, \alpha]$  and  $g''$  is monotonic on  $[0, \beta]$ .*

(a) *The number  $N$  of positive-integer lattice points inside  $\Gamma$  in the first quadrant satisfies:*

$$\begin{aligned} & |N - \text{Area}(\Gamma) + (L + M)/2| \\ & \leq 6 \left( \int_0^\alpha |f''(x)|^{1/3} dx + \int_0^\beta |g''(y)|^{1/3} dy \right) + 175 \left( \max_{[0, \alpha]} \frac{1}{|f''|^{1/2}} + \max_{[0, \beta]} \frac{1}{|g''|^{1/2}} \right) \\ & \quad + \frac{1}{4} (|f'(\alpha)| + |g'(\beta)|) + 3. \end{aligned}$$

(b) *The number  $N(r, s)$  of positive-integer lattice points lying inside  $r\Gamma(s)$  in the first quadrant satisfies (for  $r, s > 0$ ):*

$$\begin{aligned}
 (8) \quad & |N(r, s) - r^2 \text{Area}(\Gamma) + r(s^{-1}L + sM)/2| \\
 & \leq 6r^{2/3} \left( \int_0^\alpha |f''(x)|^{1/3} dx + \int_0^\beta |g''(y)|^{1/3} dy \right) \\
 & \quad + 175r^{1/2} \left( \max_{[0, \alpha]} \frac{s^{-3/2}}{|f''|^{1/2}} + \max_{[0, \beta]} \frac{s^{3/2}}{|g''|^{1/2}} \right) + \frac{1}{4}(s^2|f'(\alpha)| + s^{-2}|g'(\beta)|) + 3.
 \end{aligned}$$

Proposition 11 and its proof are closely related to work of Krätzel [16, Theorem 1]. We give a direct proof below for two reasons: we want the estimate (8) that depends explicitly on the stretching parameter  $s$ , and we want a proof that can be modified to use a weaker monotonicity hypothesis, in Proposition 12.

A better bound on the right side of (8), giving order  $O(r^{\theta+\varepsilon})$  with  $\theta = 131/208 \approx 0.63 < 2/3$ , can be found in work of Huxley [13], with precursors in [12, Theorems 18.3.2 and 18.3.3]. That bound is difficult to prove, though, and the improvement is not important for our purposes since it leads to only a slight improvement in the rate of convergence for  $S(r)$ , namely from  $O(r^{-1/6})$  to  $O(r^{(\theta+\varepsilon-1)/2})$  in Theorem 3.

*Proof.* Part (a). We divide the region under  $\Gamma$  into three parts. Let  $N_1$  count the lattice points lying to the left of the line  $x = \alpha$  and above  $y = \beta$ , and  $N_2$  count the lattice points to the right of  $x = \alpha$  and below  $y = \beta$ , and  $N_3$  count the lattice points in the remaining rectangle  $(0, \alpha] \times (0, \beta]$ . That is,

$$\begin{aligned}
 N_1 &= \sum_{0 < m \leq \alpha} \sum_{\beta < n \leq f(m)} 1 = \sum_{0 < m \leq \alpha} (\lfloor f(m) \rfloor - \lfloor \beta \rfloor), \\
 N_2 &= \sum_{0 < n \leq \beta} \sum_{\alpha < m \leq g(n)} 1 = \sum_{0 < n \leq \beta} (\lfloor g(n) \rfloor - \lfloor \alpha \rfloor), \\
 N_3 &= \lfloor \alpha \rfloor \lfloor \beta \rfloor.
 \end{aligned}$$

In terms of the *sawtooth* function  $\psi$ , defined by

$$\psi(x) = x - \lfloor x \rfloor - 1/2,$$

one can evaluate

$$N_1 = \sum_{0 < m \leq \alpha} (f(m) - \psi(f(m)) - 1/2 - \lfloor \beta \rfloor).$$

Then we apply the Euler–Maclaurin summation formula

$$\sum_{0 < m \leq \alpha} f(m) = \int_0^\alpha f(x) dx - \psi(\alpha)f(\alpha) + \psi(0)f(0) + \int_0^\alpha f'(x)\psi(x) dx$$

(which we observe for later reference holds whenever  $f$  is piecewise  $C^1$ -smooth) to deduce that

$$\begin{aligned} N_1 &= \int_0^\alpha f(x) \, dx - \psi(\alpha)f(\alpha) + \psi(0)f(0) + \int_0^\alpha f'(x)\psi(x) \, dx \\ &\quad - \sum_{0 < m \leq \alpha} \psi(f(m)) - \lfloor \alpha \rfloor (1/2 + \lfloor \beta \rfloor) \\ &= \int_0^\alpha f(x) \, dx - \psi(\alpha)\beta - M/2 + \int_0^\alpha f'(x)\psi(x) \, dx \\ &\quad - \sum_{0 < m \leq \alpha} \psi(f(m)) - \lfloor \alpha \rfloor (1/2 + \lfloor \beta \rfloor). \end{aligned}$$

Similarly

$$\begin{aligned} N_2 &= \int_0^\beta g(y) \, dy - \psi(\beta)\alpha - L/2 + \int_0^\beta g'(y)\psi(y) \, dy \\ &\quad - \sum_{0 < n \leq \beta} \psi(g(n)) - \lfloor \beta \rfloor (1/2 + \lfloor \alpha \rfloor), \end{aligned}$$

and so

$$\begin{aligned} (9) \quad N &= N_1 + N_2 + N_3 \\ &= \int_0^\alpha f(x) \, dx + \int_0^\beta g(y) \, dy - \lfloor \alpha \rfloor \lfloor \beta \rfloor - (L+M)/2 \\ &\quad - \psi(\alpha)\beta - \lfloor \alpha \rfloor / 2 - \psi(\beta)\alpha - \lfloor \beta \rfloor / 2 \\ &\quad + \int_0^\alpha f'(x)\psi(x) \, dx + \int_0^\beta g'(y)\psi(y) \, dy \\ &\quad - \sum_{0 < m \leq \alpha} \psi(f(m)) - \sum_{0 < n \leq \beta} \psi(g(n)) \\ &= \text{Area}(\Gamma) - (L+M)/2 + \int_0^\alpha f'(x)\psi(x) \, dx + \int_0^\beta g'(y)\psi(y) \, dy \\ &\quad - \sum_{0 < m \leq \alpha} \psi(f(m)) - \sum_{0 < n \leq \beta} \psi(g(n)) + \text{remainder} \end{aligned}$$

where

$$(10) \quad \text{remainder} = -(\alpha - \lfloor \alpha \rfloor)(\beta - \lfloor \beta \rfloor) + (\alpha - \lfloor \alpha \rfloor + \beta - \lfloor \beta \rfloor)/2.$$

This remainder lies between 0 and 1, since  $0 \leq -xy + (x+y)/2 \leq 1$  when  $x, y \in [0, 1]$ .



We estimate the sum of sawtooth functions in (9) by using Theorem 18 (which is due to van der Corput): since  $f''$  is monotonic and nonzero on  $[0, \alpha]$ , the theorem implies

$$(11) \quad \left| \sum_{0 < m \leq \alpha} \psi(f(m)) \right| \leq 6 \int_0^\alpha |f''(x)|^{1/3} dx + 175 \max_{[0, \alpha]} \frac{1}{|f''|^{1/2}} + 1$$

and similarly

$$(12) \quad \left| \sum_{0 < n \leq \beta} \psi(g(n)) \right| \leq 6 \int_0^\beta |g''(y)|^{1/3} dy + 175 \max_{[0, \beta]} \frac{1}{|g''|^{1/2}} + 1.$$

To estimate the integrals of  $f'\psi$  and  $g'\psi$  in (9), we introduce the antiderivative of the sawtooth function,  $\Psi(t) = \int_0^t \psi(z) dz$ , and observe that  $-1/8 \leq \Psi(t) \leq 0$  for all  $t \in \mathbb{R}$ . By integration by parts and the fact that  $f'' < 0$ , we have

$$(13) \quad \begin{aligned} \left| \int_0^\alpha f'(x)\psi(x) dx \right| &= \left| [f'(x)\Psi(x)]_{x=0}^{x=\alpha} - \int_0^\alpha f''(x)\Psi(x) dx \right| \\ &\leq \frac{1}{8}|f'(\alpha)| + \frac{1}{8} \left| \int_0^\alpha f''(x) dx \right| \\ &= \frac{1}{8}|f'(\alpha)| + \frac{1}{8}(f'(0) - f'(\alpha)) \\ &\leq \frac{1}{4}|f'(\alpha)| \end{aligned}$$

since  $f'(\alpha) \leq f'(0) \leq 0$ . The same argument gives

$$(14) \quad \left| \int_0^\beta g'(y)\psi(y) dy \right| \leq \frac{1}{4}|g'(\beta)|.$$

Combining (9)–(14) completes the proof of Part (a).

Part (b). Simply apply Part (a) to the curve  $r\Gamma(s)$  by replacing  $L, M, f(x), g(y), \alpha, \beta$  with  $rs^{-1}L, rsM, rsf(sx/r), rs^{-1}g(s^{-1}y/r), rs^{-1}\alpha, rs\beta$  respectively.  $\square$

*Remark* Proposition 11 continues to hold if the point  $(\alpha, \beta) = (L, 0)$  lies at the right endpoint of the curve. One simply removes all mention of  $g, \beta_j$  and  $\varepsilon$  from the hypotheses of the proposition, and removes all such terms from the conclusions, as can be justified by inspecting the proof above. The same remark holds for the advanced counting estimate in the following Proposition 12.

### Advanced counting estimate

The hypotheses in the last result are somewhat restrictive. In particular, we would like to handle infinite curvature at the intercepts of the curve  $\Gamma$ , meaning  $f''$

must be allowed to blow up at  $x=0$ . Further, we would like to relax the monotonicity assumption on  $f''$ . The next result achieves these goals.

Two numbers  $\delta$  and  $\varepsilon$  appear in the next Proposition. Their role in the proof is that on the intervals  $0 < x \leq \delta$  and  $0 < y \leq \varepsilon$  we bound the sawtooth function trivially with  $|\psi| \leq 1/2$ . On the remaining intervals we seek cancellations.

**Proposition 12.** (Two-term counting estimate for more general curve) *Take a point  $(\alpha, \beta) \in \Gamma$  lying in the first quadrant, and assume  $f \in PC^2(0, \alpha]$  with  $f' < 0$  and  $f'' < 0$ , and that  $f''$  is monotonic on  $(\alpha_{i-1}, \alpha_i]$  for  $i=1, \dots, l$ . Similarly assume  $g \in PC^2(0, \beta]$  with  $g' < 0$  and  $g'' < 0$ , and that  $g''$  is monotonic on  $(\beta_{j-1}, \beta_j]$  for  $j=1, \dots, \ell$ .*

(a) *If  $\delta \in (0, \alpha)$  and  $\varepsilon \in (0, \beta)$  then the number  $N$  of positive-integer lattice points inside  $\Gamma$  in the first quadrant satisfies:*

$$\begin{aligned} & |N - \text{Area}(\Gamma) + (L+M)/2| \\ & \leq 6 \left( \int_0^\alpha |f''(x)|^{1/3} dx + \int_0^\beta |g''(y)|^{1/3} dy \right) \\ & \quad + 175 \left( \frac{1}{|f''(\delta)|^{1/2}} + \frac{1}{|g''(\varepsilon)|^{1/2}} \right) + 350 \left( \sum_{i=1}^l \frac{1}{|f''(\alpha_i)|^{1/2}} + \sum_{j=1}^\ell \frac{1}{|g''(\beta_j)|^{1/2}} \right) \\ & \quad + \frac{1}{4} \left( \sum_{i=1}^l |f'(\alpha_i)| + \sum_{j=1}^\ell |g'(\beta_j)| \right) + \frac{1}{2} (\delta + \varepsilon) + l + \ell + 1. \end{aligned}$$

(b) *If functions*

$$\delta : (0, \infty) \longrightarrow (0, \alpha), \quad \varepsilon : (0, \infty) \longrightarrow (0, \beta),$$

*are given, then the number  $N(r, s)$  of positive-integer lattice points inside  $r\Gamma(s)$  in the first quadrant satisfies (for  $r, s > 0$ ):*

$$\begin{aligned} (15) \quad & |N(r, s) - r^2 \text{Area}(\Gamma) + r(s^{-1}L + sM)/2| \\ & \leq 6r^{2/3} \left( \int_0^\alpha |f''(x)|^{1/3} dx + \int_0^\beta |g''(y)|^{1/3} dy \right) \\ & \quad + 175r^{1/2} \left( \frac{s^{-3/2}}{|f''(\delta(r))|^{1/2}} + \frac{s^{3/2}}{|g''(\varepsilon(r))|^{1/2}} \right) \\ & \quad + 350r^{1/2} \left( \sum_{i=1}^l \frac{s^{-3/2}}{|f''(\alpha_i)|^{1/2}} + \sum_{j=1}^\ell \frac{s^{3/2}}{|g''(\beta_j)|^{1/2}} \right) \\ & \quad + \frac{1}{4} \left( \sum_{i=1}^l s^2 |f'(\alpha_i)| + \sum_{j=1}^\ell s^{-2} |g'(\beta_j)| \right) + \frac{r}{2} (s^{-1}\delta(r) + s\varepsilon(r)) + l + \ell + 1. \end{aligned}$$

The integral of  $|f''|^{1/3}$  appearing in the conclusion of Proposition 12 is finite, because by Hölder's inequality and the fact that  $f'' < 0$  and  $f$  is decreasing, we have

$$\int_0^{\alpha_1} |f''(x)|^{1/3} dx \leq \alpha_1^{2/3} \left| \int_0^{\alpha_1} f''(x) dx \right|^{1/3} = \alpha_1^{2/3} |f'(0^+) - f'(\alpha_1^-)|^{1/3} < \infty.$$

The integral of  $|g''|^{1/3}$  is finite for similar reasons.

*Proof.* Part (a). The lattice point counting (9) holds just as in the proof of Proposition 11, and so the task is to estimate each of the terms on the right side of that equation.

Estimate (11) on the sum of the sawtooth function is no longer valid, because  $f''$  is no longer assumed to be monotonic on the whole interval  $[0, \alpha]$ . To control this sawtooth sum, we first observe

$$\left| \sum_{0 < m \leq \delta} \psi(f(m)) \right| \leq \frac{1}{2} \delta$$

since  $|\psi| \leq 1/2$  everywhere. Next, we have  $\delta \in (\alpha_{j-1}, \alpha_j]$  for some  $j \in \{1, \dots, l\}$ , and

$$\left| \sum_{\delta < m \leq \alpha_j} \psi(f(m)) \right| \leq 6 \int_{\delta}^{\alpha_j} |f''(x)|^{1/3} dx + 175 \max \left\{ \frac{1}{|f''(\delta)|^{1/2}}, \frac{1}{|f''(\alpha_j)|^{1/2}} \right\} + 1$$

by Theorem 18 applied on the interval  $[\delta, \alpha_j]$ . Applying that theorem again on each interval  $[\alpha_{i-1}, \alpha_i]$  with  $i = j+1, \dots, l$  gives that

$$\begin{aligned} \left| \sum_{\alpha_{i-1} < m \leq \alpha_i} \psi(f(m)) \right| &\leq 6 \int_{\alpha_{i-1}}^{\alpha_i} |f''(x)|^{1/3} dx \\ &\quad + 175 \max \left\{ \frac{1}{|f''(\alpha_{i-1})|^{1/2}}, \frac{1}{|f''(\alpha_i)|^{1/2}} \right\} + 1. \end{aligned}$$

By summing the last three displayed inequalities, we deduce a sawtooth bound

(16)

$$\begin{aligned} &\left| \sum_{0 < m \leq \alpha} \psi(f(m)) \right| \\ &\leq \frac{1}{2} \delta + 6 \int_{\delta}^{\alpha} |f''(x)|^{1/3} dx + \frac{175}{|f''(\delta)|^{1/2}} + \sum_{i=j}^{l-1} \frac{350}{|f''(\alpha_i)|^{1/2}} + \frac{175}{|f''(\alpha)|^{1/2}} + l - j + 1 \\ &\leq \frac{1}{2} \delta + 6 \int_0^{\alpha} |f''(x)|^{1/3} dx + \frac{175}{|f''(\delta)|^{1/2}} + \sum_{i=1}^l \frac{350}{|f''(\alpha_i)|^{1/2}} + l. \end{aligned}$$

Next, we adapt estimate (13) on the integral of  $f'\psi$  by simply applying the same argument on each interval  $[\alpha_{i-1}, \alpha_i]$ , hence finding

$$(17) \quad \left| \int_0^\alpha f'(x)\psi(x) dx \right| \leq \sum_{i=1}^l \left[ \frac{1}{8} |f'(\alpha_i)| + \frac{1}{8} (f'(\alpha_{i-1}) - f'(\alpha_i)) \right] \\ \leq \frac{1}{4} \sum_{i=1}^l |f'(\alpha_i)|.$$

By combining (9), (10) with (16), (17) and the analogous estimates on  $g$ , we complete the proof of Part (a).

Part (b). Apply Part (a) to the curve  $r\Gamma(s)$  by replacing  $L, M, f(x), g(y), \alpha, \beta, \delta, \varepsilon$  with  $rs^{-1}L, rsM, rsf(sx/r), rs^{-1}g(s^{-1}y/r), rs^{-1}\alpha, rs\beta, rs^{-1}\delta(r), rs\varepsilon(r)$  respectively.  $\square$

## 6. A unified approach

The next proposition provides a unified framework for proving our theorems later in the paper. It adapts the scheme of proof employed by Antunes and Freitas [2].

**Proposition 13.** *Let  $A \in \mathbb{R}, L > 0$ , and  $0 < \theta < 1$ . Consider a real valued function  $H(r, s)$  (for  $r, s > 0$ ) such that for each closed interval  $[s_1, s_2] \subset (0, \infty)$  one has*

$$(18) \quad H(r, s) = Ar^2 - Lr(s + s^{-1})/2 + O(r^\theta),$$

*with  $s \in [s_1, s_2]$  allowed to vary as  $r \rightarrow \infty$ . Assume the function  $s \mapsto H(r, s)$  attains its maximum value, for each  $r > 0$ , and write  $S(r) = \operatorname{argmax}_{s > 0} H(r, s)$  for the set of maximizing points. Suppose*

$$(19) \quad S(r) \subset [s_1, s_2] \quad \text{for all large } r > 0,$$

*for some constants  $s_1, s_2 > 0$ .*

*Then the maximizing set  $S(r)$  converges to the point  $\{1\}$  as  $r \rightarrow \infty$ , with*

$$S(r) \subset [1 - O(r^{-(1-\theta)/2}), 1 + O(r^{-(1-\theta)/2})],$$

*and the maximum value of  $H$  has asymptotic formula*

$$\max_{s > 0} H(r, s) = Ar^2 - Lr + O(r^\theta).$$

The error term  $O(r^\theta)$  in (18) has implied constant depending on the interval  $[s_1, s_2]$ .

*Proof.* Since  $S(r) \subset [s_1, s_2]$  by hypothesis (19), the asymptotic estimate (18) implies

$$\begin{aligned} H(r, s) &= Ar^2 - Lr(s + s^{-1})/2 + O(r^\theta), \\ H(r, 1) &= Ar^2 - Lr + O(r^\theta), \end{aligned}$$

for  $s \in S(r)$  and  $r \rightarrow \infty$ . Since  $s$  is a maximizing value, we have  $H(r, 1) \leq H(r, s)$  and so

$$(20) \quad s + s^{-1} \leq 2 + O(r^{-(1-\theta)}).$$

Hence  $s = 1 + O(r^{-(1-\theta)/2})$  by Lemma 19, which proves the first claim in the theorem. For the second claim, when  $s \in S(r)$  we have  $H(r, s) = Ar^2 - Lr + O(r^\theta)$  as  $r \rightarrow \infty$ , by (18) and using also that  $1 \leq (s + s^{-1})/2 \leq 1 + O(r^{-(1-\theta)})$  by (20).  $\square$

### 7. Proof of Theorem 3 and Theorem 4

#### Proof of Theorem 3

The theorem follows directly from Proposition 13 with  $H(r, s)$  being the lattice counting function  $N(r, s)$ . The hypotheses of the proposition are verified as follows.

Suppose  $0 < s_1 < s_2 < \infty$ . By Proposition 11(b) with  $L = M$  one has

$$(21) \quad N(r, s) = \text{Area}(\Gamma)r^2 - Lr(s + s^{-1})/2 + O(r^{2/3}),$$

with  $s \in [s_1, s_2]$  as  $r \rightarrow \infty$ . Thus hypothesis (18) holds for  $N(r, s)$  with the choices  $A = \text{Area}(\Gamma)$ ,  $\theta = 2/3$ , and  $L$  equalling the intercept value of  $\Gamma$ .

The boundedness hypothesis (19) holds by Theorem 2.

#### Proof of Theorem 4

Again let  $H(r, s)$  be the lattice counting function  $N(r, s)$ , take  $A = \text{Area}(\Gamma)$ , let  $L$  be the intercept value of  $\Gamma$ , and note the boundedness hypothesis (19) holds by Theorem 2. To finish verifying the hypotheses of Proposition 13, we suppose  $0 < s_1 < s_2 < \infty$  and show that (18) holds.

Take  $\theta = 1 - 2e$ , where the number  $e = \min\{\frac{1}{6}, a_1, a_2, b_1, b_2\}$  was defined in Theorem 4. Hypothesis (18) is the assertion that

$$(22) \quad N(r, s) = \text{Area}(\Gamma)r^2 - Lr(s + s^{-1})/2 + O(r^{1-2e}),$$

with  $s \in [s_1, s_2]$  as  $r \rightarrow \infty$ . To verify this asymptotic, we will estimate the remainder terms in Proposition 12(b) as follows. In that proposition take  $L = M$ , and note

$\delta(r) < \alpha$  and  $\varepsilon(r) < \beta$  for all large  $r$  by assumptions (1) and (2). We will show the right side of estimate (15) in Proposition 12(b) is bounded by

$$O(r^{2/3}) + s^{-3/2}O(r^{1-2a_2}) + s^{3/2}O(r^{1-2b_2}) + (s^{-3/2} + s^{3/2})O(r^{1/2}) + (s^2 + s^{-2})O(1) + s^{-1}O(r^{1-2a_1}) + sO(r^{1-2b_1}) + O(1)$$

for large enough  $r$ , where the implied constants in the  $O(\cdot)$ -terms depend only on the curve  $\Gamma$  and are independent of  $s$ . Since each one of these  $O(\cdot)$ -terms is bounded by  $O(r^{1-2e})$ , and  $s$  and  $s^{-1}$  are bounded when  $s \in [s_1, s_2]$ , hypothesis (18) will hold as desired.

Examining now the right side of (15), we see the first two terms are obviously  $O(r^{2/3})$ . For the next term, observe by assumption in (1) that

$$\frac{r^{1/2}s^{-3/2}}{|f''(\delta(r))|^{1/2}} = s^{-3/2}O(r^{1-2a_2}),$$

and similarly for the analogous term involving  $g''$ . Since  $f''(\alpha_i)$  and  $g''(\beta_j)$  are constant, the corresponding terms in (15) can be estimated by  $(s^{-3/2} + s^{3/2})O(r^{1/2})$ . Similarly, the terms in (15) involving  $f'(\alpha_i)$  and  $g'(\beta_j)$  can be estimated by  $(s^2 + s^{-2})O(1)$ . Next,  $s^{-1}r\delta(r) = s^{-1}O(r^{1-2a_1})$  by the assumption in (1), and similarly for  $\varepsilon(r)$ . And, of course,  $l + \ell + 1$  is constant, which completes the verification of hypothesis (18).

### 8. Proof of Theorem 6 and Theorem 7

First we need a two-term bound on the counting function in the closed first quadrant, as provided by the next proposition. The result is an analogue of Proposition 10, although the constant  $\mathcal{C}$  is slightly different than in that result.

Assume  $f$  is concave and strictly decreasing on  $[0, L]$ , with  $f(0) = M, f(L) = 0$ . The intercepts  $L$  and  $M$  need not be equal.

**Proposition 14.** (Two-term lower bound on counting function) *Let  $\mathcal{C} = M - f(L/4)$ .*

(a) *The number  $\mathcal{N}$  of nonnegative-integer lattice points lying inside  $\Gamma$  in the closed first quadrant satisfies:*

$$(23) \quad \mathcal{N} \geq \text{Area}(\Gamma) + \frac{1}{2}\mathcal{C}.$$

(b) *The number of nonnegative-integer lattice points lying inside  $r\Gamma(s)$  in the closed first quadrant satisfies (for  $r, s > 0$ ):*

$$\mathcal{N}(r, s) \geq r^2 \text{Area}(\Gamma) + \frac{1}{2}\mathcal{C}rs.$$

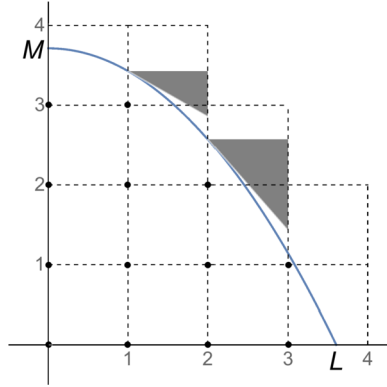


Figure 5. Nonnegative integer lattice count satisfies  $\mathcal{N} \geq \text{Area}(\Gamma) + \text{Area}(\text{triangles})$ , in proof of Proposition 14(a) when  $L \geq 2$ .

*Proof.* Part (a). Clearly  $\mathcal{N}$  equals the total area of the squares of sidelength 1 having lower left vertices at nonnegative-integer lattice points inside the curve  $\Gamma$ . The union of these squares contains  $\Gamma$ , since the curve is decreasing.

We separate the proof into cases according to the value of  $L$ .

*Case (i)* Suppose  $L \leq 2$ , so that  $L/4 \leq 1/2$ . Consider a rectangle whose lower left vertex sits on the curve at  $x=L/4$ , and has vertices

$$(L/4, f(L/4)), \quad (1, f(L/4)), \quad (1, M), \quad (L/4, M).$$

By construction, this rectangle lies inside the union of squares of sidelength 1, and it lies above  $\Gamma$  because the curve is decreasing. Hence

$$\begin{aligned} \mathcal{N} &\geq \text{Area}(\Gamma) + \text{Area}(\text{rectangle}) \\ &= \text{Area}(\Gamma) + (1 - L/4)(M - f(L/4)) \\ &\geq \text{Area}(\Gamma) + \frac{1}{2}(M - f(L/4)) \end{aligned}$$

as desired.

*Case (ii)* Suppose  $L \geq 2$ . Consider the right triangles of width 1 formed by tangent lines from the right on  $\Gamma$ , that is, the triangles with vertices  $(i, f(i)), (i+1, f(i)), (i+1, f(i) + f'(i^+))$ , where  $i=0, 1, \dots, \lfloor L \rfloor - 1$ . These triangles all lie above the horizontal axis, since by concavity  $f(i) + f'(i^+) \geq f(i+1) \geq 0$ ; the last inequality explains why the biggest  $i$ -value we consider is  $\lfloor L \rfloor - 1$ .

Thus these triangles lie inside the union of squares of sidelength 1, and lie above  $\Gamma$  by concavity. Hence

$$\mathcal{N} \geq \text{Area}(\Gamma) + \text{Area}(\text{triangles}).$$

To complete the proof of Case (ii), we estimate

$$\begin{aligned}
 \text{Area}(\text{triangles}) &\geq \frac{1}{2} \sum_{i=1}^{\lfloor L \rfloor - 1} |f'(i^+)| \\
 &\geq \frac{1}{2} \sum_{i=1}^{\lfloor L \rfloor - 1} (f(i-1) - f(i)) \quad \text{by concavity} \\
 &= \frac{1}{2} (f(0) - f(\lfloor L \rfloor - 1)) \\
 &\geq \frac{1}{2} (M - f(L/4)),
 \end{aligned}$$

because  $\lfloor L \rfloor - 1 \geq L/2 \geq L/4$  when  $L \geq 2$ .

Part (b). Replace  $\Gamma$  in Part (a) with the curve  $r\Gamma(s)$ , meaning we replace  $L, M, f(x)$  with  $rs^{-1}L, rsM, rsf(sx/r)$  respectively.  $\square$

### Proof of Theorem 6

Since  $N(r, s) \leq r^2 \text{Area}(\Gamma)$ , taking  $s=1$  and  $L=M$  in Lemma 8 gives that

$$\mathcal{N}(r, 1) \leq r^2 \text{Area}(\Gamma) + 2rL + 1.$$

Now suppose  $s \in \mathcal{S}(r)$  is a minimizing value, so that  $\mathcal{N}(r, s) \leq \mathcal{N}(r, 1)$ . Since

$$\mathcal{N}(r, s) \geq r^2 \text{Area}(\Gamma) + \frac{1}{2} \mathcal{C}rs$$

by Proposition 14(b), we conclude from above that

$$\frac{1}{2} \mathcal{C}rs \leq 2rL + 1 \leq \frac{5}{2} rL,$$

where the last inequality holds for  $r \geq 2/L$ . Hence  $s \leq 5L/\mathcal{C}$ , and so the set  $\mathcal{S}(r)$  is bounded above. Interchanging the horizontal and vertical axes and recalling  $L=M$  (*i.e.*, the intercepts are equal in this theorem), one finds similarly that  $s^{-1} \leq 5L/\tilde{\mathcal{C}}$ . Hence  $\mathcal{S}(r)$  is bounded below away from 0, which completes the proof.

### Proof of Theorem 7

The theorem will follow from Proposition 13 with the choice  $H(r, s) = -\mathcal{N}(r, s)$ , since maximizing  $s \mapsto H(r, s)$  corresponds to minimizing  $s \mapsto \mathcal{N}(r, s)$ . The boundedness hypothesis (19) of the proposition holds by Theorem 6. The other hypothesis (18) is verified as follows.



Taking  $L=M$  in the relation between  $\mathcal{N}(r, s)$  and  $N(r, s)$  in Lemma 8, and calling on the asymptotic for  $N(r, s)$  in either (21) (under the assumptions of Theorem 3) or (22) (under the assumptions of Theorem 4), we deduce

$$\mathcal{N}(r, s) = \text{Area}(\Gamma)r^2 + Lr(s^{-1} + s)/2 + O(r^\theta)$$

with  $s \in [s_1, s_2]$  allowed to vary as  $r \rightarrow \infty$ , where

$$\theta = \begin{cases} 2/3 & \text{under the assumptions of Theorem 3,} \\ 1-2e & \text{under the assumptions of Theorem 4.} \end{cases}$$

That is, we have verified hypothesis (18) with  $H = -\mathcal{N}$ ,  $A = -\text{Area}(\Gamma)$ , and  $L$  the intercept value of  $\Gamma$ .

### 9. Open problem for 1-ellipses — lattice points in right triangles

Lattice point maximization for right triangles appears to be an open problem. Consider the  $p$ -circle with  $p=1$ , which is a diamond with vertices at  $(\pm 1, 0)$  and  $(0, \pm 1)$ . It intersects the first quadrant in the line segment  $\Gamma$  joining the points  $(0, 1)$  and  $(1, 0)$ . Here  $L=M=1$ . Stretching the 1-circle in the  $x$ - and  $y$ -directions gives a 1-ellipse

$$|sx| + |s^{-1}y| = 1,$$

which together with the coordinate axes forms a right triangle of area  $1/2$  in the first quadrant, with one vertex at the origin and hypotenuse  $\Gamma(s)$  joining the vertices at  $(s^{-1}, 0)$  and  $(0, s)$ . As previously, we write  $S(r)$  for the set of  $s$ -values that maximize the number of positive-integer (first quadrant) lattice points below or on  $r\Gamma(s)$ , when  $r > 0$ .

First of all, the 45–45–90 degree triangle ( $s=1$ ) does not always enclose the most lattice points: Figure 6 shows an example.

The open problem is to understand the limiting behavior of the maximizing  $s$ -values. Does  $S(r)$  converge to  $\{1\}$  as  $r \rightarrow \infty$ ? We proved the answer is “Yes” for  $p$ -ellipses when  $1 < p < \infty$  (Example 5), but for  $p=1$  we suggest the answer is “No”. Numerical evidence in Figure 7 suggests that the set  $S(r)$  does not converge to  $\{1\}$  as  $r \rightarrow \infty$ . Indeed, the plotted heights appear to cluster at a large number of values, possibly dense in some interval around  $s=1$ . These cluster values presumably have some number theoretic significance.

In the remainder of the section we remark that maximizing  $s$ -values are  $\leq 3$  in the limit as  $r \rightarrow \infty$ , and we describe the numerical scheme that generates Figure 7. Lastly, we explain why  $s = \sqrt{2}$  is a good candidate for a cluster value as  $r \rightarrow \infty$ .

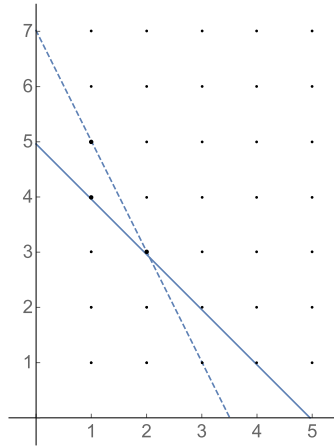


Figure 6. The 1-ellipse  $sx + s^{-1}y = r$  with  $r = 4.96$ , for  $s = 1$  (solid) and  $s = \sqrt{2}$  (dashed). The dashed line encloses three more lattice points (shown in bold) than the solid line.

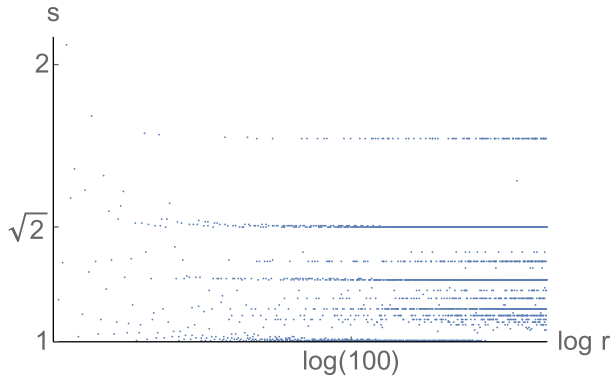


Figure 7. Optimal  $s$ -values for maximizing the number of lattice points in the 1-ellipse (triangle). The graph plots  $\sup S(r)$  versus  $\log r$ . The plotted  $r$ -values are multiples of  $\sqrt{3}/10$ , and the horizontal axis is at height  $s = 1$ .

**The bound on maximizing  $s$ -values for right triangles ( $p = 1$ )**

Given  $\varepsilon > 0$ , we claim

$$S(r) \subset \left[ \frac{1}{3+\varepsilon}, 3+\varepsilon \right] \quad \text{for all large } r.$$

This bound is slightly better than the one in Theorem 2 (which had 4 instead of 3), and can be proved in the same way with the help of a special formula for  $N(r, 1)$ :

$$\begin{aligned}
 (24) \quad N(r, 1) &= \# \text{ first-quadrant lattice points under the line } y = r - x \\
 &= \frac{1}{2} [r] [r - 1] \\
 &\geq \frac{1}{2} (r - 1)(r - 2) = \frac{1}{2} r^2 - \frac{3}{2} r + 1.
 \end{aligned}$$

**How can one efficiently maximize the lattice counting function for the 1-ellipse?**

A brute force method of counting how many lattice points lie under the line  $r\Gamma(s)$ , and then varying  $s$  to maximize that number of lattice points, is simply unworkable in practice. The counting function  $N(r, s)$  jumps up and down in value as  $s$  varies, sometimes jumping quite rapidly, and a brute force method of sampling at a finite collection of  $s$ -values can never be expected to capture all such jump points or their precise locations.

Instead, for a given  $r$  we should pre-identify the possible jump values of  $s$ , and use that information to count the lattice points. We start with the simple observation that a lattice point  $(j, k)$  lies under the line  $r\Gamma(s)$  if and only if

$$sj + s^{-1}k \leq r,$$

which is equivalent to

$$(25) \quad js^2 - rs + k \leq 0.$$

For this quadratic inequality to have a solution, the discriminant must be non-negative,  $r^2 - 4jk \geq 0$ , and thus we need only consider lattice points beneath the hyperbola  $r^2 = 4xy$ . For each such lattice point, equality holds in (25) for two positive  $s$ -values, namely

$$s_{\min}(j, k; r) = \frac{r - \sqrt{r^2 - 4jk}}{2j}, \quad s_{\max}(j, k; r) = \frac{r + \sqrt{r^2 - 4jk}}{2j}.$$

The geometrical meaning of these values can be understood, as follows: as  $s$  increases from 0 to  $\infty$ , one endpoint of the line segment  $r\Gamma(s)$  slides up on the  $y$ -axis while the other endpoint moves left on the  $x$ -axis. The line segment passes through the point  $(j, k)$  twice: first when  $s = s_{\min}(j, k; r)$  and again when  $s = s_{\max}(j, k; r)$ . The point  $(j, k)$  lies below the line when  $s$  belongs to the closed interval between these two values.

Thus the counting function is

$$\begin{aligned}
 N(r, s) &= \# \{ (j, k) : s_{\min}(j, k; r) \leq s \leq s_{\max}(j, k; r) \} \\
 &= \sum_{j, k > 0} \mathbb{1}_{s_{\min}(j, k; r) \leq s} - \sum_{j, k > 0} \mathbb{1}_{s_{\max}(j, k; r) < s}
 \end{aligned}$$

where we sum only over positive-integer lattice points with  $4jk \leq r^2$ .

The last formula says that the counting function  $N(r, s)$  equals the number of values  $s_{\min}(j, k; r)$  that are less than or equal to  $s$  minus the number of values  $s_{\max}(j, k; r)$  that are less than  $s$ . To facilitate the evaluation in practice, one should sort the list of values of  $s_{\min}(j, k; r)$  into increasing order, and similarly sort the list of values of  $s_{\max}(j, k; r)$ . The numbers in these two lists are the only numbers where  $N(r, s)$  can change value, as  $s$  increases. In particular, when  $s$  increases to  $s_{\min}(j, k; r)$ , the point  $(j, k)$  is picked up by the line segment for the first time and so  $N(r, s)$  increases by 1. When  $s$  increases strictly beyond  $s_{\max}(j, k; r)$ , the point  $(j, k)$  is dropped by the line segment and so  $N(r, s)$  decreases by 1. Note the counting function might increase or decrease by more than 1 at some  $s$ -values, if the sorted lists of  $s_{\min}$  and  $s_{\max}$  values have repeated entries (arising from lattice points that are picked up by, or else dropped by, the line segment at the same  $s$ -value).

After sorting the  $s_{\min}$  and  $s_{\max}$  lists, we evaluate the maximum of  $N(r, s)$  by scanning through the two lists, increasing a counter by 1 at each number in the sorted  $s_{\min}$  list, and decreasing the counter just after each number in the sorted  $s_{\max}$  list. The largest value achieved by the counter is the maximum of  $N(r, s)$ , and  $S(r)$  consists of the closed interval or intervals of  $s$ -values on which this maximum count is attained.

By this method, we can maximize the lattice counting function for the 1-ellipse in a computationally efficient manner, for any given  $r > 0$ . The code is available in [20, Appendix B].

When presenting the results of this method graphically, in Figure 7, we plot only the largest  $s$  value in  $S(r)$ , because the family of 1-ellipses is invariant under the map  $s \mapsto 1/s$  and so the smallest value in  $S(r)$  will be just the reciprocal of the largest value.

### Why is the 1-ellipse not covered by our theorems?

For the  $p$ -ellipse with  $p=1$ , Theorem 4 does not apply because  $f$  is linear and so  $f'' \equiv 0$ . Specifically, in the proof we see inequalities (11) and (12) are no longer useful, since their right sides are infinite. The situation cannot easily be rescued, because the left side of (8) need not even be  $o(r)$ . For example, when  $s=1$  and  $r$  is an integer, by evaluating the number  $N(r, 1)$  of lattice points under the curve  $y=r-x$  we find

$$N(r, 1) - r^2 \text{Area}(\Gamma) + r(L+M)/2 = \frac{1}{2}r(r-1) - \frac{1}{2}r^2 + r = \frac{1}{2}r,$$

which is of order  $r$  and hence has the same order as the ‘‘boundary term’’  $r(L+M)/2$  on the left side. Thus the method breaks down completely for  $p=1$ . We seek instead to illuminate the situation through numerical investigations.

**A cluster value at  $s=\sqrt{2}$ ?**

Inspired by the numerical calculations in Figure 7, we will show that  $s=\sqrt{2}$  gives a substantially higher count of lattice points than  $s=1$ , for a certain sequence of  $r$ -values tending to infinity. This observation suggests (but does not prove) that  $\sqrt{2}$  or some number close to it should belong to  $S(r)$  for those  $r$ -values. To be clear: we have not found a proof of this claim. Doing so would provide a counterexample to the idea that the set  $S(r)$  converges to  $\{1\}$  as  $r\rightarrow\infty$ .

To compare the counting functions for  $s=1$  and  $s=\sqrt{2}$ , we first notice that for  $s=1$  the counting function for the 1-circle is given by

$$N(r, 1) = \lfloor r \rfloor \lfloor r-1 \rfloor / 2, \quad r > 0.$$

At  $s=\sqrt{2}$  the slope of the 1-ellipse is  $-2$ , and for the special choice  $r=\sqrt{2}(m+1/2)$  with  $m\geq 1$  the counting function can be evaluated explicitly as

$$N(r, \sqrt{2}) = m^2.$$

We further choose  $m$  such that  $r\in(n-1/4, n)$  for some integer  $n$ , noting that an increasing sequence of such  $m$ -values can be found due to the density in the unit interval of multiples of  $\sqrt{2} \pmod 1$ . Then, writing  $r=n-\varepsilon$  where  $\varepsilon<1/4$ , we have

$$\begin{aligned} N(r, \sqrt{2}) - N(r, 1) &= m^2 - (n-1)(n-2)/2 \\ &= \frac{1}{2}(r^2 - \sqrt{2}r + 1/2) - \frac{1}{2}(r+\varepsilon-1)(r+\varepsilon-2) \\ &\geq \frac{1}{2}r - (\text{constant}). \end{aligned}$$

Hence  $\limsup_{r\rightarrow\infty} (N(r, \sqrt{2}) - N(r, 1))/r \geq 1/2$ , and so  $s=\sqrt{2}$  can give (for certain choices of  $r$ ) a substantially higher count of lattice points than  $s=1$ , as we wanted to show.

The work above implies that  $1\notin S(r)$  for a sequence of  $r$ -values tending to infinity. More generally, Marshall and Steinerberger showed that if  $x>0$  is rational then  $\sqrt{x}\notin S(r)$  for a sequence of  $r$ -values tending to infinity (see [22, Theorem 1]), while if  $x>0$  is irrational then  $\sqrt{x}\notin S(r)$  for all sufficiently large  $r$  (see [22, Lemma 2] and its associated discussion).

**Conjecture for  $p=1$**

To finish the chapter, we state some of our numerical observations as a conjecture. Let

$$S = \{(r, s) : r > 0, s \in S(r)\} \subset (0, \infty) \times (0, \infty),$$

$$\begin{aligned}\bar{S} &= \text{closure of } S \text{ in } [0, \infty) \times [0, \infty), \\ S(\infty) &= \{s \in [0, \infty) : (\infty, s) \in \bar{S}\}.\end{aligned}$$

Earlier in the chapter we proved that  $S(\infty) \subset [1/3, 3]$ .

The clustering behavior of  $S(r)$  observed in Figure 7 suggests the following conjecture.

**Conjecture 15.** ( $p=1$ ) *The limiting set  $S(\infty)$  is countably infinite, and is contained in*

$$[1/3, 3] \cap \{\sqrt{x} : x \in \mathbb{Q}, x > 0\}.$$

Marshall and Steinerberger [22, Theorem 2] recently proved that  $S(\infty)$  contains (countably) infinitely many square roots of rational numbers and is contained in  $[1/\sqrt{5}, \sqrt{5}]$ . For example, they showed the set  $S(\infty)$  contains 1 and  $\sqrt{3/2}$ . Yet a precise characterization of the set remains elusive. One would like a characterization in terms of some number theoretic condition.

## 10. Connection between counting function maximization and eigenvalue minimization

Maximizing a counting function is morally equivalent to minimizing the size of the things being counted. Let us apply this general principle to the case of the circle

$$\Gamma : x^2 + y^2 = 1 \quad \text{in the first quadrant,}$$

and its associated ellipses  $r\Gamma(s)$ . In this section,  $L=M=1$  and  $\text{Area}(\Gamma)=\pi/4$ .

### Minimizing eigenvalues of the Dirichlet Laplacian on rectangles

Write

$$(26) \quad \{\lambda_n(s) : n = 1, 2, 3, \dots\} = \{(js)^2 + (ks^{-1})^2 : j, k = 1, 2, 3, \dots\}$$

so that  $\lambda_n(s)$  is the  $n$ th eigenvalue of the Dirichlet Laplacian on a rectangle of side lengths  $s^{-1}\pi$  and  $s\pi$ . (The eigenfunctions have the form  $\sin(jsx) \sin(ks^{-1}y)$ .) Then the lattice point counting function is the eigenvalue counting function, because

$$\begin{aligned}N(r, s) &= \#\{(j, k) : (js)^2 + (ks^{-1})^2 \leq r^2\} \\ &= \#\{n : \lambda_n(s) \leq r^2\}.\end{aligned}$$

Define

$$S_*(n) = \operatorname{argmin}_{s>0} \lambda_n(s),$$

so that  $S_*(n)$  is the set of  $s$ -values that minimize the  $n$ th eigenvalue.

The next result says that the rectangle minimizing the  $n$ th eigenvalue will converge to a square as  $n \rightarrow \infty$ .

**Corollary 16.** (Optimal Dirichlet rectangle is asymptotically balanced, due to Antunes and Freitas [2, Theorem 2.1]; Gittins and Larson [11]) *The optimal stretch factor for minimizing  $\lambda_n(s)$  approaches 1 as  $n \rightarrow \infty$ , with*

$$S_*(n) \subset [1 - O(n^{-1/12}), 1 + O(n^{-1/12})],$$

and the minimal Dirichlet eigenvalue satisfies the asymptotic formula

$$\min_{s>0} \lambda_n(s) = \frac{4}{\pi}n + \left(\frac{4}{\pi}\right)^{3/2} n^{1/2} + O(n^{1/3}).$$

The proof is a modification of our Theorem 3. Full details are provided in the ArXiv version of this paper [18, Corollary 10]. In the proof one relies on Proposition 10 to bound the stretch factor  $s$  of the optimal rectangle. Proposition 10 is simpler in both statement and proof than the corresponding Theorem 3.1 of Antunes and Freitas [2], which contains an additional lower order term with an unhelpful sign.

*Remark* One would like to prove using only the definition of the counting function that

$$S_*(n) \longrightarrow 1 \quad \text{if and only if} \quad S(r) \longrightarrow 1,$$

or in other words that the rectangle minimizing the  $n$ th eigenvalue will converge to a square if and only if the ellipse maximizing the number of lattice points converges to a circle. Then Corollary 16 would follow qualitatively from Theorem 3, without needing any additional proof. Our attempts to find such an abstract equivalence have failed due to possible multiplicities in the eigenvalues. Perhaps an insightful reader will see how to succeed where we have failed.

## Maximizing eigenvalues of the Neumann Laplacian on rectangles

If one considers lattice points in the closed first quadrant, that is, allowing also the lattice points on the axes, then one obtains the Neumann eigenvalues of the rectangle having side lengths  $s^{-1}\pi$  and  $s\pi$ :

$$\{\mu_n(s) : n = 1, 2, 3, \dots\} = \{(js)^2 + (ks^{-1})^2 : j, k = 0, 1, 2, \dots\}.$$

Notice the first eigenvalue is always zero:  $\mu_1(s)=0$  for all  $s$ . The lattice point counting function is once again an eigenvalue counting function, because

$$\mathcal{N}(r, s) = \#\{(j, k) : (js)^2 + (ks^{-1})^2 \leq r^2\} = \#\{n : \mu_n(s) \leq r^2\}.$$

The appropriate problem is to maximize the  $n$ th eigenvalue (rather than minimizing as in the Dirichlet case), and so we let

$$\mathcal{S}_*(n) = \operatorname{argmax}_{s>0} \mu_n(s).$$

The corollary below says that the rectangle maximizing the  $n$ th Neumann eigenvalue will converge to a square as  $n \rightarrow \infty$ .

**Corollary 17.** (Optimal Neumann rectangle is asymptotically balanced, due to van den Berg, Bucur and Gittins [5]; Gittins and Larson [11]) *The optimal stretch factor for maximizing  $\mu_n(s)$  approaches 1 as  $n \rightarrow \infty$ , with*

$$\mathcal{S}_*(n) \subset [1 - O(n^{-1/12}), 1 + O(n^{-1/12})],$$

*and the maximal Neumann eigenvalue satisfies the asymptotic formula*

$$\max_{s>0} \mu_n(s) = \frac{4}{\pi} n - \left(\frac{4}{\pi}\right)^{3/2} n^{1/2} + O(n^{1/3}).$$

One adapts the arguments used for Theorem 7. A complete proof is in [18, Corollary 11]. Note that our lower bound on the counting function in Proposition 14, which one uses to control the stretch factor  $s$  of the optimal rectangle, is simpler in both statement and proof than the corresponding Lemma 2.2 by van den Berg *et al.* [5]. Further, our Proposition 14 holds for all  $r > 0$ , whereas [5, Lemma 2.2] holds only for  $r \geq 2s$ . Consequently we need not establish an *a priori* bound on  $s$  as was done in [5, Lemma 2.3].

Those authors did obtain a slightly better rate of convergence than we do, by calling on sophisticated lattice counting estimates of Huxley; see the comments after Proposition 11.

## Acknowledgments

This research was supported by grants from the Simons Foundation (# 204296 and #429422 to Richard Laugesen) and the University of Illinois Scholars' Travel Fund, and also by travel funding from the conference *Fifty Years of Hearing Drums: Spectral Geometry and the Legacy of Mark Kac* in Santiago, Chile, May 2016. At that conference, Michiel van den Berg generously discussed his papers [5] and [6] with Laugesen. The anonymous referee made useful suggestions that improved the organization of the paper. The material in this paper forms part of Shiya Liu's Ph.D. dissertation at the University of Illinois, Urbana–Champaign [20].



### A. The van der Corput sum

The next theorem is due to van der Corput [9, Satz 5], and is central to the proofs of Proposition 11 and Proposition 12. We formulate the theorem as in Krätzel [15, Korollar zu Satz 1.5, p. 24]. The constants in the theorem are interesting only for being modest in size.

Recall the sawtooth function  $\psi(x) = x - [x] - 1/2$ .

**Theorem 18.** *Suppose  $a < b$  and  $h \in C^2[a, b]$  with  $h''$  monotonic and nonzero. Then*

$$\left| \sum_{a < n \leq b} \psi(h(n)) \right| \leq 6 \int_a^b |h''(t)|^{1/3} dt + 175 \max_{[a, b]} \frac{1}{|h''|^{1/2}} + 1.$$

Krätzel's result has “+2” as the final term. We obtain the theorem with “+1” by correcting a gap in the proof and arguing carefully in one case, as explained in [20, Theorem A.8]. The constant term is irrelevant for our purposes in this paper, in any case.

### B. An elementary lemma

**Lemma 19.** *If  $s > 0$  and  $0 < t < 1$  then*

$$s + s^{-1} \leq 2 + t \implies |s - 1| \leq 3\sqrt{t}.$$

*Proof.* By taking the square root on both sides of the inequality

$$(s^{1/2} - s^{-1/2})^2 = s + s^{-1} - 2 \leq t,$$

and then using that the number 1 lies between  $s^{1/2}$  and  $s^{-1/2}$ , we find

$$|s^{1/2} - 1| \leq t^{1/2}.$$

Hence  $1 - t^{1/2} \leq s^{1/2} \leq 1 + t^{1/2}$ , and now squaring both sides and using that  $t < t^{1/2}$  (when  $t < 1$ ) proves the lemma.  $\square$

### References

1. ANTUNES, P. R. S. and FREITAS, P., Numerical optimisation of low eigenvalues of the Dirichlet and Neumann Laplacians, *J. Optim. Theory Appl.* **154** (2012), 235–257.
2. ANTUNES, P. R. S. and FREITAS, P., Optimal spectral rectangles and lattice ellipses, *Proc. R. Soc., Math. Phys. Eng. Sci.* **469** (2013), 20120492–20120492.
3. ANTUNES, P. R. S. and FREITAS, P., Optimisation of eigenvalues of the Dirichlet Laplacian with a surface area restriction, *Appl. Math. Optim.* **73** (2016), 313–328.

4. ARITURK, S. and LAUGESSEN, R. S., Optimal stretching for lattice points under convex curves, *Port. Math.* **74** (2017), 91–114.
5. van den BERG, M., BUCUR, D. and GITTINS, K., Maximizing Neumann eigenvalues on rectangles, *Bull. Lond. Math. Soc.* **48** (2016), 877–894.
6. van den BERG, M. and GITTINS, K., Minimising Dirichlet eigenvalues on cuboids of unit measure, *Mathematika* **63** (2017), 469–482.
7. BUCUR, D. and FREITAS, P., Asymptotic behaviour of optimal spectral planar domains with fixed perimeter, *J. Math. Phys.* **54**, 053504 (2013).
8. COLBOIS, B. and EL SOUFI, A., Extremal eigenvalues of the Laplacian on Euclidean domains and closed surfaces, *Math. Z.* **278** (2014), 529–549.
9. van der CORPUT, J. G., Zahlentheoretische Abschätzungen mit Anwendung auf Gitterpunktprobleme, *Math. Z.* **17** (1923), 250–259.
10. FREITAS, P., Asymptotic behaviour of extremal averages of Laplacian eigenvalues, *J. Stat. Phys.* **167** (2017), 1511–1518.
11. GITTINS, K. and LARSON, S., Asymptotic behaviour of cuboids optimising Laplacian eigenvalues, *Integral Equ. Oper. Theory* **89** (2017), 607–629.
12. HUXLEY, M. N., *Area, Lattice Points, and Exponential Sums*, London Mathematical Society Monographs. New Series **13**, The Clarendon Press, Oxford University Press, New York, 1996. Oxford Science Publications.
13. HUXLEY, M. N., Exponential sums and lattice points. III, *Proc. Lond. Math. Soc.* **87** (2003), 591–609.
14. IVIĆ, A., KRÄTZEL, E., KÜHLEITNER, M. and NOWAK, W. G., Lattice points in large regions and related arithmetic functions: recent developments in a very classic topic, in *Elementare und analytische Zahlentheorie*, Schr. Wiss. Ges. Johann Wolfgang Goethe Univ. Frankfurt am Main **20**, pp. 89–128, Franz Steiner Verlag, Stuttgart, 2006.
15. KRÄTZEL, E., *Analytische Funktionen in der Zahlentheorie*, Teubner–Texte zur Mathematik **139**, B. G. Teubner, Stuttgart, 2000.
16. KRÄTZEL, E., Lattice points in planar convex domains, *Monatshefte Math.* **143** (2004), 145–162.
17. LARSON, S., Asymptotic shape optimization for Riesz means of the Dirichlet Laplacian over convex domains. [1611.05680](#).
18. LAUGESSEN, R. S. and LIU, S., Optimal stretching for lattice points and eigenvalues. (Expanded version of the current paper.) [1609.06172](#).
19. LAUGESSEN, R. S. and LIU, S., Shifted lattices and asymptotically optimal ellipses, *J. Anal.* (2018) <https://doi.org/10.1007/s41478-017-0070-5>.
20. LIU, S., Asymptotically optimal shapes for counting lattice points and eigenvalues, Ph.D. dissertation, University of Illinois, Urbana–Champaign, 2017. <http://hdl.handle.net/2142/98364>.
21. MARSHALL, N. F., Stretching convex domains to capture many lattice points. *Int. Math. Res. Not. IMRN*, to appear. [1707.00682](#).
22. MARSHALL, N. F. and STEINERBERGER, S., Triangles capturing many lattice points. *Mathematika*, to appear. [1706.04170](#).
23. OUDET, É., Numerical minimization of eigenmodes of a membrane with respect to the domain, *ESAIM Control Optim. Calc. Var.* **10** (2004), 315–330.

Richard S. Laugesen  
Department of Mathematics  
University of Illinois  
Urbana, IL US-61801  
U.S.A.  
[Laugesen@illinois.edu](mailto:Laugesen@illinois.edu)

Shiya Liu  
Department of Mathematics  
University of Illinois  
Urbana, IL US-61801  
U.S.A.  
[sliu63@illinois.edu](mailto:sliu63@illinois.edu), [shiyaliu.zju@gmail.com](mailto:shiyaliu.zju@gmail.com)

*Received January 23, 2017  
in revised form May 8, 2017*