ASYMPTOTICALLY OPTIMAL SHAPES FOR COUNTING LATTICE POINTS AND EIGENVALUES

BY

SHIYA LIU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Associate Professor Lee DeVille, Chair
Professor Richard S. Laugesen, Director of Research
Associate Professor Eduard Kirr
Associate Professor Xiaochun Li

ABSTRACT

In Part I, we aim to maximize the number of first-quadrant lattice points under a concave (or convex) curve with respect to reciprocal stretching in the coordinate directions. The optimal domain is shown to be asymptotically balanced, meaning that the optimal stretch factor approaches 1 as the "radius" approaches infinity. In particular, the result implies when $1 < p < \infty$ that among all p-ellipses (or Lamé curves), the p-circle $x^p + y^p = r^p$ is asymptotically optimal for enclosing the most first-quadrant lattice points as the radius approaches infinity.

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: which right triangles in the first quadrant (with two sides along the axes) will enclose the most lattice points for given area, and what are the limiting shapes of those triangles as the area tends to infinity?

In Part II, we translate the positive-integer lattice points in the first quadrant by some amount in the horizontal and vertical directions. We seek to identify the limiting shape of the curve that encloses the greatest number of shifted lattice points in the same family of reciprocal stretching curves as in Part I.

The limiting shape is shown to depend explicitly on the lattice shift. The result holds for all positive shifts, and for negative shifts satisfying a certain condition. When the shift becomes too negative, the optimal curve no longer converges to a limiting shape, and instead it degenerates.

Our results handle the p-circle when $p > 1$ (concave) and also when $0 < p < 1$ (convex). The straight line case ($p = 1$) generates an open problem about minimizing high eigenvalues of quantum harmonic oscillators with normalized parabolic potentials.

To my parents and my advisor, for their love and support.

ACKNOWLEDGMENTS

Most results in this thesis are published in [20], [21] and [22].

Support from the Simons Foundation (#429422 to Richard Laugesen) and the University of Illinois Research Board award RB17002 is gratefully acknowledged.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION: LATTICE POINTS AND EIGENVALUES

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 1.1. Nonetheless, optimal ellipses must become more and more circular as the area increases to infinity, by a striking result of Antunes and Freitas [2].

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 πr^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 $s(r) \to 1$ as $r \to \infty$, as illustrated in Figure 1.2. 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. Chapter 9 explains the connection.) The analogous result for optimal ellipsoids becoming asymptotically spherical was proved recently in three dimensions by van den Berg and Gittins [7] and in higher dimensions by Gittins and Larson [12], once again in the eigenvalue formulation.

This thesis 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 3.5 says the maximizing $s(r)$

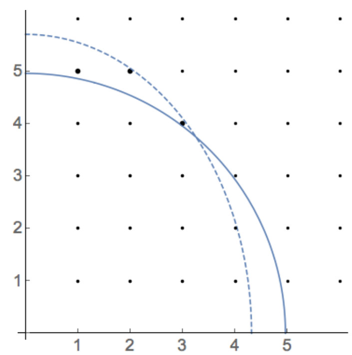

Figure 1.1: Circle $s = 1$ and ellipse $s = 1.15$, for radius $r = 4.96$. The ellipse encloses three more points than the circle, as shown in bold.

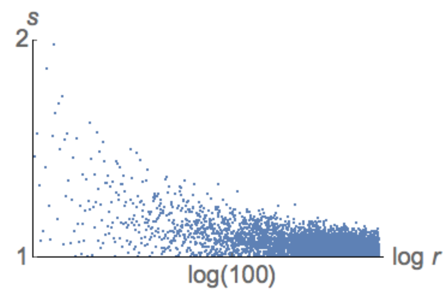

Figure 1.2: Optimal s-values for maximizing the number of lattice points in the 2-ellipse. The graph plots the largest value $s(r)$ versus $\log r$. The plotted r-values are multiples of $\sqrt{3}/10$, an irrational number chosen in the hope of exhibiting generic behavior. The horizontal axis is at height $s = 1$.

is bounded. Theorem 5.3 shows under a mild monotonicity hypothesis on the second derivative of the curve that $s(r) \to 1$ as $r \to \infty$. Thus the most "balanced" curve in the family will enclose the most lattice points in the limit. Marshall [23] recently extended this result to higher dimensions, using different methods.

Theorem 5.4 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 Example 5.5. In each case the p-circle is asymptotically optimal for the lattice counting problem in the first quadrant. The case $p = 1$ is an open problem. Our numerical evidence in Chapter 6 suggests that the first-quadrant right triangle enclosing the most lattice points does *not* approach a 45–45–90 degree triangle as $r \to \infty$. Instead one seems to get an infinite limit set of optimal triangles. Partial results in this direction were obtained recently by

Marshall and Steinerberger [24].

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 8.3 shows that the value $s = s(r)$ minimizing the number of enclosed lattice points will tend to 1 as $r \to \infty$. In the case of circles and ellipses, this result was obtained recently by van den Berg, Bucur and Gittins [6] (and in higher dimensions by Gittins and Larson [12], generalized by Marshall [23]). As explained in Chapter 9, they showed that the maximizing rectangle for the n-th eigenvalue of the Neumann Laplacian must approach a square as $n \to \infty$.

This work 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 3.4) in order to control the stretch factor $s(r)$. Then we prove more precise lattice counting estimates (Proposition 4.1) 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 [5] by building on this thesis's results. We include those results in Chapter 7, especially Theorem 7.6, Corollary 7.7 and Example 7.8.

Remark. The lattice point counting estimates in this thesis are similar to those used for the Gauss circle problem, which aims for accurate asymptotics on the counting function inside the circle and other closed curves as the area grows to infinity. Huxley [15] has the best known error estimate on the Circle Problem. The lattice counting formulas in our thesis differ somewhat from that work, because we consider only one quadrant of lattice points and later in Part II our regions contain empty strips due to the translation of the lattice. Further, we focus on proving suitable inequalities (rather than asymptotics) for the counting function, in order to prevent the maximizing shape from degenerating. In essence, we develop inequalities that trade off the empty regions in the vertical and horizontal directions. After degeneration has been ruled out we can invoke asymptotic formulas (with error terms that need not be as good as Huxley's) to prove convergence to a limiting shape.

Spectral motivations and results

This work is inspired by recent efforts to understand the behavior of high eigenvalues of the Laplacian. Write λ_n for the n-th eigenvalue of the Dirichlet Laplacian on a bounded domain Ω 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 λ_n^*, and suppose it is achieved on a domain Ω_n^*. What can one say about the shape of this minimizing domain?

For the first eigenvalue, the minimizing domain Ω_1^* is a disk, by the Faber–Krahn inequality. For the second eigenvalue, Ω_2^* is a union of two disjoint equal-area disks, as Krahn and later P. Szego showed. A long-standing conjecture says Ω_3^* should be a disk and Ω_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 Figure 1.3, and for more information consult [1, 13, 25] and references therein. Antunes and Freitas remark, though, that the "most natural guess" is Ω_n^* approaches a disk as $n \to \infty$, which is known to occur if the area normalization is strengthened to a perimeter normalization [3, 8]. This conjecture would imply the famous Pólya conjecture $\lambda_n \geq 4\pi n/|\Omega|$, as Colbois and El Soufi [9, Corollary 2.2] showed using subadditivity of $n \mapsto \lambda_n^*$.

A partial result of Freitas [11] succeeds in determining the leading order asymptotic as $n \to \infty$ of the minimum value of the eigenvalue sum $\lambda_1 + \cdots + \lambda_n$ (rather than of λ_n itself). This result provides no information on the shapes of the minimizing domains. Larson [19] 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 Ω_n^*.

A complete resolution for rectangular domains was found by Antunes and Freitas [2], using lattice counting methods as explained in Chapter 9. They proved that the minimizing domain for λ_n among rectangles approaches a square as $n \to \infty$. Similarly, the cube is asymptotically minimal in 3 and higher dimensions [7, 12].

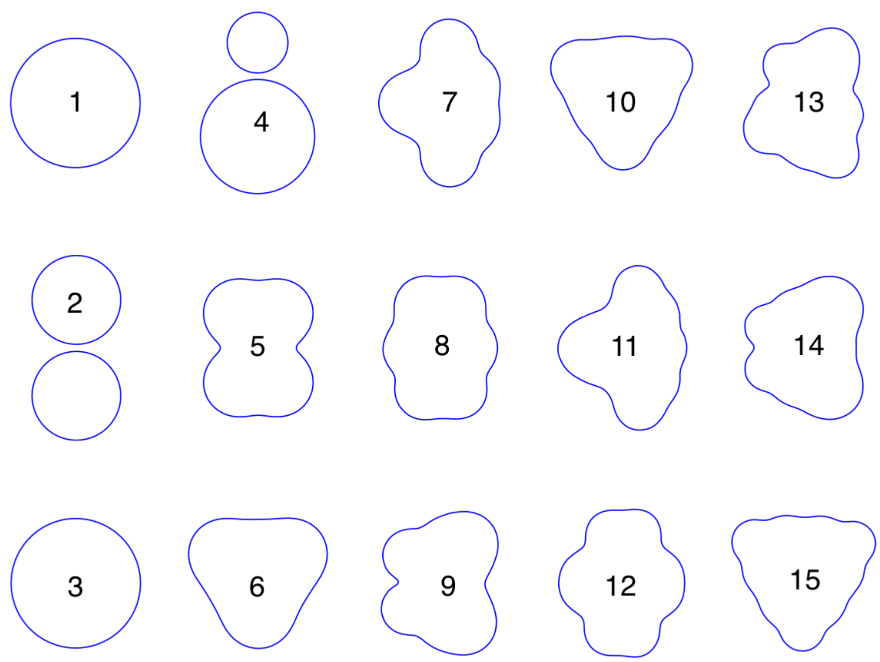

Figure 1.3: Minimizers of the first 15 Dirichlet eigenvalues. (Figure credit: Antunes and Oudet [13, Fig. 11.1], used with permission of the authors.)

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

The results on rectangular domains (which may be regarded as infinite potential wells) might suggest $s(n) \to 1$, but we think that it is not the case. Instead we believe $s(n)$ might cluster around infinitely many values as $n \to \infty$. 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 Chapter 6 for $p = 1$ except that now the lattices are shifted by 1/2 to the left and downwards. For the unshifted lattice, numerical work in Chapter 6 suggests that the optimal stretching parameter s does not converge to 1, and instead has many cluster points as

$r \to \infty$. Chapter 17 finds the same behavior for the shifted lattice too.

Part I

Integer lattice points

CHAPTER 2

ASSUMPTIONS AND DEFINITIONS

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

Assume throughout the thesis that Γ is a strictly decreasing curve in the first quadrant. Our first few theorems assume Γ is concave. The x- and y-intercepts of the curve are equal, occurring at $x = L$ and $y = L$ respectively, as shown in Figure 2.1. Write Area(Γ) for the area enclosed by the curve Γ and the x- and y-axes.

We represent the curve Γ 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 $r > 0$:

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

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

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

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

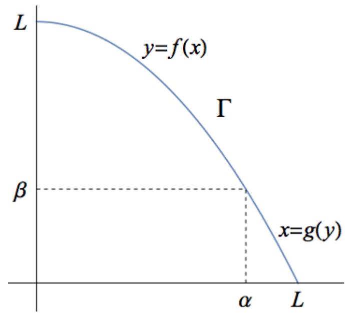

Figure 2.1: A concave decreasing curve Γ in the first quadrant, with intercepts at L. The point (α, β) on the curve is relevant to Theorem 5.3.

function for $r\Gamma(s)$ by

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

For each $r > 0$, consider the set

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

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 $s > 0$.

Analogously, when we count nonnegative-integer lattice points, which means we include lattice points on the axes, the counting function for $r\Gamma(s)$ is

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

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

Piecewise smooth curves

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 $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_l = \alpha$ exists such that $f \in C^2(0, \alpha_1]$ and $f \in C^2[\alpha_{i-1}, \alpha_i]$ for $i = 2, \ldots, l$. 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 $i = 2, \ldots, l$, with the derivative being taken in the one-sided senses at the partition points $\alpha_1, \ldots, \alpha_l$. 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 $0 = \beta_0 < \cdots < \beta_\ell = \beta$.

CHAPTER 3

BOUNDEDNESS OF THE OPTIMAL STRETCH PARAMETER

Our first task is to bound the set of optimal stretch parameters.

Lemma 3.1 (r-dependent bound on optimal stretch factors)**.** *If Γ is a concave, strictly decreasing curve in the first quadrant with equal intercepts (as in Figure 2.1), then*

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

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 after the definition of $S(r)$ in Chapter 2. Thus the maximum can only be attained when s lies in the interval $\big[(rL)^{-1}, rL\big]$. □

Next we will improve Lemma 3.1 to show the maximizing set $S(r)$ is bounded, and the bounds can be evaluated explicitly in the limit as $r \to \infty$. To control the stretch factors and hence prove Theorem 3.5, 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 Γ is 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 Γ, while $\mathcal{N}(r,s)$ counts nonnegative-integer lattice points.

Lemma 3.2 (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. □

Lemma 3.3 (Rough lower bound)**.** *The number $N(r,s)$ of positive-integer lattice points lying inside $r\Gamma(s)$ in the first quadrant satisfies*

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

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 $\mathcal{N}(r,s) \geq r^2 \operatorname{Area}(\Gamma)$, and so

$$\begin{aligned} N(r,s) &\geq \mathcal{N}(r,s) - r(s^{-1}L + sM) - 1 \qquad \text{by (3.2)} \\ &\geq r^2 \operatorname{Area}(\Gamma) - r(s^{-1}L + sM) - 1. \end{aligned}$$

□

For the upper bound in the next proposition, remember Γ 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 3.4 (Two-term upper bound on counting function)**.** *Let $C = M - f(L/2)$.*
(a) The number N of positive-integer lattice points lying inside Γ in the first quadrant satisfies

$$N \leq \operatorname{Area}(\Gamma) - \frac{1}{2}C \tag{3.1}$$

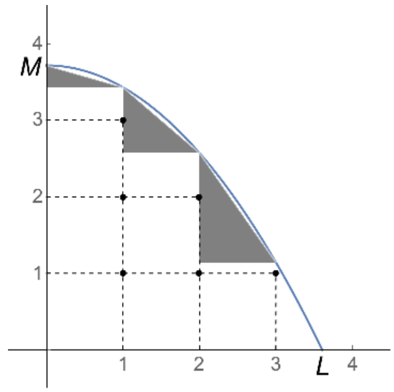

Figure 3.1: Positive integer lattice count satisfies
$N \leq \text{Area}(\Gamma) - \text{Area}(\text{triangles})$, in proof of Proposition 3.4(a).

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 Γ. Consider also the right triangles of width 1 formed by secant lines on Γ (see Figure 3.1), that is, the triangles with vertices $\big(i-1, f(i-1)\big)$, $\big(i, f(i)\big)$, $\big(i-1, f(i)\big)$, where $i = 1, \dots, \lfloor L \rfloor$. These triangles lie above the squares by construction, and lie below Γ by concavity. Hence

$$N + \text{Area}(\text{triangles}) \leq \text{Area}(\Gamma). \tag{3.2}$$

Since f is decreasing, we find

$$\begin{aligned}
\text{Area}(\text{triangles}) &= \sum_{i=1}^{\lfloor L \rfloor} \frac{1}{2}\big(f(i-1) - f(i)\big) \\
&= \frac{1}{2}\big(f(0) - f(\lfloor L \rfloor)\big) \qquad (3.3)\\
&\geq \frac{1}{2}\big(M - f(L/2)\big) = \frac{1}{2}C, \qquad (3.4)
\end{aligned}$$

because $\lfloor L \rfloor \geq L/2$ when $L \geq 1$. Combining (3.2) and (3.4) proves Part (a).

Part (b). Simply replace Γ 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. □

The curve Γ has x- and y-intercepts at L, in the theorems in that follow; see Figure 2.1. We start by improving Lemma 3.1 to show the maximizing set $S(r)$ is bounded, and the bounds can be evaluated explicitly in the limit as $r \to \infty$.

Theorem 3.5 (Uniform bound on optimal stretch factors)**.** *If Γ is a concave, strictly decreasing curve in the first quadrant then*

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

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

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

Proof. 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 3.1, and so the upper bound in Proposition 3.4(b) gives

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

The lower bound in Lemma 3.3 with "$s = 1$" says

$$N(r,1) \geq r^2 \operatorname{Area}(\Gamma) - 2rL - 1. \tag{3.5}$$

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 $s^{-1} \leq 5L/\widetilde{C} \equiv s_1^{-1}$, 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 $r \to \infty$. Going back to the proof of Proposition 3.4(a), we see from (3.2) and (3.3) that

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

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

$$N(r,s) + \frac{1}{2}rs\big(f(0) - f(\frac{s}{r}\lfloor \frac{rL}{s} \rfloor)\big) \leq r^2 \operatorname{Area}(\Gamma).$$

Hence

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

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 (3.5) and the above inequality imply

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

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

CHAPTER 4

TWO-TERM COUNTING ESTIMATES WITH EXPLICIT REMAINDER

Next we develop asymptotic formulas for the counting function. We start with a result for C^2-smooth curves. What matters in the following proposition is that the right side of estimate (4.1) 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 chapter.

Proposition 4.1 (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 Γ in the first quadrant satisfies:

$$\begin{aligned}
&\big|N - \operatorname{Area}(\Gamma) + (L+M)/2\big| \\
&\leq 6\Big(\int_0^\alpha |f''(x)|^{1/3}\,\mathrm{d}x + \int_0^\beta |g''(y)|^{1/3}\,\mathrm{d}y\Big) + 175\big(\max_{[0,\alpha]}\frac{1}{|f''|^{1/2}} + \max_{[0,\beta]}\frac{1}{|g''|^{1/2}}\big) \\
&\qquad + \frac{1}{4}\big(|f'(\alpha)| + |g'(\beta)|\big) + 3.
\end{aligned}$$

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

the first quadrant satisfies (for $r, s > 0$):

$$
\begin{aligned}
&\big|N(r,s) - r^2 \operatorname{Area}(\Gamma) + r(s^{-1}L + sM)/2\big| \\
&\le 6r^{2/3}\Big(\int_0^\alpha |f''(x)|^{1/3}\,\mathrm{d}x + \int_0^\beta |g''(y)|^{1/3}\,\mathrm{d}y\Big) + 175r^{1/2}\big(\max_{[0,\alpha]} \frac{s^{-3/2}}{|f''|^{1/2}} \\
&\qquad + \max_{[0,\beta]} \frac{s^{3/2}}{|g''|^{1/2}}\big) + \frac{1}{4}(s^2|f'(\alpha)| + s^{-2}|g'(\beta)|) + 3. \qquad (4.1)
\end{aligned}
$$

Proposition 4.1 and its proof are closely related to work of Krätzel [18, Theorem 1]. We give a direct proof below for two reasons: we want the estimate (4.1) 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 4.2.

A better bound on the right side of (4.1), giving order $O(r^{\theta+\epsilon})$ with $\theta = 131/208 \simeq 0.63 < 2/3$, can be found in work of Huxley [15], with precursors in [14, 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 $O(r^{(\theta+\epsilon-1)/2})$ in Theorem 5.3.

Proof. Part (a). We divide the region under Γ 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,

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

In terms of the *sawtooth* function ψ, defined by

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

one can evaluate

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

Then we apply the Euler–Maclaurin summation formula

$$\sum_{0<m\le\alpha} f(m) = \int_0^\alpha f(x)\,\mathrm{d}x - \psi(\alpha)f(\alpha) + \psi(0)f(0) + \int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x$$

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

$$\begin{aligned} N_1 &= \int_0^\alpha f(x)\,\mathrm{d}x - \psi(\alpha)f(\alpha) + \psi(0)f(0) + \int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x \\ &\qquad - \sum_{0<m\le\alpha} \psi\bigl(f(m)\bigr) - \lfloor\alpha\rfloor(1/2 + \lfloor\beta\rfloor) \\ &= \int_0^\alpha f(x)\,\mathrm{d}x - \psi(\alpha)\beta - M/2 + \int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x \\ &\qquad - \sum_{0<m\le\alpha} \psi\bigl(f(m)\bigr) - \lfloor\alpha\rfloor(1/2 + \lfloor\beta\rfloor). \end{aligned}$$

Similarly

$$\begin{aligned} N_2 &= \int_0^\beta g(y)\,\mathrm{d}y - \psi(\beta)\alpha - L/2 + \int_0^\beta g'(y)\psi(y)\,\mathrm{d}y \\ &\qquad - \sum_{0<n\le\beta} \psi\bigl(g(n)\bigr) - \lfloor\beta\rfloor(1/2 + \lfloor\alpha\rfloor), \end{aligned}$$

and so

$$\begin{aligned}
N &= N_1 + N_2 + N_3 \\
&= \int_0^\alpha f(x)\,\mathrm{d}x + \int_0^\beta g(y)\,\mathrm{d}y - \lfloor\alpha\rfloor\lfloor\beta\rfloor - (L+M)/2 \\
&\qquad - \psi(\alpha)\beta - \lfloor\alpha\rfloor/2 - \psi(\beta)\alpha - \lfloor\beta\rfloor/2 \\
&\qquad + \int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x + \int_0^\beta g'(y)\psi(y)\,\mathrm{d}y \\
&\qquad - \sum_{0<m\le\alpha} \psi\big(f(m)\big) - \sum_{0<n\le\beta} \psi\big(g(n)\big) \\
&= \mathrm{Area}(\Gamma) - (L+M)/2 + \int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x + \int_0^\beta g'(y)\psi(y)\,\mathrm{d}y \\
&\qquad - \sum_{0<m\le\alpha} \psi\big(f(m)\big) - \sum_{0<n\le\beta} \psi\big(g(n)\big) + \text{remainder}
\end{aligned} \tag{4.2}$$

where

$$\text{remainder} = -(\alpha - \lfloor\alpha\rfloor)(\beta - \lfloor\beta\rfloor) + (\alpha - \lfloor\alpha\rfloor + \beta - \lfloor\beta\rfloor)/2. \tag{4.3}$$

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

We estimate the sum of sawtooth functions in (4.2) by using Theorem A.8 (which is due to van der Corput): since f'' is monotonic and nonzero on $[0,\alpha]$, the theorem implies

$$\Big|\sum_{0<m\le\alpha} \psi\big(f(m)\big)\Big| \le 6\int_0^\alpha |f''(x)|^{1/3}\,\mathrm{d}x + 175 \max_{[0,\alpha]} \frac{1}{|f''|^{1/2}} + 1 \tag{4.4}$$

and similarly

$$\Big|\sum_{0<n\le\beta} \psi\big(g(n)\big)\Big| \le 6\int_0^\beta |g''(y)|^{1/3}\,\mathrm{d}y + 175 \max_{[0,\beta]} \frac{1}{|g''|^{1/2}} + 1. \tag{4.5}$$

To estimate the integrals of $f'\psi$ and $g'\psi$ in (4.2), we introduce the antiderivative of the sawtooth function, $\Psi(t) = \int_0^t \psi(z)\,\mathrm{d}z$, and observe that $-1/8 \le \Psi(t) \le 0$ for all $t \in \mathbb{R}$. By integration by parts and the fact that

$f'' < 0$, we have

$$
\begin{aligned}
\Big|\int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x\Big| &= \Big|\big[f'(x)\Psi(x)\big]_{x=0}^{x=\alpha} - \int_0^\alpha f''(x)\Psi(x)\,\mathrm{d}x\Big| \\
&\le \frac{1}{8}|f'(\alpha)| + \frac{1}{8}\Big|\int_0^\alpha f''(x)\,\mathrm{d}x\Big| \\
&= \frac{1}{8}|f'(\alpha)| + \frac{1}{8}\big(f'(0) - f'(\alpha)\big) \\
&\le \frac{1}{4}|f'(\alpha)| \qquad (4.6)
\end{aligned}
$$

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

$$
\Big|\int_0^\beta g'(y)\psi(y)\,\mathrm{d}y\Big| \le \frac{1}{4}|g'(\beta)|. \qquad (4.7)
$$

Combining (4.2)–(4.7) completes the proof of Part (a).

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

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 Γ, 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 δ and ϵ appear in the next Proposition. Their role in the proof is that on the intervals $0 < x \le \delta$ and $0 < y \le \epsilon$ we bound the sawtooth function trivially with $|\psi| \le 1/2$. On the remaining intervals we seek cancellations.

Proposition 4.2 (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 $i = 1, \ldots, l$. 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 $j = 1, \ldots, \ell$.*

(a) If $\delta \in (0, \alpha)$ and $\epsilon \in (0, \beta)$ then the number N of positive-integer lattice

points inside Γ *in the first quadrant satisfies:*

$$\begin{aligned}
&\big|N - \mathrm{Area}(\Gamma) + (L+M)/2\big| \\
&\le 6\Big(\int_0^\alpha |f''(x)|^{1/3}\,\mathrm{d}x + \int_0^\beta |g''(y)|^{1/3}\,\mathrm{d}y\Big) \\
&\quad + 175\Big(\frac{1}{|f''(\delta)|^{1/2}} + \frac{1}{|g''(\epsilon)|^{1/2}}\Big) + 350\Big(\sum_{i=1}^{l}\frac{1}{|f''(\alpha_i)|^{1/2}} + \sum_{j=1}^{\ell}\frac{1}{|g''(\beta_j)|^{1/2}}\Big) \\
&\quad + \frac{1}{4}\big(\sum_{i=1}^{l}|f'(\alpha_i)| + \sum_{j=1}^{\ell}|g'(\beta_j)|\big) + \frac{1}{2}\big(\delta+\epsilon\big) + l + \ell + 1.
\end{aligned}$$

(b) If functions

$$\delta : (0,\infty) \to (0,\alpha), \qquad \epsilon : (0,\infty) \to (0,\beta),$$

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

$$\begin{aligned}
&\big|N(r,s) - r^2\,\mathrm{Area}(\Gamma) + r(s^{-1}L + sM)/2\big| \\
&\le 6r^{2/3}\Big(\int_0^\alpha |f''(x)|^{1/3}\,\mathrm{d}x + \int_0^\beta |g''(y)|^{1/3}\,\mathrm{d}y\Big) \\
&\quad + 175r^{1/2}\Big(\frac{s^{-3/2}}{|f''\big(\delta(r)\big)|^{1/2}} + \frac{s^{3/2}}{|g''\big(\epsilon(r)\big)|^{1/2}}\Big) \\
&\quad + 350r^{1/2}\Big(\sum_{i=1}^{l}\frac{s^{-3/2}}{|f''(\alpha_i)|^{1/2}} + \sum_{j=1}^{\ell}\frac{s^{3/2}}{|g''(\beta_j)|^{1/2}}\Big) \\
&\quad + \frac{1}{4}\big(\sum_{i=1}^{l}s^2|f'(\alpha_i)| + \sum_{j=1}^{\ell}s^{-2}|g'(\beta_j)|\big) + \frac{r}{2}\big(s^{-1}\delta(r) + s\epsilon(r)\big) + l + \ell + 1.
\end{aligned} \tag{4.8}$$

The integral of $|f''|^{1/3}$ appearing in the conclusion of Proposition 4.2 is finite, because by Hölder's inequality and the fact that $f'' < 0$ and f is decreasing, we have

$$\int_0^{\alpha_1} |f''(x)|^{1/3}\,\mathrm{d}x \le \alpha_1^{2/3}\Big|\int_0^{\alpha_1} f''(x)\,\mathrm{d}x\Big|^{1/3} = \alpha_1^{2/3}|f'(0^+) - f'(\alpha_1^-)|^{1/3} < \infty.$$

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

Proof. Part (a). The lattice point counting equation (4.2) holds just as in the proof of Proposition 4.1, and so the task is to estimate each of the terms on the right side of that equation.

Estimate (4.4) 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

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

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

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

by Theorem A.8 applied on the interval $[\delta, \alpha_j]$. Applying that theorem again on each interval $[\alpha_{i-1}, \alpha_i]$ with $i = j+1, \ldots, l$ gives that

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

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

$$\begin{aligned} &\Big| \sum_{0<m\leq\alpha} \psi\big(f(m)\big)\Big| \\ &\leq \frac{1}{2}\delta + 6 \int_{\delta}^{\alpha} |f''(x)|^{1/3} \, \mathrm{d}x + \frac{175}{|f''(\delta)|^{1/2}} + \sum_{i=j}^{l-1} \frac{350}{|f''(\alpha_i)|^{1/2}} \\ &\quad + \frac{175}{|f''(\alpha)|^{1/2}} + l - j + 1 \\ &\leq \frac{1}{2}\delta + 6 \int_{0}^{\alpha} |f''(x)|^{1/3} \, \mathrm{d}x + \frac{175}{|f''(\delta)|^{1/2}} + \sum_{i=1}^{l} \frac{350}{|f''(\alpha_i)|^{1/2}} + l. \end{aligned} \tag{4.9}$$

Next, we adapt estimate (4.6) on the integral of $f'\psi$ by simply applying the same argument on each interval $[\alpha_{i-1}, \alpha_i]$, hence finding

$$\Big| \int_0^\alpha f'(x)\psi(x)\,\mathrm{d}x \Big| \le \sum_{i=1}^{l} \Big[\frac{1}{8} |f'(\alpha_i)| + \frac{1}{8} \big(f'(\alpha_{i-1}) - f'(\alpha_i) \big) \Big] \le \frac{1}{4} \sum_{i=1}^{l} |f'(\alpha_i)|. \tag{4.10}$$

By combining (4.2), (4.3) with (4.9), (4.10) 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, $f(x)$, $g(y)$, α, β, δ, ϵ with $rs^{-1}L$, rsM, $rsf(sx/r)$, $rs^{-1}g(s^{-1}y/r)$, $rs^{-1}\alpha$, $rs\beta, rs^{-1}\delta(r)$, $rs\epsilon(r)$ respectively. □

CHAPTER 5

OPTIMAL CONCAVE CURVE IS ASYMPTOTICALLY BALANCED

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

Proposition 5.1. *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*

$$H(r, s) = Ar^2 - Lr(s + s^{-1})/2 + O(r^\theta), \tag{5.1}$$

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

$$S(r) \subset [s_1, s_2] \qquad \textit{for all large } r > 0, \tag{5.2}$$

for some constants $s_1, s_2 > 0$.

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

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

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 (5.1) has implied constant depending on the interval $[s_1, s_2]$.

Proof. Since $S(r) \subset [s_1, s_2]$ by hypothesis (5.2), the asymptotic estimate

(5.1) 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 $r \to \infty$. Since s is a maximizing value, we have $H(r,1) \leq H(r,s)$ and so

$$s + s^{-1} \leq 2 + O(r^{-(1-\theta)}). \tag{5.3}$$

Hence $s = 1 + O(r^{-(1-\theta)/2})$ by Lemma 5.2 below, 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 $r \to \infty$, by (5.1) and using also that $1 \leq (s+s^{-1})/2 \leq 1 + O(r^{-(1-\theta)})$ by (5.3). □

Lemma 5.2. *If $s > 0$ and $0 < t < 1$ then*

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

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. □

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 5.3 (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

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

and the maximal lattice count has asymptotic formula

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

Proof. The theorem follows directly from Proposition 5.1 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 4.1(b) with $L = M$ one has

$$N(r,s) = \operatorname{Area}(\Gamma) r^2 - Lr(s + s^{-1})/2 + O(r^{2/3}), \tag{5.4}$$

with $s \in [s_1, s_2]$ as $r \to \infty$. Thus hypothesis (5.1) holds for $N(r,s)$ with the choices $A = \operatorname{Area}(\Gamma), \theta = 2/3$, and L equalling the intercept value of Γ.

The boundedness hypothesis (5.2) holds by Theorem 3.5. ☐

Slight improvements to the decay rate $O(r^{-1/6})$ and the error term $O(r^{2/3})$ are possible, as explained after Proposition 4.1.

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 $\epsilon(r)$ such that as $r \to \infty$:

$$\delta(r) = O(r^{-2a_1}), \qquad f''\big(\delta(r)\big)^{-1} = O(r^{1-4a_2}), \tag{5.5}$$

$$\epsilon(r) = O(r^{-2b_1}), \qquad g''\big(\epsilon(r)\big)^{-1} = O(r^{1-4b_2}). \tag{5.6}$$

(The second condition in (5.5) says that $f''(x)$ cannot be too small as $x \to 0$.) Let

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

Now we extend Theorem 5.3 to a larger class of concave decreasing curves.

Theorem 5.4 (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 $\epsilon(r)$ satisfy conditions (5.5) and (5.6).

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

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

and the maximal lattice count has asymptotic formula

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

Proof. Again let $H(r,s)$ be the lattice counting function $N(r,s)$, take $A = \operatorname{Area}(\Gamma)$, let L be the intercept value of Γ, and note the boundedness hypothesis (5.2) holds by Theorem 3.5. To finish verifying the hypotheses of Proposition 5.1, we suppose $0 < s_1 < s_2 < \infty$ and show that (5.1) 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 5.4. Hypothesis (5.1) is the assertion that

$$N(r,s) = \operatorname{Area}(\Gamma)r^2 - Lr(s+s^{-1})/2 + O(r^{1-2e}), \tag{5.7}$$

with $s \in [s_1, s_2]$ as $r \to \infty$. To verify this asymptotic, we will estimate the remainder terms in Proposition 4.2(b) as follows. In that proposition take $L = M$, and note $\delta(r) < \alpha$ and $\epsilon(r) < \beta$ for all large r by assumptions (5.5) and (5.6). We will show the right side of estimate (4.8) in Proposition 4.2(b) is bounded by

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

for large enough r, where the implied constants in the $O(\cdot)$-terms depend only on the curve Γ 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 (5.1) will hold as desired.

Examining now the right side of (4.8), we see the first two terms are obviously $O(r^{2/3})$. For the next term, observe by assumption in (5.5) 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 (4.8) can be estimated by $(s^{-3/2} + s^{3/2})O(r^{1/2})$. Similarly, the terms in (4.8) 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 (5.5), and similarly for $\epsilon(r)$. And, of course, $l + \ell + 1$ is constant, which completes the verification of hypothesis (5.1). $\square$

Example 5.5 (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 ℓ^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 $r \to \infty$, with

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

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

Theorem 5.3 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 5.4.

To verify that the p-circle satisfies the hypotheses of Theorem 5.4, we let

$\alpha = \beta = 2^{-1/p}$ and choose

$$\delta(r) = r^{-1/p}, \qquad \epsilon(r) = r^{-1/p},$$

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

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

and hence $a_2 = 1/2p$ in (5.5). Thus f satisfies hypothesis (5.5).

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

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

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 5.4 when $1 < p < \infty$.

For $p = \infty$, the ∞-circle is a Euclidean square and the ∞-ellipse is a rectangle. Many different rectangles of given area can contain the same number of lattice points. For example, a 4×1 rectangle and 2×2 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 Chapter 6. The case $0 < p < 1$ has been handled by Ariturk and Laugesen [5] using results here, as we explain in Chapter 7.

Incidentally, an explicit estimate on the number of lattice points in the full p-ellipse in all four quadrants was obtained by Krätzel [18, Theorem 2] for $p \geq 2$. See the informative survey by Ivić *et al.* [16, §3.1].

CHAPTER 6

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 Γ 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 6.1 shows an example.

The open problem is to understand the limiting behavior of the maximizing s-values. Does $S(r)$ converge to $\{1\}$ as $r \to \infty$? We proved the answer is "Yes" for p-ellipses when $1 < p < \infty$ (Example 5.5), but for $p = 1$ we suggest the answer is "No". Numerical evidence in Figure 6.2 suggests that the set $S(r)$ does not converge to $\{1\}$ as $r \to \infty$. 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 chapter we remark that maximizing s-values are ≤ 3 in the limit as $r \to \infty$, and we describe the numerical scheme that generates Figure 6.2. Lastly, we explain why $s = \sqrt{2}$ is a good candidate for a cluster value as $r \to \infty$.

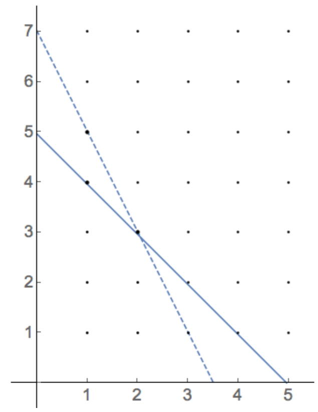

Figure 6.1: The 1-ellipse $sx + s^{-1}y = r$ with $r = 4.96$, for $s = 1$ (solid) and $s = \sqrt{2}$ (dashed). The dashed line encloses three more lattice points (shown in bold) than the solid line.

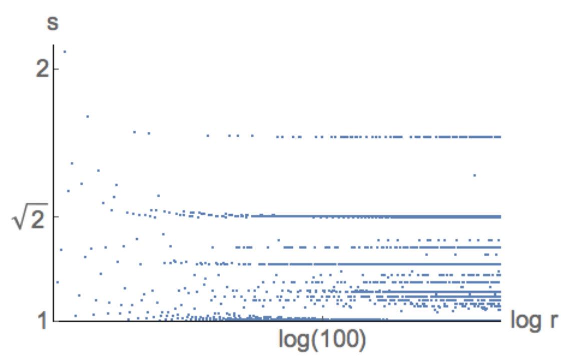

Figure 6.2: Optimal s-values for maximizing the number of lattice points in the 1-ellipse (triangle). The graph plots $\sup S(r)$ versus $\log r$. The plotted r-values are multiples of $\sqrt{3}/10$, and the horizontal axis is at height $s = 1$.

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

Given $\varepsilon > 0$, we claim

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

This bound is slightly better than the one in Theorem 3.5 (which had 4 instead of 3), and can be proved in the same way with the help of a special formula for $N(r, 1)$:

$$\begin{aligned} N(r,1) &= \#\text{ first-quadrant lattice points under the line } y = r - x \\ &= \frac{1}{2}\lfloor r \rfloor \lfloor r-1 \rfloor \\ &\geq \frac{1}{2}(r-1)(r-2) = \frac{1}{2}r^2 - \frac{3}{2}r + 1. \end{aligned} \tag{6.1}$$

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

$$js^2 - rs + k \leq 0. \tag{6.2}$$

For this quadratic inequality to have a solution, the discriminant must be nonnegative, $r^2 - 4jk \geq 0$, and thus we need only consider lattice points beneath the hyperbola $r^2 = 4xy$. For each such lattice point, equality holds in (6.2) 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 ∞, 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

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

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

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 $r > 0$. See Appendix B for the code.

When presenting the results of this method graphically, in Figure 6.2, 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 5.4 does not apply because f is linear and so $f'' \equiv 0$. Specifically, in the proof we see inequalities (4.4) and (4.5) are no longer useful, since their right sides are infinite. The situation cannot easily be rescued, because the left side of (4.1) 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

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

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 6.2, we will show that $s = \sqrt{2}$ gives a substantially higher count of lattice points than $s = 1$ (so that $1 \notin S(r)$), 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 show $S(r) \not\to \{1\}$ as $r \to \infty$.

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

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

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 $\sqrt{2}$ modulo 1. Then, writing $r = n - \epsilon$ where $\epsilon < 1/4$, we have

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

Hence $\limsup_{r\to\infty} \big(N(r, \sqrt{2}) - N(r, 1)\big)/r \geq 1/2$, 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 [24, Theorem]), while if $x > 0$ is irrational then $\sqrt{x} \notin S(r)$ for all sufficiently large r (see [24, Lemma 2] and its associated discussion).

The results of Marshall and Steinerberger are more detailed than we have indicated here. Even though they do not determine the set $S(r)$, they do identify an interesting infinite set of rational slopes for which the counting function is large on a sequence of r-values tending to infinity.

Conjecture for $p = 1$

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

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

Earlier in the chapter we proved that $S(\infty) \subset [1/3, 3]$. Also, Marshall and Steinerberger [24, Proposition 1] proved $S(n) = \{1\}$ for each $n \in \mathbb{N}, n \geq 2$, and hence $1 \in S(\infty)$.

The clustering behavior of $S(r)$ observed in Figure 6.2 suggests the following conjecture.

Conjecture 6.1 ($p = 1$)**.** *The limiting set $S(\infty)$ is countably infinite, and is contained in*

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

In order to prove this claim, one would presumably need to characterize $S(\infty)$ in terms of some number theoretic condition. The work of Marshall and Steinerberger [24] might be helpful for this purpose.

CHAPTER 7

OPTIMAL CONVEX CURVE IS ASYMPTOTICALLY BALANCED

Until now the thesis has considered concave decreasing curves. This chapter treats convex decreasing curves, and presents results of Ariturk and Laugesen [5]. Those authors applied a method of complementary domains in order to reduce from convex to concave curves. This chapter gives direct proofs instead. Also, we weaken the C^2-smoothness hypothesis from [5] to piecewise-C^2.

Optimal stretch is bounded

We develop some r-dependent bounds on the optimal stretch factors. Later, in the proof of Theorem 7.4, we will show the stretch factors are in fact uniformly bounded.

Lemma 7.1 (r-dependent bound on optimal stretch factors; Ariturk and Laugesen [5, Lemma 7.1])**.** *If*

$$r^2 \geq \frac{1}{\max\limits_{\Gamma} xy}$$

then

$$S(r) \subset \big[(rM)^{-1}, rL\big].$$

In this lemma the horizontal intercept L and vertical intercept M of the curve are allowed to differ in value.

For the proof see Ariturk and Laugesen [5, Lemma 7.1].

Lemma 7.2 (Improved r-dependent bound on optimal stretch factors; Ariturk and Laugesen [5, Lemma 7.2])**.** *A constant C exists, depending only on the curve Γ, such that if $r \geq C$ then*

$$S(r) \subset \big[2(rM)^{-1}, \frac{1}{2}rL\big].$$

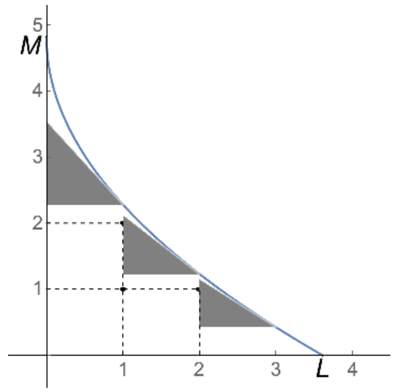

Figure 7.1: Figure taken with permission from Ariturk and Laugesen [5]. Positive integer lattice count $N(1,1) \leq \text{Area}(\Gamma) - \text{Area}(\text{triangles})$, in proof of Proposition 7.3.

For the proof refer to Ariturk and Laugesen [5, Lemma 7.2].

The curve Γ in the next proposition is the graph of $y = f(x)$, where f is convex and strictly decreasing on $[0, L]$, with $f(L) = 0, f(0) = M$. We do not assume the horizontal intercept L and vertical intercept M are equal. We also do not need differentiability of f in the next result.

Proposition 7.3 (Two-term upper bound on counting function; Ariturk and Laugesen [5, Proposition 5.1])**.** *The number $N(r,s)$ of positive-integer lattice points lying inside $r\Gamma(s)$ in the first quadrant satisfies*

$$N(r,s) \leq r^2 \,\text{Area}(\Gamma) - \frac{1}{2} f\Big(\frac{L}{2}\Big) rs$$

whenever $r \geq 2s/L$.

Proof. The following proof is taken verbatim from [5, Proposition 5.1], with permission of Ariturk and Laugesen. It is enough to prove the case $r = s = 1$ for $L \geq 2$, because then the general case of the proposition follows by applying the special case to the curve $r\Gamma(s)$ (which has horizontal intercept $rs^{-1}L$ and defining function $y = rsf(sx/r)$).

Clearly $N(1,1)$ equals the total area of the squares of sidelength 1 having upper right vertices at positive integer lattice points inside the curve Γ. The union of these squares is contained in Γ, since the curve is decreasing.

Consider the right triangles of width 1 formed by left-tangent lines on Γ, as shown in Figure 7.1. The triangles have vertices $\big(i-1, f(i)\big)$, $\big(i, f(i)\big)$, $\big(i-$

$1, f(i) - f'(i^-)\big)$, for $i = 1, \ldots, \lfloor L \rfloor$. Clearly the triangles lie under the curve by concavity, and lie outside the union of squares.

Hence

$$N(1,1) \leq \text{Area}(\Gamma) - \text{Area}(\text{triangles}).$$

To complete the proof, we estimate as follows:

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

since $L/2 \geq 1$ and $f(L) = 0$. □

The next result bounds the set of optimal stretch factors. The theorem appeared implicitly in the proofs in [5].

Theorem 7.4 (Uniform bound on optimal stretch factors)**.** *If Γ is a convex, strictly decreasing curve in the first quadrant then a constant C exists, depending only on the curve Γ, such that if $r \geq C$ then*

$$S(r) \subset [s_1, s_2],$$

for some constants $s_1, s_2 > 0$.

Proof. Recall the intercepts are assumed equal ($L = M$) in this theorem. Let $r \geq C$ where C is the constant as in Lemma 7.2 and suppose $s \in S(r)$. Then $r \geq 2s/L$ by Lemma 7.2, and so the upper bound in Proposition 7.3 gives

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

The lower bound in Lemma 3.3 with "$s = 1$" says

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

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

$$\frac{1}{2} f\big(\frac{L}{2}\big) rs \le 2rL + 1 \le \frac{5}{2} rL.$$

Hence $s \le 5L/f(L/2) \equiv s_2$, and so the set $S(r)$ is bounded above.

Interchanging the roles of the horizontal and vertical axes, we similarly find $s^{-1} \le 5L/g(L/2) \equiv s_1^{-1}$, so that the set $S(r)$ is bounded below away from 0, completing the first part of the proof. □

Two-term counting estimate

Here we prove the two-term counting estimate for general convex curves. We would like to handle infinite curvature at the intercepts of the curve Γ, meaning f'' must be allowed to blow up at $x = 0$. Further, we would like to have piecewise monotonicity assumption on f''. The next result achieves these goals.

Ariturk and Laugesen [5] give the proof using complementary regions. Here we extend their work to PC^2 curve and get a slightly improved bound using a direct proof. Two functions $\delta(r)$ and $\epsilon(r)$ appear in the next Proposition. Their role in the proof is that on the intervals $L - \delta(r) \le x < L$ and $M - \epsilon(r) \le y < M$ we bound the sawtooth function trivially with $|\psi| \le 1/2$. On the remaining intervals we seek cancellations.

Proposition 7.5 (Two-term counting estimate; Ariturk and Laugesen [5, Proposition 6.1])**.** *Take a point $(\alpha, \beta) \in \Gamma$ lying in the first quadrant with $\alpha, \beta < L$, and assume $f \in PC^2[\alpha, L)$ with $f' < 0$ and $f'' > 0$ on $[\alpha, L)$. Further suppose there is a partition $\alpha = \alpha_0 < \alpha_1 < \cdots < \alpha_l = L$ such that f'' is monotonic on each subinterval (α_{i-1}, α_i). Similarly assume $g \in PC^2[\beta, M)$ with $g' < 0$ and $g'' > 0$ on $[\beta, M)$, and there is a partition $\beta = \beta_0 < \beta_1 < \cdots < \beta_\ell = M$ such that g'' is monotonic on each subinterval (β_{j-1}, β_j).*

If functions

$$\delta : (0,\infty) \to (0, L - \alpha), \qquad \epsilon : (0,\infty) \to (0, M - \beta),$$

are given, then the number $N(r,s)$ of positive-integer lattice points inside

$r\Gamma(s)$ *in the first quadrant satisfies (for* $r, s > 0$*):*

$$\begin{aligned}
&\big|N(r,s) - r^2\,\mathrm{Area}(\Gamma) + r(s^{-1}L + sM)/2\big| \\
&\le 6r^{2/3}\Big(\int_\alpha^L f''(x)^{1/3}\,\mathrm{d}x + \int_\beta^M g''(y)^{1/3}\,\mathrm{d}y\Big) + 175r^{1/2}\Big(\frac{s^{-3/2}}{f''\big(L-\delta(r)\big)^{1/2}} \\
&\quad + \frac{s^{3/2}}{g''\big(M-\epsilon(r)\big)^{1/2}}\Big) + 350r^{1/2}\Big(\sum_{i=0}^{l-1}\frac{s^{-3/2}}{f''(\alpha_i)^{1/2}} + \sum_{j=0}^{\ell-1}\frac{s^{3/2}}{g''(\beta_j)^{1/2}}\Big) \\
&\quad + \frac{1}{4}\Big(\sum_{i=0}^{l-1} s^2|f'(\alpha_i)| + \sum_{j=0}^{\ell-1} s^{-2}|g'(\beta_j)|\Big) + \frac{r}{2}\big(s^{-1}\delta(r) + s\epsilon(r)\big) + l + \ell + 1.
\end{aligned} \tag{7.1}$$

As Ariturk and Laugesen observed, the integral of $(f'')^{1/3}$ appearing in the conclusion of Proposition 7.5 is finite, because by Hölder's inequality and the fact that $f'' > 0$ and f is decreasing, it is bounded by a constant times

$$\Big(\int_\alpha^L f''(x)\,\mathrm{d}x\Big)^{1/3} = |f'(L^-) - f'(\alpha)|^{1/3} \le \big(-f'(\alpha)\big)^{1/3} < \infty.$$

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

Proof. First we claim that if $0 < \delta < L-\alpha$, $0 < \epsilon < M-\beta$, then the number N of positive-integer lattice points inside Γ in the first quadrant satisfies

$$\begin{aligned}
&\big|N - \mathrm{Area}(\Gamma) + (L+M)/2\big| \\
&\le 6\Big(\int_\alpha^L f''(x)^{1/3}\,\mathrm{d}x + \int_\beta^M g''(y)^{1/3}\,\mathrm{d}y\Big) + 175\Big(\frac{1}{f''\big(L-\delta\big)^{1/2}} \\
&\quad + \frac{1}{g''\big(M-\epsilon\big)^{1/2}}\Big) + 350\Big(\sum_{i=0}^{l-1}\frac{1}{f''(\alpha_i)^{1/2}} + \sum_{j=0}^{\ell-1}\frac{1}{g''(\beta_j)^{1/2}}\Big) \\
&\quad + \frac{1}{4}\Big(\sum_{i=0}^{l-1}|f'(\alpha_i)| + \sum_{j=0}^{\ell-1}|g'(\beta_j)|\Big) + \frac{1}{2}\big(\delta+\epsilon\big) + l + \ell + 1.
\end{aligned} \tag{7.2}$$

As in Proposition 4.1, we will prove the claim by dividing the region under Γ into three parts N_1, N_2, N_3, but exchanging the role of horizontal and vertical axes in N_1 and N_2, so that

$$
\begin{aligned}
N_1 &= \sum_{\beta<n\le M}\ \sum_{0<m\le g(n)} 1 = \sum_{\beta<n\le M} \lfloor g(n)\rfloor = \sum_{\beta<n\le M} g(n) - \psi\big(g(n)\big) - \frac{1}{2},\\
N_2 &= \sum_{\alpha<m\le L}\ \sum_{0<n\le f(m)} 1 = \sum_{\alpha<m\le L} \lfloor f(m)\rfloor = \sum_{\alpha<m\le L} f(m) - \psi\big(f(m)\big) - \frac{1}{2},\\
N_3 &= \lfloor\alpha\rfloor\lfloor\beta\rfloor,
\end{aligned}
$$

where $\psi(x) = x - \lfloor x\rfloor - 1/2$ is the sawtooth function. We apply the Euler–Maclaurin summation formula

$$
\sum_{\beta<n\le M} g(n) = \int_\beta^M g(y)\,\mathrm{d}y - \psi(M)g(M) + \psi(\beta)g(\beta) + \int_\beta^M g'(y)\psi(y)\,\mathrm{d}y
$$

Then

$$
\begin{aligned}
N_1 &= \int_\beta^M g(y)\,\mathrm{d}y + \psi(\beta)\alpha + \int_\beta^M g'(y)\psi(y)\,\mathrm{d}y\\
&\qquad - \frac{1}{2}\big(\lfloor M\rfloor - \lfloor\beta\rfloor\big) - \sum_{\beta<n\le M}\psi\big(g(n)\big),\\
N_2 &= \int_\alpha^L f(x)\,\mathrm{d}x - \psi(\alpha)\beta + \int_\alpha^L f'(x)\psi(x)\,\mathrm{d}x\\
&\qquad - \frac{1}{2}\big(\lfloor L\rfloor - \lfloor\alpha\rfloor\big) - \sum_{\alpha<m\le L}\psi\big(f(m)\big),
\end{aligned}
$$

and so

$$
\begin{aligned}
N &= N_1 + N_2 + N_3\\
&= \mathrm{Area}(\Gamma) - (L+M)/2 + \int_\alpha^L f'(x)\psi(x)\,\mathrm{d}x + \int_\beta^M g'(y)\psi(y)\,\mathrm{d}y \qquad (7.3)\\
&\qquad - \sum_{\alpha<m\le L}\psi\big(f(m)\big) - \sum_{\beta<n\le M}\psi\big(g(n)\big) + \text{remainder} \qquad (7.4)
\end{aligned}
$$

where

$$
\begin{aligned}
\text{remainder} &= (\alpha - \lfloor\alpha\rfloor)(\beta - \lfloor\beta\rfloor) - (\alpha - \lfloor\alpha\rfloor + \beta - \lfloor\beta\rfloor)/2\\
&\qquad + (L - \lfloor L\rfloor + M - \lfloor M\rfloor)/2 \qquad (7.5)
\end{aligned}
$$

This remainder lies between -1 and 1, since $-1 \le xy - (x+y)/2 \le 0$ when

$x, y \in [0, 1]$.

To estimate the sum of sawtooth functions in (7.4), we first observe for the given positive constant δ that

$$\Big|\sum_{L-\delta<m\le L}\psi\big(f(m)\big)\Big| \le \frac{1}{2}\delta,$$

since $|\psi| \le 1/2$ everywhere. Next, we have $L-\delta \in [\alpha_j, \alpha_{j+1})$ for some $j \in \{0, \dots, l-1\}$, and

$$\begin{aligned}\Big|\sum_{\alpha_j<m\le L-\delta}\psi\big(f(m)\big)\Big| &\le 6\int_{\alpha_j}^{L-\delta} f''(x)^{1/3}\,\mathrm{d}x\\ &\qquad + 175\max\Big\{\frac{1}{f''(L-\delta)^{1/2}}, \frac{1}{f''(\alpha_j)^{1/2}}\Big\} + 1\end{aligned}$$

by Theorem A.8 applied on the interval $[\alpha_j, L-\delta]$. Applying that theorem again on each interval $[\alpha_{i-1}, \alpha_i]$ with $i = 1, \dots, j$ gives that

$$\begin{aligned}\Big|\sum_{\alpha_{i-1}<m\le \alpha_i}\psi\big(f(m)\big)\Big| &\le 6\int_{\alpha_{i-1}}^{\alpha_i} f''(x)^{1/3}\,\mathrm{d}x\\ &\qquad + 175\max\Big\{\frac{1}{f''(\alpha_{i-1})^{1/2}}, \frac{1}{f''(\alpha_i)^{1/2}}\Big\} + 1.\end{aligned}$$

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

$$\begin{aligned}&\Big|\sum_{\alpha<m\le L}\psi\big(f(m)\big)\Big|\\ &\le \frac{1}{2}\delta + 6\int_{\alpha}^{L-\delta} f''(x)^{1/3}\,\mathrm{d}x + \frac{175}{f''(L-\delta)^{1/2}} + \sum_{i=0}^{l-1}\frac{350}{f''(\alpha_i)^{1/2}} + l.\end{aligned} \tag{7.6}$$

To estimate the integrals of $f'\psi$ and $g'\psi$ in (7.4), we introduce the antiderivative of the sawtooth function, $\Psi(t) = \int_t^L \psi(z)\,\mathrm{d}z$, and an analogous proof as in Chapter 4 shows

$$\Big|\int_{\alpha}^{L} f'(x)\psi(x)\,\mathrm{d}x\Big| \le \frac{1}{4}\sum_{i=0}^{l-1}|f'(\alpha_i)|, \tag{7.7}$$

By combining (7.4), (7.5) with (7.6), (7.7) and the analogous estimates on

g, we complete the proof of the claim (7.2).

Next, to prove the proposition we rescale the curve Γ in (7.2) to $r\Gamma(s)$, by replacing L, M, $f(x)$, $g(y)$, α, β, δ, ϵ with $rs^{-1}L$, rsM, $rsf(sx/r)$, $rs^{-1}g(s^{-1}y/r)$, $rs^{-1}\alpha$, $rs\beta$, $rs^{-1}\delta(r)$, $rs\epsilon(r)$ respectively. □

Optimal curve is asymptotically balanced

Next we state Ariturk and Laugesen's result [5] on optimal stretches of convex curves. We handle piecewise smooth curves. Apart from that change, the proof is essentially the same as theirs.

Theorem 7.6 (Optimal convex curve is asymptotically balanced; Ariturk and Laugesen [5, Theorem 4.1])**.** *Assume $(\alpha, \beta) \in \Gamma$ is a point in the first quadrant with $\alpha, \beta < L$, such that $f \in PC^2[\alpha, L)$ with $f' < 0$ and $f'' > 0$ on $[\alpha, L)$, and similarly $g \in PC^2[\beta, L)$ with $g' < 0$ and $g'' > 0$ on $[\beta, L)$. Further suppose the interval (α, L) can be partitioned into finitely many subintervals on which f'' is monotonic, and similarly that (β, L) can be partitioned into subintervals on which g'' is monotonic. Moreover, assume constants $a_1, a_2, b_1, b_2 > 0$ and positive valued functions $\delta(r)$ and $\epsilon(r)$ exist such that as $r \to \infty$,*

$$\delta(r) = O(r^{-2a_1}), \qquad \frac{1}{f''\big(L - \delta(r)\big)} = O(r^{1-4a_2}),$$
$$\epsilon(r) = O(r^{-2b_1}), \qquad \frac{1}{g''\big(L - \epsilon(r)\big)} = O(r^{1-4b_2}).$$

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

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

where the exponent is $e = \min(\frac{1}{6}, a_1, a_2, b_1, b_2)$. Further, the maximal lattice count has asymptotic formula

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

Proof. Let $H(r, s)$ be the lattice counting function $N(r, s)$, take $A = \operatorname{Area}(\Gamma)$, let L be the intercept value of Γ, and note the boundedness hypothesis (5.2) holds by Theorem 7.4. To finish verifying the hypotheses of Proposition 5.1, we suppose $0 < s_1 < s_2 < \infty$ and show that (5.1) 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 7.6. Hypothesis (5.1) is the assertion that (5.7) holds with $s \in [s_1, s_2]$ as $r \to \infty$. To verify this asymptotic, we will estimate the remainder terms in Proposition 7.5 as follows. In that proposition take $L = M$, and note $\delta(r) < \alpha$ and $\epsilon(r) < \beta$ for all large r by assumptions in Theorem 7.6. The right side of estimate (7.1) in Proposition 7.5 is bounded by

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

for all $r, s > 0$, where the implied constants in the $O(\cdot)$-terms depend only on the curve Γ 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 (5.1) will hold as desired. ☐

Corollary 7.7 (Ariturk and Laugesen [5, Corollary 4.2])**.** *Assume* $(\alpha, \beta) \in \Gamma$ *is a point in the first quadrant, such that* $f \in C^2[\alpha, L]$ *with* $f' < 0, f'' > 0$ *and* f'' *monotonic, and* $g \in C^2[\beta, L]$ *with* $g' < 0, g'' > 0$ *and* g'' *monotonic.*

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

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

and the maximal lattice count satisfies

$$\max_{s>0} N(r, s) = r^2 \operatorname{Area}(\Gamma) - rL + O(r^{2/3}). \tag{7.8}$$

The corollary follows by taking $a_1 = b_1 = 1/2, a_2 = b_2 = 1/4, e = 1/6$ in the theorem and noting that $f''(L) > 0$ and $g''(L) > 0$ by assumption.

Example 7.8 (Optimal p-ellipses for lattice point counting; Ariturk and Laugesen [5, Example 4.3])**.** The p-ellipse with $0 < p < 1$ containing the maximum number of positive-integer lattice points must approach a p-circle in the limit as $r \to \infty$, with

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

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

The following proof is taken directly from Ariturk and Laugesen's paper, and is included here just for the reader's convenience. To verify that the p-circle satisfies the hypotheses of Theorem 7.6, we let $\alpha = \beta = 2^{-1/p}$. The first and second derivatives of $f(x) = (1 - x^p)^{1/p}$ are given in Example 5.5, and the third derivative is

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

If $0 < p \le 1/2$ then $f''' < 0$ on the interval $(0, 1)$, and so f'' is monotonic. If $1/2 < p < 1$ then f''' vanishes at exactly one point in the interval $(\alpha, 1)$, namely at $\alpha_1 = [(2-p)/(1+p)]^{1/p}$, and so f'' is monotonic on the subintervals (α, α_1) and $(\alpha_1, 1)$. Further, we choose $a_1 = a_2 = p/2$ and let $\delta(r) = r^{-2a_1} = r^{-p}$ for all large r, and verify directly that

$$\frac{1}{f''\big(1 - \delta(r)\big)} = O(r^{1-2p}) = O(r^{1-4a_2}).$$

The calculations are the same for g, and so the desired conclusion for p-ellipses with $0 < p < 1$ now follows from Theorem 7.6. The analogous result for $1 < p < \infty$ is in Example 5.5.

CHAPTER 8

CLOSED FIRST QUADRANT LATTICE POINTS

Our results have analogues for lattice point counting in the closed (rather than open) first quadrant for both concave and convex curves, as we now explain.

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

8.1 Concave curves: optimal stretch parameter

In this section we assume f is concave.

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 3.4, although the constant $\mathcal{C}$ is slightly different than in that result.

Proposition 8.1 (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 Γ in the closed first quadrant satisfies:

$$\mathcal{N} \geq \text{Area}(\Gamma) + \frac{1}{2}\mathcal{C}. \tag{8.1}$$

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

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

Proof. Part (a). Clearly $\mathcal{N}$ equals the total area of the squares of sidelength

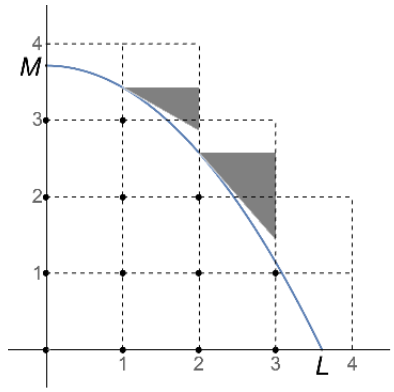

Figure 8.1: Nonnegative integer lattice count satisfies
$\mathcal{N} \geq \text{Area}(\Gamma) + \text{Area}(\text{triangles})$, in proof of Proposition 8.1(a) when $L \geq 2$.

1 having lower left vertices at nonnegative-integer lattice points inside the curve Γ. The union of these squares contains Γ, 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

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

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

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

as desired.

Case (ii): Suppose $L \geq 2$. Consider the right triangles of width 1 formed by tangent lines from the right on Γ, that is, the triangles with vertices $\big(i, f(i)\big), \big(i+1, f(i)\big), \big(i+1, f(i) + f'(i^+)\big)$, where $i = 0, 1, \ldots, \lfloor L \rfloor - 1$, see Figure 8.1. 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 Γ by concavity. Hence

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

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

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

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

Part (b). Replace Γ 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. ☐

Uniform bound on optimal stretch factors

Theorem 8.2 (Uniform bound on optimal stretch factors)**.** *If Γ is a concave, strictly decreasing curve in the first quadrant then*

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

for some constants $s_1, s_2 > 0$.

Proof. Since $N(r,s) \leq r^2 \,\text{Area}(\Gamma)$, taking $s = 1$ and $L = M$ in Lemma 3.2 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 8.1(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 $s^{-1} \leq 5L/\widetilde{\mathcal{C}}$. Hence $\mathcal{S}(r)$ is bounded below away from 0, which completes the proof. □

Optimal concave curve is asymptotically balanced

Theorem 8.3 (Optimal concave curve is asymptotically balanced)**.** *Under the assumptions of Theorem 5.3, the optimal stretch factor for minimizing $\mathcal{N}(r,s)$ approaches 1 as r tends to infinity:*

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

and under the assumptions of Theorem 5.4 we have similarly that:

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

Proof. The theorem will follow from Proposition 5.1 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 (5.2) of the proposition holds by Theorem 8.2. The other hypothesis (5.1) is verified as follows.

Taking $L = M$ in the relation between $\mathcal{N}(r,s)$ and $N(r,s)$ in Lemma 3.2, and calling on the asymptotic for $N(r,s)$ in either (5.4) (under the assumptions of Theorem 5.3) or (5.7) (under the assumptions of Theorem 5.4), we deduce

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

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

$$\theta = \begin{cases} 2/3 & \text{under the assumptions of Theorem 5.3,} \\ 1 - 2e & \text{under the assumptions of Theorem 5.4.} \end{cases}$$

That is, we have verified hypothesis (5.1) with $H = -\mathcal{N}, A = -\operatorname{Area}(\Gamma)$, and L the intercept value of Γ. □

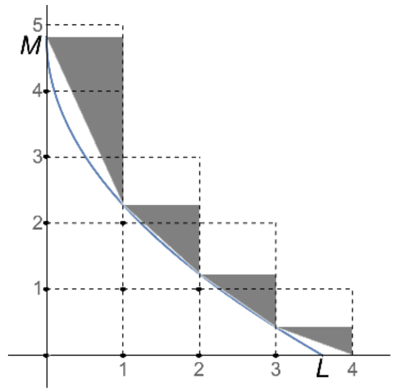

Figure 8.2: Nonnegative integer lattice count
$\mathcal{N}(1,1) \geq \text{Area}(\Gamma) + \text{Area}(\text{triangles})$, in proof of Proposition 8.4. Figure taken with permission from Ariturk and Laugesen [5, Figure 5].)

8.2 Convex curves: optimal stretch parameter

First we need a two-term bound on the counting function in the closed first quadrant. Assume in this section that f is convex. Then we have the following analogue of Proposition 7.3.

Proposition 8.4 (Two-term lower bound on counting function; Ariturk and Laugesen [5, Proposition 9.1])**.** *The number of nonnegative-integer lattice points lying inside $r\Gamma(s)$ in the closed first quadrant satisfies*

$$\mathcal{N}(r,s) \geq r^2 \,\text{Area}(\Gamma) + \frac{1}{2}Mrs, \qquad r,s > 0.$$

Proof. The proof below is taken directly from [5, Proposition 9.1]. We need only prove the special case where $r = s = 1$, because applying that case to the curve $r\Gamma(s)$ (which has vertical intercept Mrs) yields the general case of the proposition.

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

Consider the right triangles lying above chords of Γ, as shown in Figure 8.2. That is, for $i = 1, \ldots, \lfloor L \rfloor$ we take the triangle with vertices $(i-1, f(i-1)), (i, f(i)), (i, f(i-1))$, and the final triangle has vertices at $(\lfloor L \rfloor, f(\lfloor L \rfloor)), (\lceil L \rceil, 0), (\lceil L \rceil, f(\lfloor L \rfloor))$.

These triangles all lie above Γ, by concavity, and lie inside the collection of squares of sidelength 1. Hence

$$\mathcal{N}(1,1) \geq \text{Area}(\Gamma) + \text{Area}(\text{triangles}) = \text{Area}(\Gamma) + \frac{1}{2}M.$$

□

Uniform bound on optimal stretch factors

Theorem 8.5 (Uniform bound on optimal stretch factors)**.** *If Γ is a convex, strictly decreasing curve in the first quadrant then*

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

for some constants $s_1, s_2 > 0$.

Proof. Since $N(r,s) \leq r^2 \,\text{Area}(\Gamma)$, taking $s = 1$ and $L = M$ in Lemma 3.2 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 8.1(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 $s^{-1} \leq 5L/\widetilde{\mathcal{C}}$. Hence $\mathcal{S}(r)$ is bounded below away from 0, which completes the proof. □

Optimal convex curve is asymptotically balanced

Theorem 8.6 (Optimal convex curve is asymptotically balanced; Ariturk and Laugesen [5, Theorem 4.4])**.** *Under the assumptions of Theorem 7.6,* *the optimal stretch factor for minimizing $\mathcal{N}(r,s)$ approaches 1 as r tends to infinity, with*

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

and under the assumptions of Corollary 7.7, we similarly have that

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

Proof. The theorem will follow from Proposition 5.1 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 (5.2) of the proposition holds by Theorem 8.5. The other hypothesis (5.1) is verified as follows.

Taking $L = M$ in the relation between $\mathcal{N}(r,s)$ and $N(r,s)$ in Lemma 3.2, and calling on the asymptotic for $N(r,s)$ in (7.1) under either the assumptions of Theorem 7.6 or the assumptions of Corollary 7.7, we deduce

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

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

$$\theta = \begin{cases} 1 - 2e & \text{under the assumptions of Theorem 7.6,} \\ 2/3 & \text{under the assumptions of Corollary 7.7.} \end{cases}$$

That is, we have verified hypothesis (5.1) with $H = -\mathcal{N}, A = -\operatorname{Area}(\Gamma)$, and L the intercept value of Γ. $\square$

CHAPTER 9

CONNECTION WITH EIGENVALUE MINIMIZATION AND MAXIMIZATION

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 chapter, $L = M = 1$ and $\text{Area}(\Gamma) = \pi/4$.

Minimizing eigenvalues of the Dirichlet Laplacian on rectangles

Write

$$\{\lambda_n(s) : n = 1, 2, 3, \dots\} = \{(js)^2 + (ks^{-1})^2 : j, k = 1, 2, 3, \dots\} \tag{9.1}$$

so that $\lambda_n(s)$ is the nth 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

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

Define

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

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

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

Corollary 9.1 (Optimal Dirichlet rectangle is asymptotically balanced, due to Antunes and Freitas [2, Theorem 2.1]; Gittins and Larson [12]).
The optimal stretch factor for minimizing $\lambda_n(s)$ approaches 1 as $n \to \infty$, with

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

and the minimal Dirichlet eigenvalue satisfies the asymptotic formula

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

The proof is a modification of our Theorem 5.3. Full details are provided in the ArXiv version of this work [20, Corollary 10]. In the proof one relies on Proposition 3.4 to bound the stretch factor s of the optimal rectangle. Proposition 3.4 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

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

or in other words that the rectangle minimizing the nth eigenvalue will converge to a square if and only if the ellipse maximizing the number of lattice points converges to a circle. Then Corollary 9.1 would follow qualitatively from Theorem 5.3, 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 $s\pi$:

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

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

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

The appropriate problem is to maximize the nth 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 nth Neumann eigenvalue will converge to a square as $n \to \infty$.

Corollary 9.2 (Optimal Neumann rectangle is asymptotically balanced, due to van den Berg, Bucur and Gittins [6]; Gittins and Larson [12])**.**
The optimal stretch factor for maximizing $\mu_n(s)$ approaches 1 as $n \to \infty$, 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 8.3. A complete proof is in [20, Corollary 11]. Note that our lower bound on the counting function in Proposition 8.1, 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.* [6]. Further, our Proposition 8.1 holds for all $r > 0$, whereas [6, Lemma 2.2] holds only for $r \ge 2s$. Consequently we need not establish an *a priori* bound on s as was done in [6, 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 4.1.

Part II

Shifted integer lattice points

CHAPTER 10

OVERVIEW: SHIFTED LATTICE POINTS

This second part of the thesis tackles a variant of the lattice counting problem in which the lattice is translated by some increments in the x- and y-directions, and shows the asymptotically optimal ellipse is no longer a circle but an ellipse whose semi-axis ratio depends explicitly on the translation increments. This optimal ratio succeeds in "balancing" the horizontal and vertical empty strip areas created by the translation of the lattice; see Figure 10.1. The precise statement is given in Corollary 15.3.

Generalized ellipses obtained by stretching a smooth curve (either concave or convex) can be handled by our methods too, in Theorem 15.1. The results hold for all positive translations, and for small negative translations that satisfy a computable, curve-dependent criterion.

Open problems for right triangles and shifted lattice points are investigated in Chapter 17.

Chapter 18 explains the connection to eigenvalues of the quantum harmonic oscillator and for a whole family of such Schrödinger eigenvalue problems.

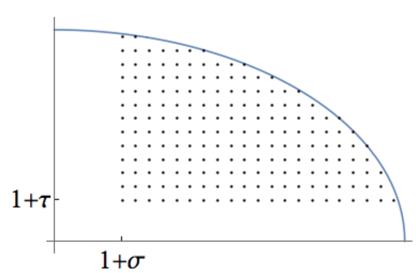

Figure 10.1: An ellipse that maximizes (among ellipses with the same area) the number of enclosed positive-integer lattice points shifted by 4 units horizontally and 2 units vertically. This optimal ellipse roughly balances the areas of the horizontal and vertical empty strips; see Corollary 15.3.

CHAPTER 11

NOTATIONS AND DEFINITIONS

Given numbers $\sigma, \tau > -1$, consider the translated or *shifted* positive-integer lattice

$$(\mathbb{N}+\sigma) \times (\mathbb{N}+\tau),$$

which lies in the open first quadrant. The first part of this thesis considered the unshifted lattice, where $\sigma = \tau = 0$. In this second part of the thesis, define the shifted-lattice counting function under the curve $s\Gamma(s)$ to be

$$\begin{aligned} N(r,s) &= \text{number of shifted positive-integer lattice points lying under } r\Gamma(s) \\ &= \#\big\{(j,k) \in \mathbb{N}\times\mathbb{N} : k+\tau \leq rsf\big((j+\sigma)s/r\big)\big\}. \end{aligned}$$

The set $S(r)$ consists of s-values that maximize $N(r,s)$, that is,

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

Next we will define parameters conditions for $\sigma, \tau > -1$.

Parameter Assumption 11.1. Γ is concave and strictly decreasing, with

$$\max\Big\{f\big(\frac{1-\sigma^-}{2-\sigma^-}L\big), g\big(\frac{1-\tau^-}{2-\tau^-}L\big)\Big\} < 2\Big(\frac{1}{2}-\sigma^- - \tau^-\Big)L. \tag{11.1}$$

Parameter Assumption 11.2. Γ is convex and strictly decreasing, with

$$\min\Big\{(1-\sigma^-)f\big(\frac{1-\sigma^-}{2-\sigma^-}L\big), (1-\tau^-)g\big(\frac{1-\tau^-}{2-\tau^-}L\big)\Big\} > 2(\sigma^- + \tau^-)L \tag{11.2}$$

and

$$\mu_f(\sigma) \stackrel{\text{def}}{=} \min\Big\{(1+\sigma)f\big(\frac{1+\sigma}{2+\sigma}x\big) - f(x) : \frac{1+\sigma}{2+\sigma}L \leq x \leq L\Big\} > 0, \tag{11.3}$$

$$\mu_g(\tau) \stackrel{\text{def}}{=} \min\Big\{(1+\tau)g\big(\frac{1+\tau}{2+\tau}y\big) - g(y) : \frac{1+\tau}{2+\tau}L \leq y \leq L\Big\} > 0. \tag{11.4}$$

These Parameter Assumptions are significant only when $\sigma < 0$ or $\tau < 0$, because when $\sigma, \tau \geq 0$ the conditions (11.1) and (11.2) hold automatically (using that $0 < f(x) < L$ and $0 < g(y) < L$ when $x, y \in (0, L)$) and also conditions (11.3) and (11.4) hold automatically (using that f and g are strictly decreasing and positive).

Next we state the smoothness conditions to be used.

Concave Condition 11.3. Γ is concave, and for some $(\alpha, \beta) \in \Gamma$ with $\alpha, \beta > 0$ one has $f \in C^2[0, \alpha], g \in C^2[0, \beta]$, with

$$
\begin{aligned}
&f' < 0 \text{ on } (0, \alpha],\ f'' < 0 \text{ on } [0, \alpha],\ f'' \text{ monotonic on } [0, \alpha], \\
&g' < 0 \text{ on } (0, \beta],\ g'' < 0 \text{ on } [0, \beta],\ g'' \text{ monotonic on } [0, \beta].
\end{aligned}
$$

Convex Condition 11.4. Γ is convex, and for some $(\alpha, \beta) \in \Gamma$ with $\alpha, \beta > 0$ one has $f \in C^2[\alpha, L], g \in C^2[\beta, L]$, with

$$
\begin{aligned}
&f' < 0 \text{ on } [\alpha, L),\ f'' > 0 \text{ on } [\alpha, L],\ f'' \text{ monotonic on } [\alpha, L], \\
&g' < 0 \text{ on } [\beta, L),\ g'' > 0 \text{ on } [\beta, L],\ g'' \text{ monotonic on } [\beta, L].
\end{aligned}
$$

Later we will need weaker smoothness conditions. Let (α, β) be a point on the curve Γ with $\alpha, \beta > 0$.

Weaker Concave Condition 11.5. Suppose Γ is concave, and:

- $f \in C^2(0, \alpha], f' < 0, f'' < 0$, and a partition $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_l = \alpha$ exists such that f'' is monotonic on (α_{i-1}, α_i) for each $i = 1, 2, \ldots, l$.
- $g \in C^2(0, \beta], g' < 0, g'' < 0$, and a partition $0 = \beta_0 < \beta_1 < \cdots < \beta_m = \beta$ exists such that g'' is monotonic on (β_{i-1}, β_i) for each $i = 1, 2, \ldots, m$.
- Positive functions $\delta(r)$ and $\epsilon(r)$ exist such that

$$
\delta(r) = O(r^{-2a_1}), \qquad f''\big(\delta(r)\big)^{-1} = O(r^{1-4a_2}), \tag{11.5}
$$

$$
\epsilon(r) = O(r^{-2b_1}), \qquad g''\big(\epsilon(r)\big)^{-1} = O(r^{1-4b_2}). \tag{11.6}
$$

  as $r \to \infty$, for some numbers $a_1, a_2, b_1, b_2 > 0$.

- Let $a_3 = b_3 = 1/2$.

(The second condition in (11.5) says that $f''(x)$ cannot be too small as $x \to 0$.)

Weaker Convex Condition 11.6. Suppose Γ is convex, and:

- $f \in C^2[\alpha, L), f' < 0, f'' > 0$, and a partition $\alpha = \alpha_0 < \alpha_1 < \cdots < \alpha_l = L$ exists such that f'' is monotonic on (α_{i-1}, α_i) for each $i = 1, 2, \ldots, l$.
- $g \in C^2[\beta, L), g' < 0, g'' > 0$, and a partition $\beta = \beta_0 < \beta_1 < \cdots < \beta_m = L$ exists such that g'' is monotonic on (β_{i-1}, β_i) for each $i = 1, 2, \ldots, m$.
- Positive functions $\delta(r)$ and $\epsilon(r)$ exist such that

$$\delta(r) = O(r^{-2a_1}), \qquad f''\big(L - \delta(r)\big)^{-1} = O(r^{1-4a_2}), \tag{11.7}$$
$$\epsilon(r) = O(r^{-2b_1}), \qquad g''\big(L - \epsilon(r)\big)^{-1} = O(r^{1-4b_2}). \tag{11.8}$$

  as $r \to \infty$, for some numbers $a_1, a_2, b_1, b_2 > 0$.
- $f(x) = L + O(x^{2a_3})$ as $x \to 0$, and $g(y) = L + O(y^{2b_3})$ as $y \to 0$, for some numbers $a_3, b_3 > 0$.

(The last condition says Γ cannot approach the axes too rapidly near the intercept points.)

CHAPTER 12

CONCAVE CURVES — COUNTING FUNCTION ESTIMATES

In this chapter, we estimate the counting function. The curve Γ is taken to be concave decreasing in the first quadrant, throughout this chapter. Denote the horizontal and vertical intercepts by $x = L$ and $y = M$ respectively, where L and M are positive but not necessarily equal. Allowing unequal intercepts is helpful for some of the results below.

We start with a preliminary r-dependent bound on the maximizing set $S(r)$. The proof of this bound also makes clear why $N(r,s)$ attains its maximum as a function of s, for each fixed r, so that the set $S(r)$ is well defined.

Lemma 12.1 (Linear-in-r bound on optimal stretch factors for concave curves)**.** *If* $\sigma, \tau > -1$ *then*

$$S(r) \subset \big[(1+\tau)/rM, rL/(1+\sigma)\big] \qquad \text{whenever } r \geq (2+\sigma+\tau)/\sqrt{LM}.$$

Proof. The curve $r\Gamma(s)$ with the particular choice $s = \sqrt{L/M}$ has horizontal and vertical intercepts equal to $r\sqrt{LM}$. That intercept value is $\geq (2+\sigma+\tau)$, by assumption on r in this lemma. Hence by concavity, $r\Gamma(s)$ encloses the point $(1+\sigma, 1+\tau)$ and so $N(r,s) > 0$ for this particular value of s, which means the maximum of $s \mapsto N(r,s)$ is greater than 0.

When $s > rL/(1+\sigma)$, the x-intercept of $r\Gamma(s)$ is less than $1+\sigma$ and so no shifted lattice points are enclosed, meaning $N(r,s) = 0$. Thus the maximum is not attained for such s-values. Arguing similarly with the y-intercept shows the maximum is also not attained when $s < (1+\tau)/rM$. The lemma follows. $\square$

The last lemma required only that Γ be concave decreasing. Smoothness was not needed. Smoothness is not used in the next proposition either, which gives an upper bound on the counting function and so extends a result from the unshifted case Proposition 3.4.

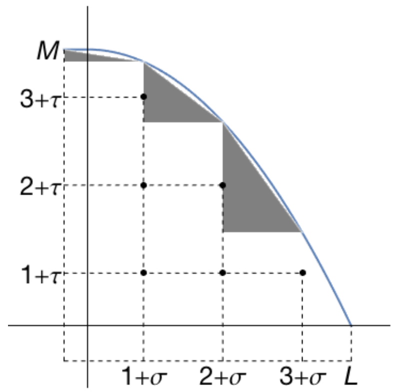

Figure 12.1: Concave curve enclosing lattice points shifted in the negative direction. The square areas represent the lattice point count, while the triangles estimate the discrepancy between that count and the area under the curve, as needed for Proposition 12.2

.

Proposition 12.2 (Two-term upper bound on counting function for concave curves)**.** *Let $\sigma, \tau > -1$. The number $N(r,s)$ of shifted lattice points lying inside $r\Gamma(s)$ satisfies*

$$N(r,s) \leq r^2 \operatorname{Area}(\Gamma) - C_1 rs + \sigma^- \tau^- \tag{12.1}$$

for all $r \geq (1-\sigma^-)s/L$ and $s \geq 1$, where

$$C_1 = C_1(\Gamma, \sigma, \tau) = \frac{1}{2}\Big(M - f(\frac{1-\sigma^-}{2-\sigma^-}L)\Big) - \sigma^- M - \tau^- L. \tag{12.2}$$

The constant C_1 might or might not be positive. Parameter Assumption 11.1 consists of the assumption $C_1 > 0$ along with the corresponding inequality for g, in the situation where $L = M$.

Proof. First suppose $\sigma \leq 0, \tau \leq 0$. Write N for the number of shifted lattice points under Γ, and suppose $L \geq 1+\sigma$ so that $\lfloor L-\sigma \rfloor \geq 1$. Extend the curve Γ horizontally from $(0, M)$ to (σ, M), so that $f(\sigma) = M$. Construct triangles with vertices at $\big(i-1+\sigma, f(i-1+\sigma)\big)$, $\big(i+\sigma, f(i+\sigma)\big)$, $\big(i-1+\sigma, f(i+\sigma)\big)$ for $i = 1, \ldots, \lfloor L-\sigma \rfloor$, as illustrated in Figure 12.1. The rightmost vertex of the final triangle has horizontal coordinate $\lfloor L-\sigma \rfloor + \sigma$, which is less than or equal to L. These triangles lie above the unit squares with upper right vertices at shifted lattice points, and lie below the curve Γ due to concavity. Hence

$$N + \text{Area(triangles)} \leq \text{Area}(\Gamma) - \sigma(M-\tau) - \tau(L-\sigma) - \sigma\tau, \qquad (12.3)$$

where the correction terms on the right side of the inequality represent the areas of the rectangular regions outside the first quadrant.

Letting $k = \lfloor L - \sigma \rfloor \geq 1$, we compute

$$\begin{aligned} \text{Area(triangles)} &= \sum_{i=1}^{k} \frac{1}{2}\big(f(i-1+\sigma) - f(i+\sigma)\big) \\ &= \frac{1}{2}\big(M - f(k+\sigma)\big) \\ &\geq \frac{1}{2}\big(M - f(\frac{1+\sigma}{2+\sigma}L)\big) \qquad (12.4) \end{aligned}$$

because f is decreasing and $k+\sigma \leq L < k+1+\sigma$ implies

$$k+\sigma > \frac{k+\sigma}{k+1+\sigma}L \geq \frac{1+\sigma}{2+\sigma}L.$$

Combining (12.3) and (12.4) proves

$$N \leq \text{Area}(\Gamma) - \sigma M - \tau L - \frac{1}{2}\Big(M - f(\frac{1+\sigma}{2+\sigma}L)\Big) + \sigma\tau. \qquad (12.5)$$

Now we replace Γ with the curve $r\Gamma(s)$, meaning we replace $N, L, M, f(x)$ with $N(r,s), rs^{-1}L, rsM, rsf(sx/r)$ respectively, thereby obtaining the desired estimate (12.1) (noting that $L/s \leq Ls$ since $s \geq 1$). The restriction $L \geq 1+\sigma$ becomes $r \geq (1+\sigma)s/L$ under the rescaling, and so we have proved the proposition in the case $\sigma \leq 0, \tau \leq 0$.

When $\sigma > 0, \tau > 0$, the number of shifted lattice points inside $r\Gamma(s)$ is less than or equal to the number when there is no shift ($\sigma = \tau = 0$), simply because the curve is decreasing. Thus this case of the proposition follows from the "$\sigma, \tau \leq 0$" case above.

When $\sigma > 0, \tau \leq 0$, the number of shifted lattice points inside $r\Gamma(s)$ is less than or equal to the number for $\sigma = 0$ with the same τ value, and so this case of the proposition follows also from the "$\sigma, \tau \leq 0$" case above. A similar argument holds when $\sigma \leq 0, \tau > 0$. □

Corollary 12.3 (Improved two-term upper bound on counting function for concave curves)**.** *Let $\sigma, \tau > -1$. If s is bounded above and bounded below away from 0, as $r \to \infty$, then the number $N(r,s)$ of shifted lattice points* *lying inside* $r\Gamma(s)$ *satisfies*

$$N(r,s) \le r^2 \operatorname{Area}(\Gamma) - r\big(s^{-1}\tau L + s(\sigma+1/2)M\big) + o(r). \tag{12.6}$$

Proof. Take $c > 1$ and suppose $c^{-1} < s < c$ throughout the rest of the proof.

Suppose $\sigma, \tau \le 0$. Let $K \ge 1$. Repeat the proof of Proposition 12.2 except with the initial supposition $L \ge 1 + \sigma$ replaced by $L \ge K + \sigma$, and do not assume $s \ge 1$. One finds

$$N(r,s) \le r^2 \operatorname{Area}(\Gamma) - D_K rs - \tau L r s^{-1} + \sigma\tau$$

for all $r \ge (K+\sigma)s/L$, where

$$D_K = D_K(\Gamma,\sigma) = \frac{1}{2}\Big(M - f(\frac{K+\sigma}{K+1+\sigma}L)\Big) + \sigma M.$$

We deduce

$$\begin{aligned}&\limsup_{r\to\infty}\sup_{s<c}\frac{1}{r}\Big(N(r,s) - r^2 \operatorname{Area}(\Gamma) + r\big(s(\sigma+1/2)M + s^{-1}\tau L\big)\Big)\\ &\le \frac{c}{2} f(\frac{K+\sigma}{K+1+\sigma}L).\end{aligned}$$

The last expression can be made arbitrarily small by choosing K sufficiently large (recall $f(L) = 0$), and so the left side is ≤ 0. That proves the corollary when $\sigma, \tau \le 0$.

Suppose $\sigma > 0, \tau \le 0$. We will relate this case to the previous one. To emphasize the dependence of the counting function on the shift parameters, write $N_{\sigma,\tau}(r,s)$ for the counting function that was previously written $N(r,s)$. Adding columns of shifted lattice points at $x = \sigma - \lceil\sigma\rceil + 1, \ldots, \sigma - 1, \sigma$ gives the counting function $N_{\widetilde{\sigma},\tau}(r,s)$ where $\widetilde{\sigma} = \sigma - \lceil\sigma\rceil \in (-1, 0]$. This counting function is related to the original one by

$$\begin{aligned}N_{\widetilde{\sigma},\tau}(r,s) &= N_{\sigma,\tau}(r,s) + \sum_{i=0}^{\lceil\sigma\rceil-1}\lfloor rsf\big(s(\sigma-i)/r\big) - \tau\rfloor,\\ &= N_{\sigma,\tau}(r,s) + \lceil\sigma\rceil rsM + o(r),\end{aligned}$$

as $r \to \infty$, since s is bounded above and f is continuous with $f(0) = M$.

Since $\widetilde{\sigma}, \tau \le 0$, we may apply (12.6) with σ replaced by $\widetilde{\sigma}$ to obtain

$$N_{\widetilde{\sigma},\tau}(r,s) \le r^2 \operatorname{Area}(\Gamma) - r\big(s^{-1}\tau L + s(\sigma - \lceil\sigma\rceil + 1/2)M\big) + o(r) \quad \text{as } r \to \infty.$$

Combining the above two formulas, we prove the corollary for $\sigma > 0, \tau \le 0$.

When $\sigma \le 0, \tau > 0$, simply add the appropriate rows instead of columns and argue like above using $\lceil\tau\rceil$ instead of $\lceil\sigma\rceil$, and using the boundedness of s^{-1}. Similarly, one can treat the case $\sigma > 0, \tau > 0$. □

The next proposition gives an asymptotic approximation to $N(r,s)$, assuming the curve is concave decreasing and has suitably monotonic second derivative.

Proposition 12.4 (Two-term counting estimate for concave curves)**.** *Let $\sigma, \tau > -1$ and $0 \le q < 1$. If Weaker Concave Condition 11.5 holds and $s + s^{-1} = O(r^q)$ then*

$$N(r,s) = r^2 \operatorname{Area}(\Gamma) - r\big(s^{-1}(\tau + 1/2)L + s(\sigma + 1/2)M\big) + O(r^Q) \quad (12.7)$$

as $r \to \infty$, where

$$Q = \max\{\tfrac{2}{3}, \tfrac{1}{2} + \tfrac{3}{2}q, 1 - 2a_1 + q, 1 - 2a_2 + \tfrac{3}{2}q, 1 - 2b_1 + q, 1 - 2b_2 + \tfrac{3}{2}q\}.$$

Special cases: (i) If $q = 0$ then $Q = 1 - 2e$ where $e = \min\{\frac{1}{6}, a_1, a_2, b_1, b_2\}$. (ii) If Concave Condition 11.3 holds then $Q = \max\{\frac{2}{3}, \frac{1}{2} + \frac{3}{2}q\}$.

The numbers a_1, a_2, b_1, b_2 come from Weaker Concave Condition 11.5. That Condition also involves a point $(\alpha, \beta) \in \Gamma$ with $\alpha, \beta > 0$, which we use in the following proof.

Proof. The idea is to translate and truncate the curve $r\Gamma(s)$ as in Figure 12.2, in order to reduce to an unshifted lattice problem. Then we invoke known results from the first part of the thesis.

Step 1 — Translating and truncating. Notice $rs \to \infty$ and $rs^{-1} \to \infty$ as $r \to \infty$, since $s = O(r^q)$ and $s^{-1} = O(r^q)$ with $q < 1$. Thus by taking r large enough, we insure

$$rs^{-1} g\Big(s^{-1}\frac{1+\tau}{r}\Big) > rs^{-1}\alpha > 1 + \sigma, \qquad rsf\Big(s\frac{1+\sigma}{r}\Big) > rs\beta > 1 + \tau.$$

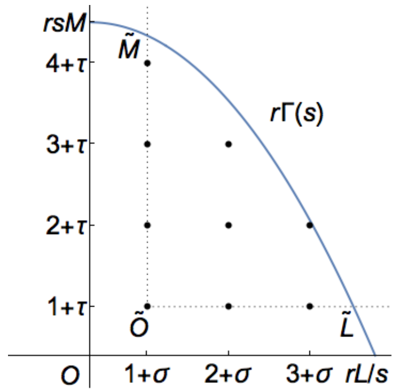

Figure 12.2: Curve $r\Gamma(s)$ enclosing positive-integer lattice points shifted by $(\sigma,\tau) = (-0.3, -0.4)$. The new origin is $\widetilde{O}$, and $\widetilde{L}$ and $\widetilde{M}$ are the new x- and y-intercepts, as defined in the proof of Proposition 12.4.

For all large r one also has $\delta(r) < \alpha$ and $\epsilon(r) < \beta$, by Weaker Concave Condition 11.5.

Given a large r satisfying the above conditions, and a corresponding $s > 0$, we let

$$\widetilde{\alpha} = rs^{-1}\alpha - (1+\sigma), \qquad \widetilde{\beta} = rs\beta - (1+\tau),$$

and

$$\widetilde{L} = rs^{-1}g\Big(s^{-1}\frac{1+\tau}{r}\Big) - (1+\sigma), \qquad \widetilde{M} = rsf\Big(s\frac{1+\sigma}{r}\Big) - (1+\tau),$$

so that

$$0 < \widetilde{\alpha} < \widetilde{L}, \qquad 0 < \widetilde{\beta} < \widetilde{M}.$$

Consider the point $\widetilde{O} = (1+\sigma, 1+\tau)$ in the first quadrant. Regard this point as the new origin, and let $\widetilde{\Gamma}$ be the portion of $r\Gamma(s)$ lying in the new first quadrant — see Figure 12.2. That is, $\widetilde{\Gamma}$ is the graph of

$$\widetilde{f}(x) = rsf\Big(s\frac{x+1+\sigma}{r}\Big) - (1+\tau), \qquad 0 \le x \le \widetilde{L},$$

and also of its inverse function

$$\widetilde{g}(y) = rs^{-1}g\Big(s^{-1}\frac{y+1+\tau}{r}\Big) - (1+\sigma), \qquad 0 \le y \le \widetilde{M}.$$

Notice $(\widetilde{\alpha}, \widetilde{\beta}) \in \widetilde{\Gamma}$, since $f(\widetilde{\alpha}) = \widetilde{\beta}$. Write $\widetilde{N}$ for the number of positive-integer lattice points under the curve $\widetilde{\Gamma}$. That is,

$$\widetilde{N} = \#\{(j,k) \in \mathbb{N} \times \mathbb{N} : k \le \widetilde{f}(j)\}.$$

This $\widetilde{N}$ does not count the lattice points in the first column or row, which arise from $j = 0$ or $k = 0$.

Weaker Concave Condition 11.5 guarantees that $\widetilde{f}$ is C^2-smooth on the interval $[0, \widetilde{\alpha}]$, with $\widetilde{f}' < 0$ and $\widetilde{f}'' < 0$ there, and similarly $\widetilde{g}$ is C^2-smooth on $[0, \widetilde{\beta}]$ with $\widetilde{g}' < 0$ and $\widetilde{g}'' < 0$ there.

Next, we partition the interval $[0, \widetilde{\alpha}]$ as $0 = \widetilde{\alpha}_0 < \widetilde{\alpha}_1 < \cdots < \widetilde{\alpha}_{\widetilde{l}} = \widetilde{\alpha}$ where the interior partition points are chosen to be the elements of

$$\{rs^{-1}\alpha_i - (1+\sigma) : i = 1, \ldots, l-1\}$$

that happen to lie between 0 and $\widetilde{\alpha}$. Observe $\widetilde{f}''$ is monotonic on each subinterval of the partition, by Weaker Concave Condition 11.5. Similarly, $\widetilde{g}''$ is monotonic on each subinterval of the corresponding partition $0 = \widetilde{\beta}_0 < \widetilde{\beta}_1 < \cdots < \widetilde{\beta}_{\widetilde{m}} = \widetilde{\beta}$ of the interval $[0, \widetilde{\beta}]$.

Let

$$\widetilde{\delta} = \big[rs^{-1}\delta(r) - (1+\sigma)\big]^+, \qquad \widetilde{\epsilon} = \big[rs\epsilon(r) - (1+\tau)\big]^+,$$

so that $0 \le \widetilde{\delta} < \widetilde{\alpha}$ and $0 \le \widetilde{\epsilon} < \widetilde{\beta}$.

To relate some of these old and new quantities, we denote antiderivatives of f, g by

$$F(x) = \int_0^x f(t)\,\mathrm{d}t, \qquad G(y) = \int_0^y g(t)\,\mathrm{d}t, \tag{12.8}$$

and observe that

$$\begin{aligned}
\mathrm{Area}(\widetilde{\Gamma}) &= r^2\,\mathrm{Area}(\Gamma) - r^2\big(F((1+\sigma)s/r) + G((1+\tau)s^{-1}/r)\big) \\
&\quad + (1+\sigma)(1+\tau), \\
\widetilde{f}'(x) &= s^2 f'\big(s\frac{x+1+\sigma}{r}\big), \qquad \widetilde{f}''(x) = \frac{s^3}{r} f''\big(s\frac{x+1+\sigma}{r}\big), \\
\int_0^{\widetilde{\alpha}} |\widetilde{f}''(x)|^{1/3}\,\mathrm{d}x &= r^{2/3}\int_{(1+\sigma)s/r}^{\alpha} |f''(x)|^{1/3}\,\mathrm{d}x \le r^{2/3}\int_0^{\alpha} |f''(x)|^{1/3}\,\mathrm{d}x, \\
\sum_{i=1}^{\widetilde{l}} \frac{1}{|\widetilde{f}''(\widetilde{\alpha}_i)|^{1/2}} &\le \sum_{i=1}^{l} \frac{r^{1/2}s^{-3/2}}{|f''(\alpha_i)|^{1/2}},
\end{aligned}$$

and similarly for $\widetilde{g}$ except with s replaced by s^{-1}.

Step 2 — Estimating the counting function. Applying part (a) of Proposition 4.1 the curve $\widetilde{\Gamma}$ and using the preceding relationships, we get

$$\begin{aligned}
&\big|\widetilde{N} - r^2\,\mathrm{Area}(\Gamma) + r^2\big(F((1+\sigma)s/r) + G((1+\tau)s^{-1}/r)\big) \\
&\qquad\qquad\qquad + \frac{r}{2}\big(sf((1+\sigma)s/r) + s^{-1}g((1+\tau)s^{-1}/r)\big)\big| \\
&\le 6r^{2/3}\Big(\int_0^\alpha |f''(x)|^{1/3}\,\mathrm{d}x + \int_0^\beta |g''(y)|^{1/3}\,\mathrm{d}y\Big) \\
&\quad + 175r^{1/2}\big(\frac{s^{-3/2}}{|f''(\delta(r))|^{1/2}} + \frac{s^{3/2}}{|g''(\epsilon(r))|^{1/2}}\big) \\
&\quad + 525r^{1/2}\big(\sum_{i=1}^{l}\frac{s^{-3/2}}{|f''(\alpha_i)|^{1/2}} + \sum_{j=1}^{m}\frac{s^{3/2}}{|g''(\beta_j)|^{1/2}}\big) + \frac{r}{2}(s^{-1}\delta(r) + s\epsilon(r)) \\
&\quad + \frac{1}{4}\big(\sum_{i=1}^{l} s^2|f'(\alpha_i)| + \sum_{j=1}^{m} s^{-2}|g'(\beta_j)|\big) \\
&\quad + l + m + \frac{1}{2}(1+\sigma) + \frac{1}{2}(1+\tau) + (1+\sigma)(1+\tau) + 1,
\end{aligned} \tag{12.9}$$

where we dealt with the term involving $|\widetilde{f}''(\widetilde{\delta})|^{-1/2}$ in (4.1) as follows. One has $\widetilde{f}''(\widetilde{\delta}) = r^{-1}s^3 f''(z)$ where $z = r^{-1}s(\widetilde{\delta}+1+\sigma) \ge \delta(r)$, and so by monotonicity of f'' on each subinterval of the partition (as assumed in Weaker Concave Condition 11.5) one concludes

$$|\widetilde{f}''(\widetilde{\delta})| \ge r^{-1}s^3 \min\{|f''\big(\delta(r)\big)|, |f''\big(\alpha_1\big)|, \dots, |f''\big(\alpha_l\big)|\}.$$

Thus the term involving $|\widetilde{f}''(\widetilde{\delta})|^{-1/2}$ can be estimated by the sum of terms involving $|f''\big(\delta(r)\big)|^{-1/2}$ and $|f''\big(\alpha_i\big)|^{-1/2}$.

The right side of (12.9) already has the desired order $O(r^Q)$, by direct estimation and using that $s + s^{-1} = O(r^q)$ and $2q < \frac{1}{2} + \frac{3}{2}q$ since $q < 1$.

Step 3 — Understanding the left side of inequality (12.9). It remains to deal with the terms on the left of (12.9). Clearly $N(r,s)$ and $\widetilde{N}$ count the same lattice points, except that $N(r,s)$ also counts the points in the first row and column. That is,

$$\begin{aligned}
\widetilde{N} &= N(r,s) - \lfloor rsf\big((1+\sigma)s/r\big) - \tau\rfloor - \lfloor rs^{-1}g\big((1+\tau)s^{-1}/r\big) - \sigma\rfloor + 1 \\
&= N(r,s) - rsf\big((1+\sigma)s/r\big) - \tau - rs^{-1}g\big((1+\tau)s^{-1}/r\big) - \sigma + \rho(r,s)
\end{aligned}$$

for some number $\rho(r,s) \in [1,3]$. Substitute this formula into the left side of (12.9). Substitute also the following expressions, which are obtained from Lemma 12.5:

$$\begin{aligned} rsf((1+\sigma)s/r) &= rsM + O(s^2), \\ r^2F((1+\sigma)s/r) &= rs(1+\sigma)M + O(s^2), \end{aligned}$$

and similarly for g and G. The proposition now follows straightforwardly, since $O(s^2) = O(r^{2q})$. □

Lemma 12.5. *If f is decreasing and concave on $[0,L]$ then*

$$f(x) = f(0) + O(x), \qquad F(x) = f(0)x + O(x^2), \qquad \text{as } x \to 0,$$

where $F(x) = \int_0^x f(t)\,\mathrm{d}t$ is the antiderivative of $f(x)$.

Proof. The difference quotient $(f(x) - f(0))/x$ is a decreasing function of x since f is concave, and it is less than or equal to 0 since f is decreasing. Hence the difference quotient is bounded, and so $f(x) = f(0) + O(x)$. Integrating completes the proof. □

CHAPTER 13

CONVEX CURVES — COUNTING FUNCTION ESTIMATES

Assume the curve Γ is convex decreasing, throughout this chapter. We will prove estimates for convex curves analogous to the work in Chapter 12 for concave curves.

Lemma 13.1 below is an improved r-dependent bound on the optimal stretch factors, generalizing Ariturk and Laugesen's lemma from the unshifted situation Lemma 7.2. By "improved" we refer to the upper and lower bounds: for instance, when $\sigma = 0$ the upper bound in Lemma 13.1 improves on the bound in Lemma 12.1 by a factor of 2. This tighter bound on the optimal stretch factor gives us more flexibility when deriving the two-term counting estimate in Proposition 13.2.

In the next lemma we assume for simplicity that the x- and y-intercepts are both L, so that we need not change the definitions of $\mu_f(\sigma)$ and $\mu_g(\tau)$ in (11.3) and (11.4).

Lemma 13.1 (Improved linear-in-r bound on optimal stretch factors for convex curves). *If $\sigma, \tau > -1$ with $\mu_f(\sigma) > 0$ and $\mu_g(\tau) > 0$, then*

$$S(r) \subset \left[\frac{2+\tau}{rL}, \frac{rL}{2+\sigma}\right]$$

whenever

$$r \geq \max\Big((2+\sigma)\sqrt{2(1+\tau)/L\mu_f(\sigma)}, (2+\tau)\sqrt{2(1+\sigma)/L\mu_g(\tau)}\Big). \tag{13.1}$$

Proof.

Claim 1: $N(r,s) = 0$ if $s \in \big(0, (1+\tau)/rL\big]$ or $s \in \big[rL/(1+\sigma), \infty\big)$. Indeed, the curve $r\Gamma(s)$ has x- and y-intercepts at rL/s and rsL, respectively, and so if $rL/s \leq 1+\sigma$ or $rsL \leq 1+\tau$ then the point $(1+\sigma, 1+\tau)$ is not enclosed by the curve and so the lattice count $N(r,s)$ is zero.

Claim 2: if (13.1) holds and $s \in \big(rL/(2+\sigma), rL/(1+\sigma)\big)$ then

$$N(r,s) < N\Big(r, \frac{1+\sigma}{2+\sigma}s\Big).$$

To prove this claim, notice the x-intercept satisfies

$$1+\sigma < \frac{rL}{s} < 2+\sigma,$$

and so only the first column of shifted lattice points (the points with x-coordinate at $1+\sigma$) can contribute to the count inside $r\Gamma(s)$. Hence $N(r,s) = \lfloor rsf((1+\sigma)s/r) - \tau \rfloor$. Meanwhile, if we count shifted lattice points in the first two columns (where $x = 1+\sigma$ and $x = 2+\sigma$) we find

$$\begin{aligned}
&N\Big(r, \frac{1+\sigma}{2+\sigma}s\Big) \qquad (13.2)\\
&\geq \Big\lfloor rs\frac{1+\sigma}{2+\sigma} f\Big(\frac{(1+\sigma)^2 s}{(2+\sigma)r}\Big) - \tau \Big\rfloor + \Big\lfloor rs\frac{1+\sigma}{2+\sigma} f\Big(\frac{(1+\sigma)s}{r}\Big) - \tau \Big\rfloor \\
&> rs\frac{1+\sigma}{2+\sigma} f\Big(\frac{(1+\sigma)^2 s}{(2+\sigma)r}\Big) + rs\frac{1+\sigma}{2+\sigma} f\Big(\frac{(1+\sigma)s}{r}\Big) - 2\tau - 2 \\
&= rsf\Big(\frac{(1+\sigma)s}{r}\Big) + \frac{rs}{2+\sigma}\Big((1+\sigma)f\Big(\frac{(1+\sigma)^2 s}{(2+\sigma)r}\Big) - f\Big(\frac{(1+\sigma)s}{r}\Big)\Big) - 2(1+\tau) \\
&\geq rsf\Big(\frac{(1+\sigma)s}{r}\Big) + \frac{rs}{2+\sigma}\mu_f(\sigma) - 2(1+\tau) \\
&> rsf\Big(\frac{(1+\sigma)s}{r}\Big) \geq N(r,s),
\end{aligned}$$

where to get the final line we use that $\frac{rs}{2+\sigma}\mu_f(\sigma) > 2(1+\tau)$, which follows from $s > rL/(2+\sigma)$ and the lower bound on r in (13.1). The proof of Claim 2 is complete.

Claim 3: if (13.1) holds and $s \in \big((1+\tau)/rL, (2+\tau)/rL\big)$ then

$$N(r,s) < N\Big(r, \frac{2+\tau}{1+\tau}s\Big).$$

The proof is analogous to Claim 2, except counting in rows instead of columns.

Claim 4: if (13.1) holds then the maximizing s-values for $N(r,s)$ lie in the interval $\big[(2+\tau)/rL, rL/(2+\sigma)\big]$. To see this, note that $N(r,s') > 0$ for some $s' > 0$, by the strict inequality in Claim 2, and so the maximum does not occur in the intervals considered in Claim 1. The maximum does not occur in the interval considered in Claim 2, as that claim itself shows, and similarly

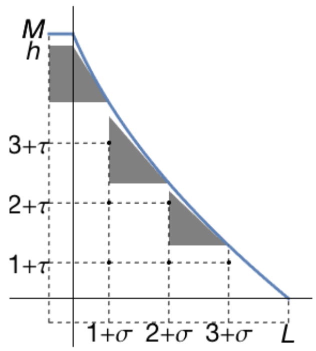

Figure 13.1: Convex curve enclosing lattice points shifted in the negative direction. The square areas represent the lattice point count, while the triangles and trapezoid estimate the discrepancy between that count and the area under the curve in Proposition 13.2

for Claim 3. Thus the maximum must occur in the remaining interval, which proves Claim 4 and thus finishes the proof of the lemma. □

The next bound generalizes work of Ariturk and Laugesen Proposition 7.3 from the unshifted situation ($\sigma = \tau = 0$) to the shifted case.

Proposition 13.2 (Two-term upper bound on counting function for convex curves)**.** *Let $\sigma, \tau > -1$. The number $N(r,s)$ of shifted lattice points lying inside $r\Gamma(s)$ satisfies*

$$N(r,s) \leq r^2 \operatorname{Area}(\Gamma) - C_2 rs + \sigma^- \tau^- \tag{13.3}$$

for all $r \geq (2-\sigma^-)s/L$ and $s \geq 1$, where

$$C_2 = C_2(\Gamma, \sigma, \tau) = \frac{1}{2}(1-\sigma^-) f\Big(\frac{1-\sigma^-}{2-\sigma^-} L\Big) - \sigma^- M - \tau^- L.$$

The constant C_2 need not be positive. That is why hypothesis (11.2) in Parameter Assumption 11.2 includes (for $L = M$) the assertion that $C_2 > 0$.

Proof. First consider $\sigma \leq 0, \tau \leq 0$. Write N for the number of shifted lattice points under Γ. Suppose $L \geq 2+\sigma$. Extend the curve horizontally from $(0, M)$ to (σ, M), so that $f(\sigma) = M$. Construct a trapezoid (see Figure 13.1) with vertices at $\big(\sigma, f(1+\sigma)\big)$, $\big(1+\sigma, f(1+\sigma)\big)$, $(0,h)$, (σ,h) where $h = f(1+\sigma) - (1+\sigma)f'(1+\sigma)$. Also construct triangles with vertices $\big(i-1+\sigma, f(i+\sigma)\big)$,

$\big(i+\sigma, f(i+\sigma)\big)$, $\big(i-1+\sigma, f(i+\sigma)-f'(i+\sigma)\big)$, where $i=2,\dots,\lfloor L-\sigma\rfloor$. These triangles lie above the squares with upper right vertices at the shifted lattice points, and like below the curve by convexity, as Figure 13.1 illustrates. Hence

$$N + \text{Area(trapezoid and triangles)} \leq \text{Area}(\Gamma) - \sigma(M-\tau) - \tau(L-\sigma) - \sigma\tau \tag{13.4}$$

Let $k = \lfloor L-\sigma \rfloor \geq 2$, so that $k+\sigma \leq L < k+\sigma+1$. Then

$$\begin{aligned}\text{Area(trapezoid)} &= \frac{1}{2}(\text{base}+\text{top})\cdot(\text{height}) \\ &= -\frac{1}{2}(1-\sigma)\cdot(1+\sigma)f'(1+\sigma) \\ &\geq \frac{1}{2}(1+\sigma)\big(f(1+\sigma)-f(2+\sigma)\big)\end{aligned}$$

by convexity, and using that $1-\sigma \geq 1$. Further, convexity implies

$$\begin{aligned}\text{Area(triangles)} &= -\frac{1}{2}\sum_{i=2}^{k} f'(i+\sigma) \\ &\geq \frac{1}{2}\sum_{i=2}^{k-1}\big(f(i+\sigma)-f(i+1+\sigma)\big) + \frac{1}{2}\big(f(k+\sigma)-f(L)\big) \\ &= \frac{1}{2}f(2+\sigma).\end{aligned} \tag{13.5}$$

Hence

$$\begin{aligned}&\text{Area(trapezoid)} + \text{Area(triangles)} \qquad (13.6)\\ &\geq \frac{1}{2}(1+\sigma)f(1+\sigma) - \frac{1}{2}\sigma f(2+\sigma) \\ &\geq \frac{1}{2}(1+\sigma)f(\frac{1+\sigma}{2+\sigma}L) - \frac{1}{2}\sigma f(\frac{2+\sigma}{2+\sigma}L) \qquad (13.7)\end{aligned}$$

since f is decreasing and $L/(2+\sigma) \geq 1$. Combining (13.4) and (13.7) and using $f(L)=0$ proves

$$N \leq \text{Area}(\Gamma) - \sigma M - \tau L - \frac{1}{2}(1+\sigma)f\big(\frac{1+\sigma}{2+\sigma}L\big) + \sigma\tau.$$

Now replace Γ with the curve $r\Gamma(s)$, meaning replace $N, L, M, f(x)$ with $N(r,s)$, $rs^{-1}L$, rsM, $rsf(sx/r)$ respectively. Using $s \geq 1$, we know $L/s \leq$

Ls; the assumption $L \geq 2 + \sigma$ becomes $r \geq (2 + \sigma)s/L$. Thus we obtain (13.3) in the case $\sigma \leq 0, \tau \leq 0$.

One may now deduce the remaining cases as was done in the proof of Proposition 12.2. □

Corollary 13.3 (Improved two-term upper bound on counting function for convex curves)**.** *Let* $\sigma, \tau > -1$. *If* s *is bounded above and bounded below away from* 0*, as* $r \to \infty$*, then the number* $N(r, s)$ *of shifted lattice points lying inside* $r\Gamma(s)$ *satisfies*

$$N(r, s) \leq r^2 \operatorname{Area}(\Gamma) - r\big(s^{-1}\tau L + s(\sigma + 1/2)M\big) + o(r). \tag{13.8}$$

Proof. Fix $c > 1$ and assume $c^{-1} < s < c$ in the rest of the proof.

Suppose $\sigma, \tau \leq 0$, and let $K \geq 2$. Repeat the proof of Proposition 13.2 except with the initial requirement $L \geq 2 + \sigma$ replaced by $L \geq K + \sigma$, and do not assume $s \geq 1$. The argument gives

$$N(r, s) \leq r^2 \operatorname{Area}(\Gamma) - E_K rs - \tau L r s^{-1} + \sigma\tau.$$

for all $r \geq (K + \sigma)s/L$, where

$$E_K = E_K(\Gamma, \sigma) = \frac{1}{2}(1 + \sigma) f\big(\frac{1 + \sigma}{K + \sigma} L\big) - \frac{1}{2}\sigma f\big(\frac{2 + \sigma}{K + \sigma} L\big) + \sigma M.$$

Hence

$$\begin{aligned} &\limsup_{r \to \infty} \sup_{s < c} \frac{1}{r}\Big(N(r, s) - r^2 \operatorname{Area}(\Gamma) + r\big(s(\sigma + 1/2)M + s^{-1}\tau L\big)\Big) \\ &\leq \frac{c}{2}\Big|M - (1 + \sigma) f\big(\frac{1 + \sigma}{K + \sigma} L\big) + \sigma f\big(\frac{2 + \sigma}{K + \sigma} L\big)\Big|. \end{aligned}$$

The last expression can be made arbitrarily small by choosing K sufficiently large (recall $f(0) = M$), and so the left side is ≤ 0, which proves the corollary when $\sigma, \tau \leq 0$.

By arguing as in the proof of Corollary 12.3, one handles the other three cases for σ and τ.

□

In the next proposition we state a two-term asymptotic for lattice point counting under convex curves.

Proposition 13.4 (Two-term counting estimate for convex curves)**.** *Let $\sigma, \tau > -1$. If Weaker Convex Condition 11.6 holds and $s + s^{-1} = O(1)$ then*

$$N(r,s) = r^2 \operatorname{Area}(\Gamma) - r\big(s^{-1}(\tau + 1/2)L + s(\sigma + 1/2)M\big) + O(r^{1-2\mathcal{E}}) \quad (13.9)$$

as $r \to \infty$, where $\mathcal{E} = \min\{\frac{1}{6}, a_1, a_2, a_3, b_1, b_2, b_3\}$. In particular, if Convex Condition 11.4 holds then (13.9) holds with $\mathcal{E} = \frac{1}{6}$.

Proposition 13.4 does not assume the intercepts L and M are equal, and so we modify Weaker Convex Condition 11.6 by taking each occurrence of "L" that relates to the function g and changing it to "M", and changing the a_3-condition to $f(x) = M + O(x^{2a_3})$.

Proof. We use the idea from Proposition 12.4: translate and truncate the curve $r\Gamma(s)$ to reduce to an unshifted lattice problem, and then use results from Ariturk and Laugesen's paper [5].

Assume $r\Gamma(s)$ does not pass through any point in the shifted lattice. This assumption will be removed in the final step of the proof.

Step 1 — Translating and truncating. Keep the notation from the proof of Proposition 12.4, except redefine the quantities $\widetilde{\delta}$ and $\widetilde{\epsilon}$ to be

$$\widetilde{\delta} = \big[\widetilde{L} + 1 + \sigma - rs^{-1}(L - \delta(r))\big]^+, \qquad \widetilde{\epsilon} = \big[\widetilde{M} + 1 + \tau - rs(M - \epsilon(r))\big]^+.$$

Arguing as in Step 1 of that proof, we have

$$0 < \widetilde{\alpha} < \lfloor \widetilde{L} \rfloor, \qquad 0 < \widetilde{\beta} < \lfloor \widetilde{M} \rfloor,$$

by taking r large enough, and also

$$0 \le \widetilde{\delta} < \lfloor \widetilde{L} \rfloor - \widetilde{\alpha}, \qquad 0 \le \widetilde{\epsilon} < \lfloor \widetilde{M} \rfloor - \widetilde{\beta}.$$

Step 2 — Estimating the counting function. Recall F represents the antiderivative of f, defined in (12.8). Applying part (a) of Proposition 7.5 to the curve $\widetilde{\Gamma}$ and using the relationships between the unshifted and shifted quantities as in the proof of Proposition 12.4, we get

$$
\begin{aligned}
&\Big|\widetilde{N} - r^2\,\mathrm{Area}(\Gamma) + r^2\Big(F\big((1+\sigma)s/r\big) + G\big((1+\tau)s^{-1}/r\big)\Big) \\
&\qquad + \frac{r}{2}\Big(sf\big((1+\sigma)s/r\big) + s^{-1}g\big((1+\tau)s^{-1}/r\big)\Big)\Big| \\
&\le 6r^{2/3}\Big(\int_\alpha^L f''(x)^{1/3}\,\mathrm{d}x + \int_\beta^M g''(y)^{1/3}\,\mathrm{d}y\Big) + 175r^{1/2}\Big(\frac{s^{-3/2}}{f''\big(L-\delta(r)\big)^{1/2}} \\
&\qquad + \frac{s^{3/2}}{g''\big(M-\epsilon(r)\big)^{1/2}}\Big) + 700r^{1/2}\Big(\sum_{i=0}^{l-1}\frac{s^{-3/2}}{f''(\alpha_i)^{1/2}} + \sum_{j=0}^{m-1}\frac{s^{3/2}}{g''(\beta_j)^{1/2}}\Big) \\
&\qquad + \frac{1}{4}\Big(\sum_{i=0}^{l-1}s^2|f'(\alpha_i)| + \sum_{j=0}^{m-1}s^{-2}|g'(\beta_j)|\Big) + \frac{1}{2}r\big(s^{-1}\delta(r) + s\epsilon(r)\big) \\
&\qquad + l + m + \frac{1}{2}(1+\sigma) + \frac{1}{2}(1+\tau) + (1+\sigma)(1+\tau) + 5 \\
&\qquad + \frac{rs^{-1}g((1+\tau)/rs) - (1+\sigma)}{rsf((1+\sigma)s/r) - (1+\tau)} + \frac{rsf((1+\sigma)s/r) - (1+\tau)}{rs^{-1}g((1+\tau)/rs) - (1+\sigma)},
\end{aligned}
\tag{13.10}
$$

where we estimated the term involving $\widetilde{f}''(\widetilde{L}-\widetilde{\delta})^{-1/2}$ as follows. One has $\widetilde{f}''(\widetilde{L}-\widetilde{\delta}) = r^{-1}s^3 f''(z)$ where

$$
z = r^{-1}s(\widetilde{L} - \widetilde{\delta} + 1 + \sigma) \le L - \delta(r),
$$

and so by monotonicity of f'' on each subinterval of the partition (as assumed in Weaker Convex Condition 11.6) one concludes

$$
\widetilde{f}''(\widetilde{L}-\widetilde{\delta}) \ge r^{-1}s^3 \min\{f''\big(L-\delta(r)\big), f''\big(\alpha_0\big), \ldots, f''\big(\alpha_{l-1}\big)\}.
$$

Thus the term involving $\widetilde{f}''(\widetilde{L}-\widetilde{\delta})^{-1/2}$ can be estimated by the sum of terms involving $f''(L-\delta(r))^{-1/2}$ and $f''(\alpha_i)^{-1/2}$.

The right side of (13.10) has the form $O(r^{1-2e})$, by arguing directly with $s + s^{-1} = O(1)$ and the assumptions in Weaker Convex Condition 11.6, and estimating the last two terms in (13.10) by

$$
\frac{rs^{-1}g((1+\tau)/rs) - (1+\sigma)}{rsf((1+\sigma)s/r) - (1+\tau)} = \frac{s^{-1}L - o(1)}{sM - o(1)} = O(1)
$$

and similarly with f and g interchanged.

Step 3 — Understanding the left side of inequality (13.10). The terms on the left of (13.10) are dealt with in the same manner as in Step 3 of Proposition 12.4, except replacing Lemma 12.5 with the last assumption in Weaker Convex Condition 11.6, as follows. Substituting $x = (1+\sigma)s/r$ into $f(x) = M + O(x^{2a_3})$ and into $F(x) = Mx + O(x^{1+2a_3})$ gives

$$\begin{aligned} rsf((1+\sigma)s/r) &= rsM + O(r^{1-2a_3}), \\ r^2F((1+\sigma)s/r) &= rs(1+\sigma)M + O(r^{1-2a_3}), \end{aligned}$$

since $s + s^{-1} = O(1)$. One argues similarly for g and G. Thus we have finished the proof under the assumption that $r\Gamma(s)$ passes through no lattice points.

Step 4 — Finishing the proof. Now drop the assumption that $r\Gamma(s)$ passes through no lattice points. Notice the counting function $N(r,s)$ is increasing in the r-variable. Fix the r and s values, and modify the functions $\delta(\cdot)$ and $\epsilon(\cdot)$ to be continuous at r. For sufficiently small $\eta > 0$ we have $N(r+\eta,s) = N(r,s)$, because the r-variable would have to increase by some positive amount for the curve $r\Gamma(s)$ to reach any new lattice points. Since no lattice points lie on the curve $(r+\eta)\Gamma(s)$, Steps 1–3 above apply to that curve. Hence by continuity as $\eta \to 0$, the conclusion of the proposition holds also for $r\Gamma(s)$. $\square$

CHAPTER 14

OPTIMAL CONCAVE OR CONVEX CURVE IS BOUNDED

Our first theorem Theorem 14.2 will say that the maximizing set $S(r)$ is bounded, under either of the Parameter Assumption 11.1 or Parameter Assumption 11.2.

In order to prove boundedness of the maximizing set, we need a rough lower bound on the counting function,. Assume the curve Γ is strictly decreasing in the first quadrant, and has x- and y-intercepts at L and M. The intercepts need not be equal, in the next lemma.

Lemma 14.1 (Rough lower bound for decreasing curve)**.** *The number $N(r,s)$ of shifted lattice points lying inside $r\Gamma(s)$ satisfies*

$$N(r,s) \geq r^2 \operatorname{Area}(\Gamma) - r\big(s^{-1}(1+\tau)L + s(1+\sigma)M\big), \qquad r,s>0. \quad (14.1)$$

Proof. We split the proof into two cases, and later rescale to handle the general curve. Write N for the number of shifted lattice points under Γ.

Case I: The point $(1+\sigma, 1+\tau)$ lies outside the curve Γ, and so $N = 0$. Then the rectangles with vertices $(0,0),(L,0),(L,1+\tau),(0,1+\tau)$ and $(0,0),(1+\sigma,0),(1+\sigma,M),(0,M)$ cover Γ since the curve is decreasing, and so by comparing areas one has

$$N + (1+\tau)L + (1+\sigma)M \geq \operatorname{Area}(\Gamma). \quad (14.2)$$

Case II: The point $(1+\sigma,1+\tau)$ lies inside the curve. We shift the origin to $\widetilde{O} = (1+\sigma,1+\tau)$ and draw new axes, denoting the x- and y-intercepts on the new axes by $\widetilde{L}$ and $\widetilde{M}$; see Figure 14.1. The part of Γ lying in the new first quadrant is $\widetilde{\Gamma}$. Each lattice point corresponds to a square whose lower left vertex sits at that point. These squares cover $\widetilde{\Gamma}$ since the curve is strictly decreasing. The remaining area under Γ is covered by the two rectangles described in Case I. The sum of the areas of the squares and rectangles must

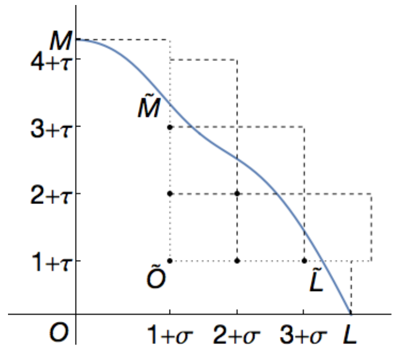

Figure 14.1: Decreasing curve Γ enclosing positive integer lattice points shifted by amount $(\sigma, \tau) = (0.4, -0.2)$. We shift the origin by $1+\sigma, 1+\tau$, obtaining a new origin $\widetilde{O}$, with $\widetilde{L}$ and $\widetilde{M}$ being the new x- and y-intercepts. The lattice point count equals the area of the squares, as used in proving Lemma 14.1.

exceed the area under Γ, and so (14.2) holds once again.

To complete the proof, simply replace the curve Γ with $r\Gamma(s)$, meaning that in (14.2) we replace N, L, M with $N(r,s)$, $rs^{-1}L$, rsM respectively. The lemma follows. □

Theorem 14.2 (Uniform bound on optimal stretch factors)**.** *If the curve Γ and the shift parameters $\sigma, \tau > -1/2$ satisfy Parameter Assumption 11.1 or 11.2, then for each $\varepsilon > 0$ one has*

$$S(r) \subset \left[B(\tau,\sigma)^{-1} - \varepsilon, B(\sigma,\tau) + \varepsilon\right] \qquad \textit{for all large } r,$$

where

$$B(\sigma,\tau) = \frac{2+\sigma+\tau+\sqrt{(2+\sigma+\tau)^2 - 4(\sigma+1/2)\tau)}}{2(\sigma+1/2)}.$$

The bounding constant $B(\sigma,\tau)$ depends only on the shift parameters, not on the curve Γ. The bounding constant $B(0,0) = 4$ in the unshifted case is consistent with our earlier work [21, Theorem 2].

Proof. We prove the theorem in two parts: first for concave curves, and then for convex curves. When Γ is concave, we will utilize the bound on $S(r)$ in Lemma 12.1 and the two-term upper bound on the counting function in Proposition 12.2, along with the improved upper bound in Corollary 12.3 and the rough lower bound on the counting function in Lemma 14.1.

Recall the intercepts are assumed equal ($L = M$) in this theorem.

Part 1: Γ is concave and Parameter Assumption 11.1 holds

The proof has two steps. Step 1 shows $S(r)$ is bounded above and below away from 0, for large r. Step 2 uses this boundedness to improve the asymptotic bound on $S(r)$, revealing that it depends only on σ and τ and not the curve Γ.

Step 1. Take $s \in S(r)$ and suppose $r \geq (2 + \sigma + \tau)/L$. Then Lemma 12.1 says $s \leq rL/(1 + \sigma)$, so that

$$r \geq \frac{(1+\sigma)s}{L} \geq \frac{(1-\sigma^-)s}{L}.$$

If $s \geq 1$ then Proposition 12.2 implies

$$N(r,s) \leq r^2 \operatorname{Area}(\Gamma) - C_1 rs + \sigma^- \tau^-.$$

Parameter Assumption 11.1 guarantees here that $C_1 > 0$.

The lower bound in Lemma 14.1 with "$s = 1$" says

$$N(r,1) \geq r^2 \operatorname{Area}(\Gamma) - (2 + \sigma + \tau)Lr. \tag{14.3}$$

Since $s \in S(r)$ is a maximizing value, one has $N(r,s) \geq N(r,1)$, and so the preceding two inequalities give

$$s \leq \frac{(2+\sigma+\tau)L}{C_1} + \frac{\sigma^-\tau^- L}{(2+\sigma+\tau)C_1}$$

when $r \geq (2 + \sigma + \tau)/L$ and $s \geq 1$. Thus $S(r)$ is bounded above for all large r.

Similarly if $s \in S(r)$ then s^{-1} is bounded above, by interchanging the roles of the horizontal and vertical axes in the argument above. Thus the set $S(r)$ is bounded below away from 0, for large r.

Step 2. The number

$$\overline{s} = \limsup_{s \in S(r), r \to \infty} s$$

is finite and positive by Step 1. Combining the inequality $N(r,s) \geq N(r,1)$ with estimate (14.3) and Corollary 12.3 (which relies on the boundedness of $S(r)$) we obtain

$$(\sigma + 1/2)\overline{s}^2 - (2 + \sigma + \tau)\overline{s} + \tau \leq 0$$

after letting $r \to \infty$. Notice $\sigma + 1/2 > 0$ by hypothesis in Theorem 14.2. Hence $\overline{s}$ is bounded above by the larger root of the quadratic; that is,

$$\overline{s} \leq B(\sigma, \tau) = \frac{2 + \sigma + \tau + \sqrt{(2 + \sigma + \tau)^2 - 4(\sigma + 1/2)\tau}}{2(\sigma + 1/2)}.$$

Similarly $\limsup_{r \to \infty} s^{-1} \leq B(\tau, \sigma)$, by interchanging the roles of the axes. The proof of Theorem 14.2 is complete, in the concave case.

Part 2: Γ is convex and Parameter Assumption 11.2 holds

Take $s \in S(r)$ and suppose r satisfies (13.1), recalling there that $\mu_f(\sigma)$ and $\mu_g(\tau)$ are positive by Parameter Assumption 11.2. Now proceed as in Part 1 of the proof, simply replacing Lemma 12.1, Proposition 12.2 and Corollary 12.3 with Lemma 13.1, Proposition 13.2 and Corollary 13.3, respectively. □

CHAPTER 15

OPTIMAL CONCAVE OR CONVEX CURVE IS ASYMPTOTICALLY BALANCED

If the curve is smooth, then the optimal stretch set $S(r)$ for maximizing the lattice count is not only bounded but converges asymptotically to a computable value, as stated in the next theorem.

Theorem 15.1 (Weaker conditions for asymptotic balance of optimal curve)**.** *If the curve Γ and shift parameters $\sigma, \tau > -1/2$ satisfy either Parameter Assumption 11.1 and Weaker Concave Condition 11.5, or Parameter Assumption 11.2 and Weaker Convex Condition 11.6, then the stretch factors maximizing $N(r,s)$ approach*

$$s^* = \sqrt{\frac{\tau + 1/2}{\sigma + 1/2}}$$

as $r \to \infty$, with

$$S(r) \subset \left[s^* - O(r^{-\mathcal{E}}), s^* + O(r^{-\mathcal{E}})\right]$$

where

$$\mathcal{E} = \min\{\tfrac{1}{6}, a_1, a_2, a_3, b_1, b_2, b_3\}.$$

Further, the maximal lattice count has asymptotic formula

$$\max_{s>0} N(r,s) = r^2 \operatorname{Area}(\Gamma) - 2rL\sqrt{(\sigma + 1/2)(\tau + 1/2)} + O(r^{1-2\mathcal{E}}). \quad (15.1)$$

We call the optimally stretched curve ($s = s^*$) "asymptotically balanced" in terms of the shift parameters, because the optimal shape balances the areas of the empty strips that are created by translation of the lattice: a horizontal rectangle of width rL/s^* and height $\tau + 1/2$ has the same area as a vertical rectangle of height rs^*L and width $\sigma + 1/2$. (The "+1/2" arises from thinking of each lattice point as the center of a unit square.) Further, subtracting these two areas, each of which equals $rL\sqrt{(\sigma + 1/2)(\tau + 1/2)}$, gives a heuristic derivation of the order-r correction term in the theorem.

Proof. Recall the intercepts are equal, $L = M$, in this theorem.

The optimal stretch parameters are bounded above and bounded below away from 0 as $r \to \infty$, by Theorem 14.2. (It suffices to use the curve-dependent bound from Step 1 of that proof; we do not need the curve-independent bound $B(\sigma, \tau)$ from Step 2.)

Hence by Proposition 12.4 (if Γ is concave) or Proposition 13.4 (if Γ is convex),

$$N(r,s) = r^2 \operatorname{Area}(\Gamma) - rL\big(s^{-1}(\tau + 1/2) + s(\sigma + 1/2)\big) + O(r^{1-2\mathcal{E}}) \tag{15.2}$$

when $s \in S(r)$; this estimate holds also when $s > 0$ is any fixed value. Thus for $s \in S(r)$ and $s^* = \sqrt{(\tau + 1/2)/(\sigma + 1/2)}$ we have

$$\begin{aligned} N(r,s) &\le r^2 \operatorname{Area}(\Gamma) - rL\big(s^{-1}(\tau + 1/2) + s(\sigma + 1/2)\big) + O(r^{1-2\mathcal{E}}), \\ N(r,s^*) &\ge r^2 \operatorname{Area}(\Gamma) - 2rL\sqrt{(\tau + 1/2)(\sigma + 1/2)} + O(r^{1-2\mathcal{E}}), \end{aligned}$$

as $r \to \infty$. Notice $N(r,s^*) \le N(r,s)$ because $s \in S(r)$ is a maximizing value, and so

$$s^{-1}(\tau + 1/2) + s(\sigma + 1/2) \le 2\sqrt{(\tau + 1/2)(\sigma + 1/2)} + O(r^{-2\mathcal{E}}). \tag{15.3}$$

Therefore $s = s^* + O(r^{-\mathcal{E}})$, by Lemma 15.2 below with $a = \tau + 1/2, b = \sigma + 1/2$.

For the final statement of the theorem, when $s \in S(r)$ one has

$$\begin{aligned} 2\sqrt{(\tau + 1/2)(\sigma + 1/2)} &\le s^{-1}(\tau + 1/2) + s(\sigma + 1/2) \\ &\le 2\sqrt{(\tau + 1/2)(\sigma + 1/2)} + O(r^{-2\mathcal{E}}) \end{aligned}$$

by the arithmetic–geometric mean inequality and (15.3). Multiplying by rL and substituting into (15.2) gives the asymptotic formula (15.1). $\square$

Lemma 15.2. *When $a, b, s > 0$ and $0 \le t \le \sqrt{ab}$,*

$$s^{-1}a + sb \le 2\sqrt{ab} + t \quad \Longrightarrow \quad \big|s - \sqrt{a/b}\big| \le \frac{3(ab)^{1/4}}{b}\sqrt{t}.$$

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

$$((s^{-1}a)^{1/2} - (sb)^{1/2})^2 = s^{-1}a + sb - 2\sqrt{ab} \le t$$

and then using that the number $(ab)^{1/4}$ lies between $(s^{-1}a)^{1/2}$ and $(sb)^{1/2}$ (because it is their geometric mean), we find

$$|(ab)^{1/4} - (sb)^{1/2}| \leq t^{1/2}.$$

Hence $(ab)^{1/4} - t^{1/2} \leq (sb)^{1/2} \leq (ab)^{1/4} + t^{1/2}$. Squaring and using that $t \leq (ab)^{1/4} t^{1/2}$ (when $t \leq \sqrt{ab}$) proves the lemma. □

The C^2-smoothness hypothesis could be weakened to piecewise smoothness, as was done for concave curves in Part I. The theorem simplifies somewhat when the second derivatives are negative (or positive) and monotonic all the way up to the endpoints:

Corollary 15.3 (Sufficient conditions for asymptotic balance of optimal curve)**.** *If the curve* Γ *and shift parameters* $\sigma, \tau > -1/2$ *satisfy either Parameter Assumption 11.1 and Concave Condition 11.3, or Parameter Assumption 11.2 and Convex Condition 11.4, then the stretch factors maximizing* $N(r,s)$ *approach*

$$s^* = \sqrt{\frac{\tau + 1/2}{\sigma + 1/2}}$$

as $r \to \infty$*, with*

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

and the maximal lattice count has asymptotic formula

$$\max_{s>0} N(r,s) = r^2 \operatorname{Area}(\Gamma) - 2rL\sqrt{(\sigma + 1/2)(\tau + 1/2)} + O(r^{2/3}).$$

In particular, when the shift parameters σ *and* τ *are equal, the optimal stretch factors for maximizing* $N(r,s)$ *approach* $s^* = 1$ *as* $r \to \infty$*.*

Proof. Concave Condition 11.3 implies Weaker Concave Condition 11.5, by choosing $\delta(r) = \epsilon(r) = r^{-1}$, $a_1 = b_1 = 1/2$ and $a_2 = b_2 = 1/4$, and noting that $f''(0) \neq 0, g''(0) \neq 0$. The same reasoning shows Convex Condition 11.4 implies Weaker Convex Condition 11.6 with $a_3 = b_3 = 1/4$, since

$$g(L) = 0,\ g'(L) \leq 0,\ g''(L) > 0 \quad \Longrightarrow \quad g(L - y) \geq cy^2 \text{ for small } y > 0$$

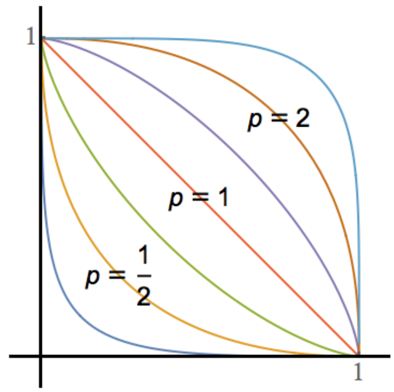

Figure 15.1: The family of p-circles $x^p + y^p = 1$, for $0 < p < \infty$.
Example 15.4 and Example 15.5 consider $p = 2$ and $p = 1/2$, respectively.

where $c > 0$, and substituting $y = \sqrt{x/c}$ gives $L - f(x) \leq \sqrt{x/c}$ for small $x > 0$, and similarly for g.

Thus Corollary 15.3 follows immediately from Theorem 15.1. □

Example 15.4 (Sufficient condition on shift parameters for the circle)**.** When the curve Γ is the portion of the unit circle in the first quadrant, one takes $L = 1, f(x) = \sqrt{1 - x^2}$, and $\alpha = \beta = 1/\sqrt{2}$. Notice f is smooth and concave, with monotonic second derivative. By symmetry it suffices to consider $\sigma \leq \tau$. When $\sigma \leq \tau \leq 0$, Parameter Assumption 11.1 says

$$\sqrt{1 - (\frac{1+\sigma}{2+\sigma})^2} < 2\sigma + 2\tau + 1.$$

When $\sigma \leq 0 \leq \tau$, equality in Parameter Assumption 11.1 would give a straight line. The resulting allowable region of (σ, τ)-shift parameters for Corollary 15.3 is plotted on the left side of Figure 15.2.

Example 15.5 (Sufficient condition on shift parameters for p-circle with $p = 1/2$)**.** Suppose Γ is the part of the 1/2-circle lying in the first quadrant, so that $L = 1, f(x) = (1 - x^{1/2})^2$, and take $\alpha = \beta = 1/4$. Notice f is smooth and convex, with monotonic second derivative $f''(x) = \frac{1}{2}x^{-3/2}$. The region of allowable shift parameters for Corollary 15.3 can be found numerically from Parameter Assumption 11.2, as shown on the right side of Figure 15.2.

Example 15.6 (p-circles)**.** Suppose Γ is the part of the p-circle $|x|^p + |y|^p = 1$ lying in the first quadrant. When $p > 1$ the curve is concave, and satisfies

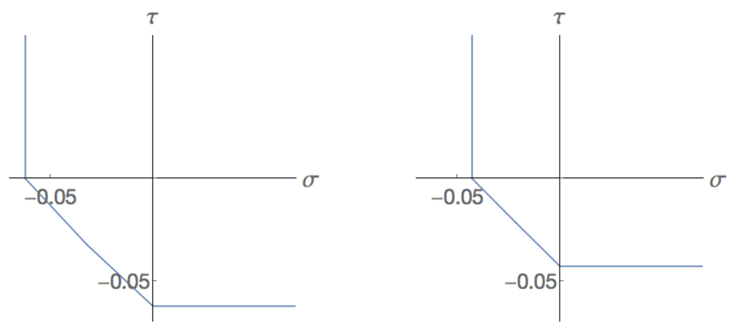

Figure 15.2: The allowable shift parameters (σ, τ) for Corollary 15.3 form the regions above the plotted curves, in the special cases where Γ is a circle (figure on the left) and a p-circle with $p = 1/2$ (figure on the right). The intercepts are at approximately -0.06 (left figure) and -0.04 (right figure). The straight lines in the second and fourth quadrants are vertical and horizontal, respectively. The curves joining the intercepts are not quite straight lines. See Example 15.4 and Example 15.5.

Weaker Concave Condition 11.5 by Example 5.5. When $0 < p < 1$ it is convex and satisfies Weaker Convex Condition 11.6 by Example 7.8, noting that $a_3 = b_3 = p/2$ since $f(x) = 1 + O(x^p)$ as $x \to 0$ and $g(y) = 1 + O(y^p)$ as $y \to 0$.

Thus Theorem 15.1 applies to each p-circle, $p \neq 1$. The allowable shift parameters can be determined numerically from Parameter Assumption 11.1 or 11.2, as in Example 15.4 and Example 15.5.

CHAPTER 16

NEGATIVE SHIFTS

In this chapter we show there can be no "universal" allowable region of negative shifts for Corollary 15.3. Specifically, for each choice of negative shifts $\sigma, \tau < 0$, no matter how close to zero, a curve exists whose optimal stretch parameters grow to infinity or else shrink to 0 as $r \to \infty$. That is, the optimal curve degenerates in the limit.

Theorem 16.1 (Negative shift: optimal concave curve can degenerate)**.** *If $-1 < \sigma < 0, \tau > -1$, then a concave C^2-smooth curve Γ exists, with intercepts at $L = 1$, such that for each $\epsilon \in (0, 1)$ one has*

$$S(r) \subset (0, r^{\epsilon-1}) \cup (r^{1-\epsilon}, \infty) \tag{16.1}$$

for all large r.

Proof. Fix $\sigma \in (-1, 0)$ and $\tau > -1$. Since $0 < 1 + \sigma < 1$, we may choose $m \in \mathbb{N}$ large enough that

$$(1 + \sigma)^{2m} < \frac{1}{2m + 1}. \tag{16.2}$$

Defining $\phi(x) = 1-x^{2m}$ for $0 \leq x \leq 1$, one checks $\phi(1+\sigma) >$ area under graph of ϕ. Thus one may choose $0 < \delta < 1$ small enough that the function

$$f(x) = 1 - \delta x^2 - (1 - \delta)x^{2m}, \qquad 0 \leq x \leq 1,$$

satisfies

$$f(1 + \sigma) > \text{area under graph of } f.$$

Observe f is smooth and strictly decreasing, with $f'' < 0$ on $[0, 1]$, so that its graph Γ is concave. The inverse function g satisfies the same conditions.

The curve $r\Gamma(r)$ is the graph of $r^2 f(x)$ for $0 \leq x \leq 1$. This curve contains only the first column of shifted lattice points (the points with x-coordinate

$1 + \sigma)$, and so

$$\begin{aligned} N(r,r) &= \lfloor r^2 f(1+\sigma) - \tau \rfloor \\ &\geq r^2 f(1+\sigma) - \tau - 1. \end{aligned}$$

Now fix $0 < \epsilon < 1$. If $s \in [r^{\epsilon-1}, r^{1-\epsilon}]$ then $s + s^{-1} = O(r^{1-\epsilon})$, and so Proposition 12.4 with $q = 1 - \epsilon$ and $L = M = 1$ gives that

$$\begin{aligned} N(r,s) &= r^2 \operatorname{Area}(\Gamma) - r\big(s(\sigma + 1/2) + s^{-1}(\tau + 1/2)\big) + O(r^{2-3\epsilon/2}) \\ &= r^2 \operatorname{Area}(\Gamma) + o(r^2). \end{aligned}$$

Since $\operatorname{Area}(\Gamma) < f(1+\sigma)$, we conclude that for all large r,

$$N(r,s) < N(r,r)$$

and so $s \notin S(r)$, which proves the theorem. □

The point of the theorem is that as soon as one of the shift parameters is negative, a concave curve exists for which the maximizing stretch parameters approach either 0 or ∞ as $r \to \infty$.

For convex curves, we do not know an analogue of Theorem 16.1: does a universal allowable region of (σ, τ) parameters exist in which Corollary 15.3 holds for all C^2-smooth convex decreasing curves?

The curve in Theorem 16.1 can even be a quarter circle:

Proposition 16.2 (Negative shift: the optimal ellipse can degenerate)**.** *If the curve Γ is the quarter unit circle and $\sigma = \tau = -2/5$, then for each $\epsilon \in (0,1)$ one has*

$$S(r) \subset (0, r^{\epsilon-1}) \cup (r^{1-\epsilon}, \infty) \qquad \textit{for all large } r.$$

Proof. The argument is the same as for Theorem 16.1, except now the curve is a quarter circle, described by $f(x) = \sqrt{1-x^2}$. The only point to check in the proof is that

$$f(1+\sigma) > \operatorname{Area}(\Gamma)$$

when $\sigma = -2/5$, which reduces to the fact that $4/5 > \pi/4$. □

CHAPTER 17

NUMERICAL EVIDENCE, AND CONJECTURES FOR TRIANGLES $(P = 1)$

Figure 17.1(a) illustrates the convergence of $s \in S(r)$ to s^*, when Γ is a quarter circle and the shifts are positive. The convergence is erratic, while still obeying the decay rate $O(r^{-1/6})$ as promised by Corollary 15.3.

Figure 17.1(b) shows the degeneration that can occur when the shifts are negative, as explained in Proposition 16.2.

Quite different behavior occurs when Γ is a straight line with slope -1, in other words, when the curve is the 1-ellipse described by $f(x) = 1 - x$, which is not covered by our results in Example 15.6. Here $N(r, s)$ counts the shifted lattice points inside the right triangle with vertices at $(r/s, 0), (0, rs)$ and the origin. Theorem 14.2 insures the maximizing set $S(r)$ is bounded above and below, being contained in $\left[B(\tau, \sigma)^{-1} - \varepsilon, B(\sigma, \tau) + \varepsilon\right]$ for all large r. This boundedness depends on Parameter Assumption 11.1 holding, which in this case says

$$(2 - \max(\sigma^-, \tau^-))(1 - 2\sigma^- - 2\tau^-) > 1.$$

In particular, $S(r)$ is bounded for the 1-ellipse if $\sigma = \tau > -0.117$. Our convergence theorems for $S(r)$ do not apply, though, since $f'' \equiv 0$ for the straight line. In fact, the numerical plots in Figure 6.2 suggest $S(r)$ might not converge, and might instead cluster at many different heights, as discussed in the unshifted case in Chapter 6. Are those heights determined by a number theoretic property of some kind?

The numerical method that generated the figures is described in Chapter 6 for $p = 1$. It adapts easily to handle other values of p, in particular $p = 2$ (the circle). See Appendix B for the code.

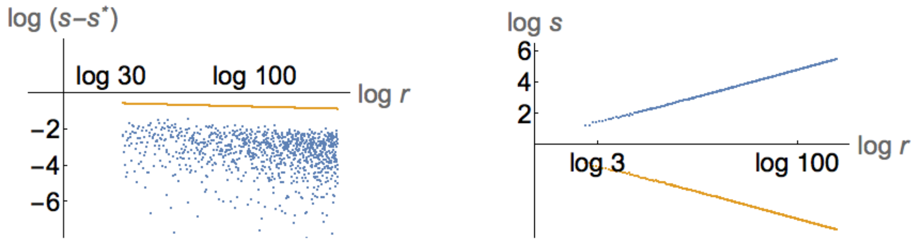

Figure 17.1: Maximizing s-values for the number of lattice points in the 2-ellipse. (a) Left figure: positive shift $\sigma = 1, \tau = 3$. The plot shows $\log(\sup S(r) - s^*)$ versus $\log r$. The line $-1/6 \log r$ indicates the guaranteed convergence rate in Corollary 15.3. (b) Right figure: negative shifts $\sigma = \tau = -2/5$. The plot shows $\log(\sup S(r))$ and $\log(\inf S(r))$ versus $\log r$. Linear fitting gives $\log \sup S(r) \simeq 0.982 \log r + 0.254$, which is consistent with the growth rate $s \gtrsim r^{1-\epsilon}$ proved in Proposition 16.2. In both plots, the r-values are multiples of $\sqrt{3}/10$, an irrational number chosen in the hope of exhibiting generic behavior.

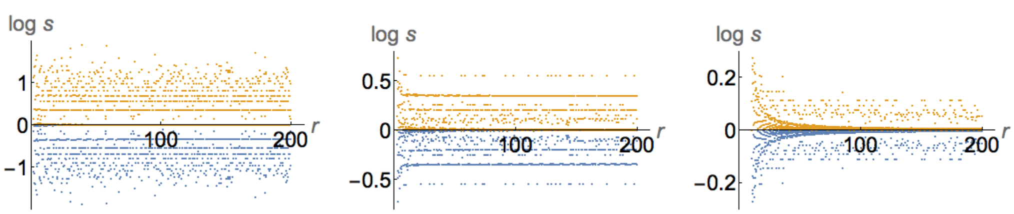

Figure 17.2: Maximizing s-values for the number of lattice points in the 1-ellipse (that is, the right triangle). The upper plots show $\log \sup S(r)$ versus r and the lower plots are $\log \inf S(r)$ versus r. The figure on the left is for shift parameters $\sigma = \tau = -1/2$, which corresponds to counting eigenvalues of the harmonic oscillator (Chapter 18). The middle figure has $\sigma = \tau = 0$, and the figure on the right has $\sigma = \tau = 4$. Notice the optimal stretch parameters are bounded in a narrower and narrower band as the shift parameters increase.

CHAPTER 18

FUTURE DIRECTIONS — OPTIMAL QUANTUM OSCILLATORS

Schrödinger equations

A natural quantum analogue of the eigenvalue minimization theorem for the Dirichlet Laplacian on rectangles is to minimize the n-th energy level among a family of harmonic oscillators. For each $s > 0$, consider

$$-\Delta u + \frac{1}{4}\big((sx)^2 + (s^{-1}y)^2\big)u = Eu, \qquad x, y \in \mathbb{R}, \tag{18.1}$$

with boundary condition $u \to 0$ as $|(x,y)| \to \infty$. Write $S(n)$ for the set of s-values that minimize the n-th eigenvalue E_n. By analogy with the result for rectangles, one might conjecture that $s \in S(n)$ must approach 1 as $n \to \infty$.

Let us translate from the harmonic oscillator into a shifted lattice point counting problem. The 1-dimensional oscillator equation $-u'' + \frac{1}{4}x^2u = Eu$ has eigenvalues $E_j = j - 1/2$ for $j = 1, 2, 3, \dots$. By separating variables and rescaling, one finds equation (18.1) has spectrum

$$\{E_n\} = \{s(j - 1/2) + s^{-1}(k - 1/2) : j, k = 1, 2, 3, \dots\}.$$

Hence the number of eigenvalues less than or equal to r is the number of points in the shifted lattice $(\mathbb{N} - 1/2) \times (\mathbb{N} - 1/2)$ lying below the straight line $sx + s^{-1}y = r$, which is precisely the definition of our counting function $N(r, s)$ when Γ is the straight line $y = 1 - x$ (the 1-ellipse) and the shift parameters are $\sigma = \tau = -1/2$.

The numerical evidence in the left part of Figure 17.2 does not suggest that the s-values maximizing the counting function $N(r, s)$ will converge to 1 as $r \to \infty$. Rather, the optimal s-values seem to cluster at various heights. Thus Antunes and Freitas's theorem for Dirichlet rectangles does not seem to carry over to harmonic oscillators.

Future directions

One could investigate the Schrödinger equation with potential $|sx|^q + |s^{-1}y|^q$, where $2 < q < \infty$. The endpoint case $q = 2$ gives the harmonic oscillator treated above, while $q = \infty$ gives an infinite potential well corresponding to the Dirichlet Laplacian on a rectangular domain. We conjecture that for each q, the s-values maximizing the eigenvalue counting function will converge to 1 as $r \to \infty$. The conjecture would provide a 1-parameter family of quantum oscillators for which the analogue of the Antunes–Freitas theorem holds true.

The difficulty is that the eigenvalues of the 1-dimensional oscillator with potential $|x|^q$ do not grow at a precisely regular rate. Hence to tackle the conjecture, one will need to extend the current thesis from shifted lattices, where each row and column of the lattice is translated by the same amount, and find a way to handle *deformed* lattices, where the amount of translation varies with the rows and columns.

APPENDIX A

THE VAN DER CORPUT TYPE THEOREM

A theorem due to van der Corput [10, Satz 5] is central to the proofs of Proposition 4.1 and Proposition 4.2. We state it as Theorem A.8 at the end of this Appendix. The theorem will be proved in the remainder of the Appendix, by following closely the steps of Krätzel [17] and fixing a gap in his proof (see the comments after Corollary A.7).

We start with a simple exponential sum estimate.

Theorem A.1 (Krätzel [17, Satz 1.1])**.** *Let* $a, b \in \mathbb{Z}$, $a < b$, $\alpha \in \mathbb{R}$. *Then*

$$\Big|\sum_{a<n\le b} e^{2\pi i\alpha n}\Big| \le \min\big(b-a, \frac{1}{2\|\alpha\|}\big), \tag{A.1}$$

where $\|x\| = \min(x - \lfloor x\rfloor, 1-(x-\lfloor x\rfloor))$.

Proof. When α is an integer, the inequality holds trivially. Let us assume α is not an integer. Then

$$\Big|\sum_{a<n\le b} e^{2\pi i\alpha n}\Big| = \Big|\frac{e^{2\pi i\alpha b}-e^{2\pi i\alpha a}}{e^{2\pi i\alpha}-1}\Big| \le \frac{2}{|e^{\pi i\alpha}-e^{-\pi i\alpha}|} = \frac{1}{|\sin\pi\alpha|} \le \frac{1}{2\|\alpha\|}.$$

This proves the theorem. □

Now we consider the general exponential sum where h is a real function. We obtain a generalization of the inequality (A.1), which is called the Kusmin–Landau inequality.

Corollary A.2 (Krätzel [17, Korollar zu Satz 1.2])**.** *Assume that* $h \in C^1[a,b]$, $a, b \in \mathbb{R}$, *and* $h'(t)$ *is monotonic. If* $N \in \mathbb{N}$ *and* $0 < \varphi, \vartheta < 1$ *satisfy* $N + \vartheta \le h'(t) \le N + 1 - \varphi$ *for all* t, *then*

$$\Big|\sum_{a<n\le b} e^{2\pi i h(n)}\Big| \le \frac{1}{2}\big(\cot\frac{\pi}{2}\vartheta + \cot\frac{\pi}{2}\varphi\big). \tag{A.2}$$

Proof. Krätzel gives the proof for [17, Satz 1.2], which we adapt to obtain the following proof.

Case I: $(a, b]$ does not contain an integer. The exponential sum is 0 while the right side is nonnegative.

Case II: $(a, b]$ contains one integer. We have $1 = \cot \frac{\pi}{4} < \frac{1}{2}(\cot \frac{\pi}{2}\vartheta + \cot \frac{\pi}{2}\varphi)$ by convexity, since $\vartheta + \varphi < 1$.

Case III: $(a, b]$ contains more than one integer. If we replace $h'(t)$ by $h'(t) + N$ and $h(t)$ by $h(t) + tN$, then the exponential sum value does not change. This means we could take $N = 0$ without loss of generality. Further, when $h'(t)$ is monotone decreasing, $1 - h'(t) = \big(t - h(t)\big)'$ is monotone increasing, and $0 < \varphi \le 1 - h'(t) \le 1 - \vartheta < 1$. Therefore we can also assume without loss of generality that $h'(t)$ is monotone increasing. Thus h is convex and increasing, with $\vartheta \le h'(t) \le 1 - \varphi$. We write

$$\begin{aligned}
S &= \sum_{a<n\le b} e^{2\pi i h(n)} \\
&= e^{2\pi i h(\lfloor a\rfloor+1)} + \sum_{n=\lfloor a\rfloor+2}^{\lfloor b\rfloor} \frac{e^{2\pi i h(n)}}{e^{2\pi i h(n)} - e^{2\pi i h(n-1)}}\big(e^{2\pi i h(n)} - e^{2\pi i h(n-1)}\big) \\
&= e^{2\pi i h(\lfloor a\rfloor+1)} + \frac{1}{2}\sum_{n=\lfloor a\rfloor+2}^{\lfloor b\rfloor} \big(1 - i\cot\pi \int_{n-1}^{n} h'(t)\,\mathrm{d}t\big)\big(e^{2\pi i h(n)} - e^{2\pi i h(n-1)}\big) \\
&= \frac{1}{2} e^{2\pi i h(\lfloor a\rfloor+1)}\big(1 + i\cot\pi\int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t\big) \\
&\qquad + \frac{1}{2} e^{2\pi i h(\lfloor b\rfloor)}\big(1 - i\cot\pi\int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t\big) \\
&\qquad - \frac{i}{2}\sum_{n=\lfloor a\rfloor+2}^{\lfloor b\rfloor-1} e^{2\pi i h(n)}\big(\cot\pi\int_{n-1}^{n} h'(t)\,\mathrm{d}t - \cot\pi\int_{n}^{n+1} h'(t)\,\mathrm{d}t\big).
\end{aligned}$$

Therefore

$$\begin{aligned}
2|S| \le{}& |1 + i\cot\pi\int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t| + |1 - i\cot\pi\int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t| \\
&+ \sum_{n=\lfloor a\rfloor+2}^{\lfloor b\rfloor-1} \big|\cot\pi\int_{n-1}^{n} h'(t)\,\mathrm{d}t - \cot\pi\int_{n}^{n+1} h'(t)\,\mathrm{d}t\big|.
\end{aligned}$$

Since $h'(t)$ is monotone increasing, $n \mapsto \cot\pi\int_{n-1}^{n} h'(t)\,\mathrm{d}t$ is monotone decreasing, and so by telescoping we have

$$\begin{aligned}
2|S| \le &\frac{1}{\sin \pi \int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t} + \frac{1}{\sin \pi \int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t} \\
&\quad + \cot \pi \int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t - \cot \pi \int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t \\
&= \frac{1 + \cos \pi \int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t}{\sin \pi \int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t} + \frac{1 - \cos \pi \int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t}{\sin \pi \int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t} \\
&= \cot \frac{\pi}{2} \int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t + \cot \frac{\pi}{2}\big(1 - \int_{\lfloor b\rfloor-1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t\big) \\
&\le \cot \frac{\pi}{2}\vartheta + \cot \frac{\pi}{2}\varphi.
\end{aligned}$$

This is inequality (A.2). □

Lemma A.3 (Krätzel [17, Hilfssatz 1.1])**.** *Let $a < b$ be real numbers. Suppose h is a real function with $h''(t) > 0$ for $t \in [a,b]$ (or $h''(t) < 0$ for $t \in [a,b]$). Further suppose $N \in \mathbb{Z}$ and $0 < \lambda < 1/2$ with*

$$N \le h'(t) \le N+1, \quad |h''(t)| \ge \lambda \qquad \textit{for } t \in [a,b].$$

Then

$$\Big|\sum_{a<n\le b} e^{2\pi i h(n)}\Big| \le \frac{5}{\sqrt{\lambda}}. \tag{A.3}$$

Proof. Case I: $(a,b]$ contains at most two integers. This case is trivial since

$$\Big|\sum_{a<n\le b} e^{2\pi i h(n)}\Big| \le 2 < 5/\sqrt{\lambda}.$$

Case II: $(a,b]$ contains more than two integers, so $b-a>2$. Without loss of generality we assume $h''(t) \ge \lambda > 0$, so that $h'(t)$ is monotone increasing. Let $0 < \delta < 1/2$ be a suitable number which is defined later. We decompose the sum according to the value of $h'(t)$. Let $a_1, b_1 \in \mathbb{Z}$ be determined by δ as follows. Let

$$
\begin{aligned}
M_1 &= \{\lfloor a\rfloor+1,\dots,a_1\} = \{n\in(a,b-1] : N \le \int_n^{n+1} h'(t)\,\mathrm{d}t < N+\delta\},\\
M_2 &= \{a_1+1,\dots,b_1\} = \{n\in(a,b-1] : N+\delta \le \int_n^{n+1} h'(t)\,\mathrm{d}t \le N+1-\delta\},\\
M_3 &= \{b_1+1,\dots,\lfloor b\rfloor-1\}\\
&= \{n\in(a,b-1] : N+1-\delta < \int_n^{n+1} h'(t)\,\mathrm{d}t \le N+1\}.
\end{aligned}
$$

We denote $S_i = \sum_{n\in M_i} e^{2\pi i h(n)}$, for $i=1,2,3$. Then

$$S = S_1+S_2+S_3+e^{2\pi i h(\lfloor b\rfloor)}.$$

For the estimation of S_2, we use Corollary A.2 with $0<\vartheta=\varphi=\delta<1/2$. We have

$$|S_2| \le \cot\frac{\pi\delta}{2} \le \frac{1}{\delta}.$$

For S_1, we use trivial estimation as follows. If M_1 has two or more elements, then

$$
\begin{aligned}
\delta &> \int_{a_1}^{a_1+1} h'(t)\,\mathrm{d}t - N \ge \int_{a_1}^{a_1+1} h'(t)\,\mathrm{d}t - \int_{\lfloor a\rfloor+1}^{\lfloor a\rfloor+2} h'(t)\,\mathrm{d}t\\
&= \sum_{n=\lfloor a\rfloor+2}^{a_1}\big(\int_n^{n+1} h'(t)\,\mathrm{d}t - \int_{n-1}^{n} h'(t)\,\mathrm{d}t\big) = \sum_{n=\lfloor a\rfloor+2}^{a_1}\int_n^{n+1}\int_{t-1}^{t} h''(x)\,\mathrm{d}x\,\mathrm{d}t\\
&\ge \lambda(a_1-\lfloor a\rfloor-1).
\end{aligned}
$$

So

$$|S_1| \le a_1-\lfloor a\rfloor \le \frac{\delta}{\lambda}+1.$$

If M_1 has at most one element, then $|S_1|\le 1\le \delta/\lambda+1$.

We also estimate S_3 trivially. Suppose there are at least two elements in

M_3. Then

$$\begin{aligned}\delta > N+1-\int_{b_1+1}^{b_1+2} h'(t)\,\mathrm{d}t \geq \int_{\lfloor b\rfloor -1}^{\lfloor b\rfloor} h'(t)\,\mathrm{d}t - \int_{b_1+1}^{b_1+2} h'(t)\,\mathrm{d}t \\ = \sum_{n=b_1+2}^{\lfloor b\rfloor -1}\int_n^{n+1}\int_{t-1}^{t} h''(x)\,\mathrm{d}x\,\mathrm{d}t \\ \geq (\lfloor b\rfloor - b_1 - 2)\lambda,\end{aligned}$$

and therefore

$$|S_3| \leq \lfloor b\rfloor - b_1 - 1 \leq \frac{\delta}{\lambda}+1.$$

If there is at most one element in M_3, then $|S_3| \leq 1 \leq \delta/\lambda + 1$.

Combining the above estimates, we have

$$|S| \leq \frac{1}{\delta} + \frac{2\delta}{\lambda} + 3.$$

Choose $\delta = \sqrt{\lambda/2}$, the condition $\delta < 1/2$ is obtained because $\lambda < 1/2$. Since $\sqrt{2} < 1/\sqrt{\lambda}$, we have

$$|S| < (2+\frac{3}{2})\sqrt{\frac{2}{\lambda}} < \frac{5}{\sqrt{\lambda}},$$

which is (A.3). ☐

Corollary A.4 (Krätzel [17, Korollar zu Satz 1.3])**.** *Let $a < b$ be real numbers such that the interval $(a, b]$ contains more than one integer. If $h \in C^2[a, b]$ with $|h''(t)| \geq \lambda > 0$, then*

$$\Big|\sum_{a<n\leq b} e^{2\pi i h(n)}\Big| \leq |h'(b) - h'(a)|\frac{5}{\sqrt{\lambda}} + \frac{11}{\sqrt{\lambda}}. \tag{A.4}$$

Proof. Without loss of generality we assume $h''(t) \geq \lambda > 0$ so that $h'(t)$ is monotone increasing. We first suppose $\lambda < 1/2$. Write

$$|S| = \Big|\sum_{a<n\leq b} e^{2\pi i h(n)}\Big| \leq \Big|\sum_{n=\lfloor a\rfloor+1}^{\lfloor b\rfloor -1} e^{2\pi i h(n)}\Big| + 1,$$

and decompose the interval $\big[\lfloor a\rfloor + 1, \lfloor b\rfloor - 1\big]$ into consecutive subintervals

$I(a), I_k, I(b)$, as follows. The subinterval $I(a)$ is determined by the inequality

$$h'(\lfloor a \rfloor + 1) \leq h'(n) \leq \lfloor h'(\lfloor a \rfloor + 1) \rfloor + 1.$$

The subinterval I_k is determined by

$$\lfloor h'(\lfloor a \rfloor + 1) \rfloor + k < h'(n) \leq \lfloor h'(\lfloor a \rfloor + 1) \rfloor + k + 1.$$

$k = 1, 2, \ldots, \lfloor h'(\lfloor b \rfloor - 1) \rfloor - \lfloor h'(\lfloor a \rfloor + 1) \rfloor - 1$, and finally $I(b)$ is determined by

$$\lfloor h'(\lfloor b \rfloor - 1) \rfloor < h'(n) \leq h'(\lfloor b \rfloor - 1).$$

In each subinterval, we apply Lemma A.3. The number of subintervals is

$$\lfloor h'(\lfloor b \rfloor - 1) \rfloor - \lfloor h'(\lfloor a \rfloor + 1) \rfloor + 1 \leq h'(b) - h'(a) + 2,$$

and so using (A.3) we get

$$|S| \leq (h'(b) - h'(a) + 2) \frac{5}{\sqrt{\lambda}} + 1.$$

Since $\lambda < 1/2$, we obtain (A.4).

For $\lambda \geq 1/2$ we estimate trivially. Because there is more than one integer in $(a, b]$, we have $b - a > 1$. Then

$$\begin{aligned} |S| &\leq \lfloor b \rfloor - \lfloor a \rfloor < b - a + 1 < 2(b - a) \\ &< \frac{2\sqrt{2}}{\sqrt{\lambda}} \int_a^b h''(t)\, \mathrm{d}t \\ &< \frac{5}{\sqrt{\lambda}} \big(h'(b) - h'(a) \big). \end{aligned}$$

Hence (A.4) is true again.

□

Lemma A.5 (Krätzel [17, Hilfssatz 1.2])**.** *The Fourier series*

$$-\frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin 2\pi n t \tag{A.5}$$

converges uniformly to the sawtooth function $\psi(t) = t - \lfloor t \rfloor - 1/2$ *on every* *closed interval not containing an integer. The partial sums*

$$s_N(t) = -\frac{1}{\pi}\sum_{n=1}^{N}\frac{1}{n}\sin 2\pi nt$$

are uniformly bounded.

Proof. The sawtooth function is smooth except at the integers, and so its Fourier series converges uniformly on every closed interval not containing an integer. For the partial sums, since sin is periodic and odd it suffices to show $s_N(t)$ is uniformly bounded for $0 \le t \le 1/2$. Using the analysis of the Gibbs phenomenon from [26, Ch. II (9.2)] we have

$$s_N(t) = -\frac{1}{\pi}\int_0^{2\pi Nt}\frac{\sin s}{s}\,\mathrm{d}s + t + o(1)$$

when $0 \le t \le 1/2$. Hence the partial sums are uniformly bounded, since $\int_0^\tau (\sin s)/s\,\mathrm{d}s$ is continuous as a function of τ and tends to a limit as $\tau \to \infty$. □

Lemma A.6 (Krätzel [17, Hilfssatz 1.3])**.** *Let h denote a real valued function and let a, b, z be real numbers with $a < b$ and $z > 1$. Then*

$$\Big|\sum_{a<n\le b}\psi\big(h(n)\big)\Big| \le \frac{b-a+1}{2\pi z} + \frac{1}{\pi}\sum_{\nu=1}^{\infty}\min\big(\frac{1}{\nu}, \frac{z}{\nu^2}\big)\Big|\sum_{a<n\le b} e^{2\pi i\nu h(n)}\Big| \tag{A.6}$$

Proof. The slope of ψ is 1 (except where ψ jumps downward) and so

$$\psi(x) \le \psi(x-y) + y$$

for $y > 0$. Hence

$$\psi(x) = \pi z\int_0^{1/\pi z}\psi(x)\,\mathrm{d}y \le \pi z\int_0^{1/\pi z}\psi(x-y)\,\mathrm{d}y + \frac{1}{2\pi z}.$$

Using the Fourier series (A.5) for ψ along with uniform boundedness of the partial sums, we obtain

$$\psi(x) \le \frac{1}{2\pi z} - \frac{1}{2\pi i}\sum_{\nu=1}^{\infty}(c_\nu e^{2\pi i\nu x} + c_{-\nu}e^{-2\pi i\nu x}) \tag{A.7}$$

where

$$c_\mu = \frac{\pi z}{\mu} \int_0^{1/\pi z} e^{-2\pi i \mu y}\, \mathrm{d}y = \frac{z}{\mu^2} e^{-i\mu/z} \sin \frac{\mu}{z}.$$

Analogously,

$$\psi(x) \geq \psi(x+y) - y$$

for $y \geq 0$, and hence

$$\psi(x) \geq -\frac{1}{2\pi z} + \frac{1}{2\pi i} \sum_{\nu=1}^{\infty} (c_\nu e^{-2\pi i \nu x} + c_{-\nu} e^{2\pi i \nu x}). \tag{A.8}$$

Now we set $x = h(n)$ in (A.7) and (A.8) and sum over n. The estimate

$$|c_\mu| \leq \min(\frac{1}{|\mu|}, \frac{z}{\mu^2})$$

leads to (A.6). □

Corollary A.7 (Krätzel [17, Korollar zu Satz 1.4])**.** *Let $a < b$ be real numbers. If $h \in C^2[a,b]$ satisfies $|h''(t)| \geq \lambda > 0$ with $\lambda \leq 121$, then*

$$\Big| \sum_{a<n\leq b} \psi\big(h(n)\big) \Big| < \frac{11}{2} |h'(b) - h'(a)| \lambda^{-2/3} + \frac{11}{\sqrt{\lambda}}. \tag{A.9}$$

Krätzel did not assume $\lambda \leq 121$, but an upper bound on λ is needed to handle Case I in the proof, when a and b are very close.

Proof. Case I: There is at most one integer in $(a, b]$. Then (A.9) holds trivially since $|\psi| \leq 1/2 < 11/\sqrt{\lambda}$.

Case II: There is more than one integer in $(a, b]$. We obtain from (A.4) in Corollary A.4 that

$$\Big| \sum_{a<n\leq b} e^{2\pi i \nu h(n)} \Big| \leq |h'(b) - h'(a)| 5 \sqrt{\frac{\nu}{\lambda}} + \frac{11}{\sqrt{\nu\lambda}}, \qquad \nu > 0.$$

Hence from (A.6) in Lemma A.6, for each $z > 1$ we have

$$\begin{aligned}
&\Big| \sum_{a<n\le b} \psi\big(h(n)\big)\Big| \\
&\le \frac{b-a+1}{2\pi z} + \frac{1}{\pi}\sum_{\nu=1}^{\infty} \min\big(\frac{1}{\nu}, \frac{z}{\nu^2}\big)\Big(|h'(b)-h'(a)|5\sqrt{\frac{\nu}{\lambda}} + \frac{11}{\sqrt{\nu\lambda}}\Big) \\
&\le \frac{b-a+1}{2\pi z} + |h'(b)-h'(a)|\frac{5}{\pi\sqrt{\lambda}}\Big(\sum_{1\le\nu\le z}\frac{1}{\sqrt{\nu}} + \sum_{\nu>z}\frac{z}{\nu^{3/2}}\Big) + \frac{11}{\pi\sqrt{\lambda}}\sum_{\nu=1}^{\infty}\frac{1}{\nu^{3/2}} \\
&< \frac{b-a+1}{2\pi z} + |h'(b)-h'(a)|\frac{5}{\pi\sqrt{\lambda}}\Big(\int_0^z \frac{1}{\sqrt{t}}\,\mathrm{d}t + \int_{\lfloor z\rfloor}^{\infty}\frac{z}{t^{3/2}}\,\mathrm{d}t\Big) \\
&\quad + \frac{11}{\pi\sqrt{\lambda}}\Big(1 + \int_1^{\infty}\frac{1}{t^{3/2}}\,\mathrm{d}t\Big) \\
&\le \frac{b-a+1}{2\pi z} + |h'(b)-h'(a)|\frac{25}{\pi}\sqrt{\frac{z}{\lambda}} + \frac{33}{\pi\sqrt{\lambda}}
\end{aligned}$$

where we used that $\lfloor z\rfloor \ge z/2$ when $z > 1$. Since h'' does not change sign and $b - a > 1$ in Case II, we have

$$|h'(b) - h'(a)| = |\int_a^b h''(t)\,\mathrm{d}t| \ge \lambda(b-a) \ge \lambda\frac{(b-a+1)}{2},$$

and so

$$\Big| \sum_{a<n\le b} \psi\big(h(n)\big)\Big| < |h'(b)-h'(a)|\Big(\frac{1}{\pi z\lambda} + \frac{25}{\pi}\sqrt{\frac{z}{\lambda}}\Big) + \frac{33}{\pi\sqrt{\lambda}}.$$

We will further bound the above inequality by separating into two cases: $\lambda < 1/5^3$ and $\lambda \ge 1/5^3$. First suppose $\lambda < 1/5^3$, and choose

$$z = \frac{1}{5}\lambda^{-1/3} > 1.$$

Then (A.9) holds since

$$\begin{aligned}
\Big| \sum_{a<n\le b} \psi\big(h(n)\big)\Big| &< |h'(b)-h'(a)|\Big(\frac{5}{\pi} + \frac{25}{\pi}\sqrt{\frac{1}{5}}\Big)\lambda^{-2/3} + \frac{33}{\pi\sqrt{\lambda}} \\
&< \frac{11}{2}|h'(b)-h'(a)|\lambda^{-2/3} + \frac{11}{\sqrt{\lambda}}.
\end{aligned}$$

Next suppose $\lambda \geq 1/5^3$. Using the trivial bound $|\psi(\cdot)| \leq 1/2$, we have

$$\begin{aligned}
\Big|\sum_{a<n\leq b} \psi\big(h(n)\big)\Big| &\leq \frac{1}{2}(\lfloor b\rfloor - \lfloor a\rfloor) \\
&\leq b - a \qquad \text{since } \lfloor b\rfloor - \lfloor a\rfloor \geq 2 \text{ in Case II} \\
&\leq 5\lambda^{-2/3}\lambda(b-a) \\
&\leq 5\lambda^{-2/3}\Big|\int_a^b h''(t)\,\mathrm{d}t\Big| \\
&< \frac{11}{2}|h'(b) - h'(a)|\lambda^{-2/3}.
\end{aligned}$$

Thus (A.9) holds.

□

Theorem A.8 (Krätzel [17, Korollar zu Satz 1.5])**.** *Suppose $a < b$ are real numbers and $h \in C^2[a,b]$ with h'' monotonic and nonzero. Then*

$$\Big|\sum_{a<n\leq b} \psi\big(h(n)\big)\Big| \leq 6\int_a^b |h''(t)|^{1/3}\,\mathrm{d}t + 175\max_{[a,b]}\frac{1}{|h''|^{1/2}} + 1. \tag{A.10}$$

Krätzel's result has "+2" as the final term. We get "+1" by correcting a gap in the proof coming from the additional assumption $\lambda \leq 121$ needed in Corollary A.7, and by arguing more carefully in Case I below.

Proof. Without loss of generality we assume the interval $(a,b]$ contains at least one integer and $|h''(t)|$ is increasing. Let $c = (12/11)^{3/2} > 1$.

Case I: Assume that $|h''(b)| \leq 121$.

We decompose the interval $(a,b]$ into subintervals $(n_\nu, n_{\nu+1}]$ $(\nu = 0, 1, \ldots, N-1)$ and $(n_N, b]$, where $n_0 = a$, such that

$$c^\nu|h''(\lfloor a\rfloor + 1)| \leq |h''(n)| < c^{\nu+1}|h''(\lfloor a\rfloor + 1)| \leq |h''(\lfloor b\rfloor)| \leq 121 \tag{A.11}$$

for integers n with $n_\nu < n \leq n_{\nu+1}$. Applying (A.9) in Corollary A.7 to each subinterval, we get

$$\Big|\sum_{a<n\le b}\psi\big(h(n)\big)\Big| = \Big|\sum_{\nu=0}^{N-1}\sum_{n_\nu<n\le n_{\nu+1}}\psi(h(n)) + \sum_{n_N<n\le b}\psi(h(n))\Big|$$
$$< \frac{11}{2}\sum_{\nu=0}^{N-1}\frac{|h'(n_{\nu+1})-h'(n_\nu)|}{(c^\nu|h''(\lfloor a\rfloor+1)|)^{2/3}} + \frac{11}{2}\frac{|h'(b)-h'(n_N)|}{(c^N|h''(\lfloor a\rfloor+1)|)^{2/3}}$$
$$+\sum_{\nu=0}^{N}\frac{11}{\sqrt{c^\nu|h''(\lfloor a\rfloor+1)|}}.$$

Because of (A.11) we get

$$\Big|\sum_{a<n\le b}\psi\big(h(n)\big)\Big| < \frac{11}{2}\sum_{\nu=0}^{N-1}(c^\nu|h''(\lfloor a\rfloor+1)|)^{-2/3}\int_{n_\nu}^{n_{\nu+1}}|h''(t)|\,\mathrm{d}t$$
$$+\frac{11}{2}(c^N|h''(\lfloor a\rfloor+1)|)^{-2/3}\int_{n_N}^{b}|h''(t)|\,\mathrm{d}t$$
$$+\sum_{\nu=0}^{\infty}\frac{11}{\sqrt{c^\nu|h''(\lfloor a\rfloor+1)|}}$$
$$<\frac{11}{2}c^{2/3}\sum_{\nu=0}^{N-1}\int_{n_\nu}^{n_{\nu+1}}|h''(t)|^{1/3}\,\mathrm{d}t + \frac{11}{2}c^{2/3}\int_{n_N}^{b}|h''(t)|^{1/3}\,\mathrm{d}t$$
$$+\frac{\sqrt{c}}{\sqrt{c}-1}\frac{11}{\sqrt{|h''(\lfloor a\rfloor+1)|}}.$$

Now (A.10) follows.

Case II: Assume that $|h''(b)|>121$. Let $\widetilde{b}\in[a,b]$ such that $|h''(\widetilde{b})|=121$, if such a point exists, and otherwise let $\widetilde{b}=a$. Recall $|h''|$ is increasing. Then on $[a,\widetilde{b}]$, we use the previous proof to obtain

$$\Big|\sum_{a<n\le\widetilde{b}}\psi\big(h(n)\big)\Big| \le 6\int_a^{\widetilde{b}}|h''(t)|^{1/3}\,\mathrm{d}t + 175\max_{[a,\widetilde{b}]}\frac{1}{|h''|^{1/2}}.$$

And on $[\widetilde{b},b]$ we use the lower bound $|h''|\ge 121$ to obtain

$$\Big|\sum_{\widetilde{b}<n\le b}\psi\big(h(n)\big)\Big| \le \frac{1}{2}(b-\widetilde{b}+1)\le 6\int_{\widetilde{b}}^{b}|h''(t)|^{1/3}\,\mathrm{d}t+1.$$

Combining the two inequalities, we have

$$\Big|\sum_{a<n\le b}\psi\big(h(n)\big)\Big| \le 6\int_a^b |h''(t)|^{1/3}\,\mathrm{d}t + 175\max_{[a,b]}\frac{1}{|h''|^{1/2}} + 1.$$

□

APPENDIX B

CODE FOR *P*-ELLIPSE LATTICE POINT COUNTING

In this appendix, we will provide the C++ code for numerical calculation of the optimal stretch parameter. The algorithm was described in Chapter 6 for $p = 1$ with no shift. The code adapts to other values of p and also to shifted lattice counting. This code is used to generate the data for Figure 1.2, Figure 6.2, Figure 17.1 and Figure 17.2.

Lower bound for optimal stretch

```
#include<math.h>
#include <cmath>
#include<algorithm>
#include<iostream>
#include<vector>
#include<fstream>
#include<sstream>
using namespace std;

string find_opt(double p, double r, double sigma, double tau) {
  /*
  Function find_opt finds the smallest s-value that maximizes the
     counting function. Parameter "p" means the curve is a
     p-ellipse. Scale parameter is "r". Shift parameters are
     "sigma" and "tau".
  */
  if(sigma <= -1 || tau <= -1) {
    cout<<"Error: shift parameters less than -1."<<endl;
    return "shift parameters less than -1.";
  }
```

```
double max_lattice = r*r*1.0/pow(4, 1.0/p);
if(max_lattice < (1+sigma)*(1+tau)) {
   cout<< "Error: radius too small, no lattice points
       included."<<endl;
   return "radius too small.";
}
vector<double> min_arr;
vector<double> max_arr;

for(int j=1; j+sigma<= max_lattice/(1+tau); j=j+1) {
   for(int k=1; k+tau<=max_lattice/(j+sigma); k=k+1) {
      double delta =
          sqrt(pow(r,(2*p))-4*pow((j+sigma)*(k+tau),p));
      min_arr.push_back(pow((pow(r,p)-delta)*1.0/2/pow(j+sigma,p),
          1.0/p));
      max_arr.push_back(pow((pow(r,p)+delta)*1.0/2/pow(j+sigma,p),
          1.0/p));
   }
}
sort(min_arr.begin(), min_arr.end());
sort(max_arr.begin(), max_arr.end());
for(int i=0; i< max_arr.size(); i++) {
   int j = max_arr.size()-i-1;
}
int j = 0;
int k = 0;
long count = 0;
long max_num = 0;
double s = -1;
double val = -1;

while(j < min_arr.size() && k < max_arr.size()) {
   if(min_arr[j] <= max_arr[k]) {
      count += 1;
      val = min_arr[j];
      j += 1;
   }
   else {
```

```
            count -= 1;
            val = max_arr[k];
            k += 1;
        }
        // check whether we need to update max_num
        if(count > max_num) {
            max_num = count;
            s = val;
        }
    }
    return to_string(r) + "," + to_string(s);
}

int main()
{
    ofstream myfile;
    double sigma = 0;
    double tau = 0;
    double p = 2;
    double step = 3;
    myfile.open(to_string(sigma) + "_" + to_string(tau) + "_" +
        to_string(int(p)) + "_" + ".txt");

    for(double i = sqrt(pow(4*(1+sigma)*(1+tau),p))+1; i<1000;
        i=i+0.1*sqrt(step)) {
        vector<string> elems;
        stringstream ss(find_opt(p, i, sigma, tau));
            string item;
            while (getline(ss, item, ',')) {
                elems.push_back(item);
            }
        myfile<<i<<" "<< elems[1]<<endl;
    }
    myfile.close();
}
```

```
#include<math.h>
#include <cmath>
#include<algorithm>
#include<iostream>
#include<vector>
#include<fstream>
#include<sstream>
using namespace std;

string find_opt_max(double p, double r, double sigma, double tau) {
   /*
   Function find_opt_max finds the largest s-value that maximizes
       the counting function. Parameter "p" means the curve is a
       p-ellipse. Scale parameter is "r". Shift parameters are
       "sigma" and "tau".
   */
   if(sigma <= -1 || tau <= -1) {
      cout<<"Error: shift parameters less than -1."<<endl;
      return "shift parameters less than -1.";
   }
   double max_lattice = r*r*1.0/pow(4,1.0/p);
   if(max_lattice < (1+sigma)*(1+tau)) {
      cout<< "Error: radius too small, no lattice points
          included."<<endl;
      return "radius too small.";
   }
   vector<double> min_arr;
   vector<double> max_arr;

   for(int j=1; j+sigma<= max_lattice/(1+tau); j=j+1) {
      for(int k=1; k+tau<=max_lattice/(j+sigma); k=k+1) {
         double delta =
             sqrt(pow(r,(2*p))-4*pow((j+sigma)*(k+tau),p));
         min_arr.push_back(pow((pow(r,p)-delta)*1.0/2/pow(j+sigma,p),
             1.0/p));
```

```
        max_arr.push_back(pow((pow(r,p)+delta)*1.0/2/pow(j+sigma,p),
           1.0/p));
      }
   }
   sort(min_arr.begin(), min_arr.end());
   sort(max_arr.begin(), max_arr.end());
   int j = 0;
   int k = 0;
   long count = 0;
   long max_num = 0;
   double s = -1;
   double val = -1;

   while(j < min_arr.size() && k < max_arr.size()) {
      if(min_arr[j] <= max_arr[k]) {
         count += 1;
         val = min_arr[j];
         j += 1;
         max_num = max(max_num, count);
      }
      else {
         if(count == max_num) {
            s = max_arr[k];
         }
         count -= 1;
         val = max_arr[k];
         k += 1;
      }
   }
   if(count == max_num) {
      s = max_arr[k];
   }
   return to_string(r) + "," + to_string(s);
}

int main()
{
   ofstream myfile;
```

```
  double sigma = 0 ;
  double tau = 0;
  double p = 2;
  double step = 3;
  myfile.open("max"+to_string(sigma) + "_" + to_string(tau) + "_"
     + to_string(int(p)) + ".txt");

  for(double i=sqrt(pow(4*(1+sigma)*(1+tau),p))+1; i<1000;
     i=i+0.1*sqrt(step)) {
    vector<string> elems;
    stringstream ss(find_opt_max(p, i, sigma, tau));
      string item;
      while (getline(ss, item, ',')) {
      elems.push_back(item);
      }
    myfile<<i<<" "<< elems[1]<<endl;
  }
  myfile.close();
}
```

REFERENCES

[1] P. R. S. Antunes and P. Freitas. *Numerical optimisation of low eigenvalues of the Dirichlet and Neumann Laplacians.* J. Optim. Theory Appl. 154 (2012), 235–257.

[2] P. R. S. Antunes and P. Freitas. *Optimal spectral rectangles and lattice ellipses.* Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 469 (2013), no. 2150, 20120492, 15 pp.

[3] P. R. S. Antunes and P. Freitas. *Optimisation of eigenvalues of the Dirichlet Laplacian with a surface area restriction.* Appl. Math. Optim. 73 (2016), no. 2, 313–328.

[4] P. R. S. Antunes and É. Oudet. *Numerical minimization of Dirichlet–Laplacian eigenvalues of four-dimensional geometries.* SIAM J. Sci. Comput., to appear.

[5] S. Ariturk and R. S. Laugesen. *Optimal stretching for lattice points under convex curves.* Port. Math., to appear. ArXiv:1701.03217

[6] M. van den Berg, D. Bucur and K. Gittins. *Maximizing Neumann eigenvalues on rectangles.* Bull. Lond. Math. Soc. 48 (2016), no. 5, 877–894.

[7] M. van den Berg and K. Gittins. *Minimising Dirichlet eigenvalues on cuboids of unit measure.* Mathematika 63 (2017), no. 2, 469–482.

[8] D. Bucur and P. Freitas. *Asymptotic behaviour of optimal spectral planar domains with fixed perimeter.* J. Math. Phys. 54 (2013), no. 5, 053504.

[9] B. Colbois and A. El Soufi. *Extremal eigenvalues of the Laplacian on Euclidean domains and closed surfaces.* Math. Z. 278 (2014), 529–549.

[10] J. G. van der Corput. *Zahlentheoretische Abschätzungen mit Anwendung auf Gitterpunktprobleme.* Math. Z. 17 (1923), 250–259.

[11] P. Freitas. *Asymptotic behaviour of extremal averages of Laplacian eigenvalues.* J. Stat. Phys. 167 (2017), no. 6, 1511–1518.

[12] K. Gittins and S. Larson. *Asymptotic behaviour of cuboids optimising Laplacian eigenvalues.* Preprint. ArXiv:1703.10249.

[13] A. Henrot, ed. Shape Optimization and Spectral Theory. De Gruyter Open, 2017.

[14] M. N. Huxley. Area, Lattice Points, and Exponential Sums. London Mathematical Society Monographs. New Series, vol. 13, The Clarendon Press, Oxford University Press, New York, 1996.

[15] M. N. Huxley. *Exponential sums and lattice points. III.* Proc. London Math. Soc. (3) 87 (2003), 591–609.

[16] A. Ivić, E. Krätzel, M. Kühleitner and W. G. Nowak. *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, 89–128, Franz Steiner Verlag, Stuttgart, 2006.

[17] E. Krätzel. Analytische Funktionen in der Zahlentheorie. Teubner–Texte zur Mathematik, 139. B. G. Teubner, Stuttgart, 2000.

[18] E. Krätzel. *Lattice points in planar convex domains.* Monatsh. Math. 143 (2004), 145–162.

[19] S. Larson. *Asymptotic shape optimization for Riesz means of the Dirichlet Laplacian over convex domains.* ArXiv:1611.05680

[20] R. S. Laugesen and S. Liu. *Optimal stretching for lattice points and eigenvalues.* ArXiv:1609.06172 (Expanded version of paper [21].)

[21] R. S. Laugesen and S. Liu. *Optimal stretching for lattice points and eigenvalues.* Arkiv för Matematik, to appear.

[22] R. S. Laugesen and S. Liu. *Shifted lattices and asymptotically optimal ellipses.* Preprint.

[23] N. F. Marshall. *Stretching convex domains to capture many lattice points.* ArXiv:1707.00682.

[24] N. F. Marshall and S. Steinerberger. *Triangles capturing many lattice points.* ArXiv:1706.04170.

[25] É. Oudet. *Numerical minimization of eigenmodes of a membrane with respect to the domain.* ESAIM Control Optim. Calc. Var. 10 (2004), 315–330.

[26] A. Zygmund. Trigonometric Series. Cambridge University Press, Cambridge, 1959.