NON-CONVEX GEOMETRY OF NUMBERS AND CONTINUED FRACTIONS

NICKOLAS ANDERSEN, WILLIAM DUKE, ZACH HACKING, AND AMY WOODALL

Abstract. In recent work, the first two authors constructed a generalized continued fraction called the p-continued fraction, characterized by the property that its convergents (a subsequence of the regular convergents) are best approximations with respect to the L^p norm, where $p \geq 1$. We extend this construction to the region $0 < p < 1$, where now the L^p quasinorm is non-convex. We prove that the approximation coefficients of the p-continued fraction are bounded above by $\frac{1}{\sqrt{5}} + \varepsilon_p$, where $\varepsilon_p \to 0$ as $p \to 0$. In light of Hurwitz's theorem, this upper bound is sharp, in the limit. We also measure the maximum number of consecutive regular convergents that are skipped by the p-continued fraction.

1. Introduction

A rational number r/s with $\gcd(r, s) = 1$ and $s > 0$ is said to be a best approximation to an irrational real number α if, for all rationals $r'/s' \neq r/s$ with $\gcd(r', s') = 1$ and $0 < s' < s$, we have

$$|r - s\alpha| < |r' - s'\alpha|.$$

Lagrange [4] showed that each irrational α has infinitely many best approximations p_n/q_n with $n = 0, 1, 2, 3, \ldots$ and that for each of them

$$q_n|p_n - q_n\alpha| < 1.$$

The best approximations are given explicitly by the convergents

$$\frac{p_n}{q_n} = b_0 + \frac{1}{b_1+}\,\frac{1}{b_2+}\cdots\frac{1}{b_n}$$

of the regular continued fraction expansion of α:

$$\alpha = b_0 + \frac{1}{b_1+}\,\frac{1}{b_2+}\cdots := b_0 + \cfrac{1}{b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots}}}.$$

Classical results of Vahlen and Borel (Theorems 5A and 5B of [9, Ch. I]) state that among any successive pair of convergents, there is at least one that satisfies $q_n|p_n - q_n\alpha| < \frac{1}{2}$ and among any successive triple, there is at least one that satisfies the Hurwitz bound $q_n|p_n - q_n\alpha| < \frac{1}{\sqrt{5}}$.

The notion of best approximation has the following generalization. For $(x, y) \in \mathbb{R}^2$ and a fixed $0 < p < \infty$, let

$$F^{\langle p \rangle}(x, y) = (|x|^p + |y|^p)^{\frac{1}{p}}, \tag{1.1}$$

The second author is supported by NSF grant DMS 1701638 and Simons Foundation Award 554649.

while $F^{\langle\infty\rangle}(x,y)=\max\{|x|,|y|\}$. If $p\geq 1$ then $F^{\langle p\rangle}$ gives a norm on $\mathbb{R}^2$ while if $0<p<1$ it is a quasinorm in that it only satisfies the weakened triangle inequality

$$F^{\langle p\rangle}(x_1+x_2,y_1+y_2)\leq 2^{\frac{1}{p}-1}(F^{\langle p\rangle}(x_1,y_1)+F^{\langle p\rangle}(x_2,y_2)).$$

For a fixed p, we say that r/s is an L^p-best approximation to α if there is a $t>1$ depending only on r/s so that for any rational $r'/s'\neq r/s$

$$F_t^{\langle p\rangle}(s,r-s\alpha)<F_t^{\langle p\rangle}(s',r'-s'\alpha),$$

where

$$F_t^{\langle p\rangle}(x,y)=F^{\langle p\rangle}(t^{-1}x,ty).$$

It is not difficult to show that r/s is an L^∞-best approximation to α if and only if it is a best approximation to α in Lagrange's sense (see Lemma 6.1 in [1]). Generalizing the case $p=1$, which is due to Minkowski [5, 7], it is shown in [1] that for any fixed $p\geq 1$ the L^p-best approximations to α are those rationals given by the convergents

$$\frac{r_n}{s_n}=a_0+\frac{\varepsilon_1}{a_1+}\frac{\varepsilon_2}{a_2+}\cdots\frac{\varepsilon_n}{a_n}\qquad (n=0,1,2,\ldots \text{ and } \varepsilon_j=\pm 1)$$

of a uniquely determined semi-regular continued fraction expansion of α, the p-continued fraction. Furthermore, each such best approximation satisfies

$$F^{\langle p\rangle}(s_n,r_n-s_n\alpha)<\Delta_p^{-\frac{1}{2}}$$

where Δ_p is the critical determinant of the unit ball $\mathcal{B}^{\langle p\rangle}=\{(x,y)\in\mathbb{R}^2 : F^{\langle p\rangle}(x,y)<1\}$. The value of Δ_p is given in [1, Section 4] (see also the references therein). It is increasing on $1\leq p\leq\infty$ with $\Delta_1=\frac{1}{2},\Delta_2=\frac{\sqrt{3}}{2}$ and $\Delta_\infty=1$. For any $1\leq p<\infty$, the inequality between arithmetic and geometric means gives

$$s_n|r_n-s_n\alpha|\leq\left(\frac{s_n^p+|r_n-s_n\alpha|^p}{2}\right)^{\frac{2}{p}}<4^{-\frac{1}{p}}\Delta_p^{-1}. \tag{1.2}$$

The right-hand side of this inequality increases from $\frac{1}{2}$ to 1 as p goes from 1 to ∞. The convergents r_n/s_n of the 1-continued fraction, which is Minkowski's diagonal continued fraction [6], thus satisfy

$$s_n|r_n-s_n\alpha|<\tfrac{1}{2}.$$

In fact, Minkowski showed that these r_n/s_n coincide with those regular convergents p_m/q_m that satisfy $q_m|p_m-q_m\alpha|<\frac{1}{2}$. For any $p\geq 1$ the sequence of convergents of the p-continued fraction give a subsequence of the regular convergents, but in general do not give all of those that satisfy $q_m|p_m-q_m\alpha|<(4^{1/p}\Delta_p)^{-1}$. For details and references see [1].

In this paper we generalize these results to L^p-best approximations where $p<1$. In view of the above, our prime motivation is to show that the right-hand side in the generalization of (1.2) for these approximations can be as close to $\frac{1}{\sqrt{5}}$ as desired. This necessitates letting $p\to 0$. For $p\in(0,1)$ the associated generalized continued fraction

$$\alpha=a_0+\frac{\varepsilon_1}{a_1+}\frac{\varepsilon_2}{a_2+}\cdots\qquad (a_0\in\mathbb{Q}\text{ and }\gcd(\varepsilon_n,a_n,\varepsilon_{n+1})=1) \tag{1.3}$$

that we construct in Section 3 (and which we call the p-continued fraction) has integral partial numerators ε_n satisfying $|\varepsilon_n|\leq M_p$, where M_p depends only on p. However, $M_p\to\infty$ as $p\to 0$. Since now $F^{\langle p\rangle}$ is only a quasinorm, the ball $\mathcal{B}^{\langle p\rangle}$ is not convex and we must apply some results of Mordell and Watson [8, 10] from the non-convex geometry of numbers (see also the book of Gruber and Lekkerkerker [3] and the references therein). The theorem below summarizes the main properties of the p-continued fraction.

Theorem 1.1. *For each $\delta > 0$ there exists a $p = p_\delta \in (0,1)$ such that for any irrational α there is a generalized continued fraction of α of the form (1.3) with the following properties.*

(1) The convergents are precisely the best approximations to α with respect to $F^{\langle p \rangle}$.
(2) Each convergent r_n/s_n satisfies $s_n|r_n - s_n\alpha| < \frac{1}{\sqrt{5}} + \delta$.
(3) There exists a constant M_p (depending only on p) such that $|\varepsilon_n| \le M_p$ for all n.

For any $p > 0$ the p-convergents of α form a subsequence of the regular convergents. Computation confirms that more regular convergents are skipped by the p-continued fraction than just those with $s|r - s\alpha|$ large. We now describe the relationship between the p-convergents r_n/s_n and the regular convergents p_n/q_n.

Theorem 1.2. *Fix $p \in (0,1)$ and an irrational $\alpha \in \mathbb{R}$. Let n be a positive integer and let m and ℓ be such that $r_n/s_n = p_m/q_m$ and $r_{n+1}/s_{n+1} = p_{m+\ell}/q_{m+\ell}$. Then*

$$\ell \le \log_\varphi \left(2^{4/p-1}\sqrt{5}\, \frac{\Gamma\big(1+\frac{1}{p}\big)^2}{\Gamma\big(1+\frac{2}{p}\big)} \right),$$

where $\varphi = \frac{1}{2}(1+\sqrt{5})$.

We now show that as p tends to zero, arbitrarily many regular convergents can be skipped in the p-continued fraction. This is the case for every irrational for which the regular continued fraction is 1-periodic. See Table 1.1 for the case $\alpha = \frac{1}{2}(1+\sqrt{5})$.

p	First 10 convergents of the p-continued fraction									
∞	2	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{8}{5}$	$\frac{13}{8}$	$\frac{21}{13}$	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$
0.5	2	$\frac{5}{3}$	$\frac{8}{5}$	$\frac{13}{8}$	$\frac{21}{13}$	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$
0.4	$\frac{5}{3}$	$\frac{8}{5}$	$\frac{13}{8}$	$\frac{21}{13}$	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$	$\frac{377}{233}$
0.3	$\frac{8}{5}$	$\frac{13}{8}$	$\frac{21}{13}$	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$	$\frac{377}{233}$	$\frac{610}{377}$
0.25	$\frac{13}{8}$	$\frac{21}{13}$	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$	$\frac{377}{233}$	$\frac{610}{377}$	$\frac{987}{610}$
0.2	$\frac{21}{13}$	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$	$\frac{377}{233}$	$\frac{610}{377}$	$\frac{987}{610}$	$\frac{1597}{987}$
0.18	$\frac{34}{21}$	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$	$\frac{377}{233}$	$\frac{610}{377}$	$\frac{987}{610}$	$\frac{1597}{987}$	$\frac{2584}{1597}$
0.16	$\frac{55}{34}$	$\frac{89}{55}$	$\frac{144}{89}$	$\frac{233}{144}$	$\frac{377}{233}$	$\frac{610}{377}$	$\frac{987}{610}$	$\frac{1597}{987}$	$\frac{2584}{1597}$	$\frac{4181}{2584}$

TABLE 1.1. Selected values of p with the first several convergents of the p-continued fraction for $\varphi = \frac{1}{2}(1+\sqrt{5})$.

Theorem 1.3. *Let $a \geq 1$ be an integer and let $\alpha = a + \frac{1}{a+}\frac{1}{a+}\frac{1}{a+}\cdots = \frac{1}{2}(a+\sqrt{a^2+4})$. Then for each sufficiently large positive integer m there exists a $p \in (0,1)$ such that $r_0/s_0 = p_m/q_m$.*

The paper is organized as follows. In Section 2 we review Mordell's result on the geometry of numbers for the quasinorms $F^{\langle p\rangle}$. Section 3 contains the algorithm that generates the p-continued fraction and the proof of Theorem 1.1. Sections 4 and 5 address the question of how many of the regular convergents can be skipped by the p-continued fraction; they contain the proofs of Theorems 1.2 and 1.3.

In Sections 2, 3, and 4, the value of $p \in (0,1)$ is fixed, so in those sections we will usually suppress the dependence on p from the notation. In Section 5, the value of p is allowed to vary; there, the notation will reflect the dependence on p.

2. Geometry of numbers for non-convex bodies

Let $F = F^{\langle p\rangle}$, defined in (1.1). When $p \geq 1$, F defines a norm on $\mathbb{R}^2$, but when $p \in (0,1)$, F is not a norm because the triangle inequality does not hold. However, for such p we have the following quasi-triangle inequality, so F defines a quasinorm.

Lemma 2.1. *Let $p \in (0,1)$. Then for any $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$, we have*

$$F(x_1+x_2, y_1+y_2) \leq 2^{\frac{1}{p}-1}\left(F(x_1,y_1)+F(x_2,y_2)\right).$$

Proof. We start by proving that $|a+b|^p \leq |a|^p + |b|^p$ for any $a, b \in \mathbb{R}$. By homogeneity and the usual triangle inequality, it suffices to prove that $(1+t)^p \leq 1+t^p$ for $t \geq 0$, but this follows from the fact that the derivative of $t \mapsto (1+t)^p - t^p - 1$ is nonpositive when $p < 1$.

Using that $|a+b|^p \leq |a|^p + |b|^p$ we obtain

$$F(x_1+x_2, y_1+y_2) \leq \left(F(x_1,y_1)^p + F(x_2,y_2)^p\right)^{\frac{1}{p}}.$$

Now since $t \mapsto t^{1/p}$ is convex we have $(\frac{a+b}{2})^{1/p} \leq \frac{a^{1/p}+b^{1/p}}{2}$, from which it follows that

$$\left(F(x_1,y_1)^p + F(x_2,y_2)^p\right)^{\frac{1}{p}} \leq 2^{\frac{1}{p}-1}\left(F(x_1,y_1)+F(x_2,y_2)\right),$$

as desired. □

Our starting point for studying best approximations with respect to the quasinorm F is the following theorem of Mordell [8, Section 8], which we have updated to reflect later work of Watson [10]. If $L \subseteq \mathbb{R}^2$ is a full lattice, the determinant of L equals $|\det g|$, where $g \in \mathrm{GL}_2(\mathbb{R})$ is any matrix whose rows form a $\mathbb{Z}$-basis for L.

Theorem 2.2 (Mordell and Watson)**.** *Let $p \in (0,1)$ and let (a_p, b_p) be the unique solution to the equations*

$$(a+b)^p + (a-b)^p = a^p + b^p,$$
$$a^2+b^2 = 2,$$

with $a_p > b_p$. Define c_p by

$$c_p = 2^{-\frac{1}{p}} F(a_p, b_p). \tag{2.1}$$

Then every full lattice has a point $(x,y) \neq (0,0)$ satisfying

$$F(x,y) \leq 2^{\frac{1}{p}-\frac{1}{2}}(\det L)^{\frac{1}{2}} c_p. \tag{2.2}$$

Furthermore, for $p \in (0.3295\ldots, 1)$ the inequality is sharp.

Note that the maximum value of the function $(a,b) \mapsto a^p + b^p$ subject to the constraint $a^2 + b^2 = 2$ is 2, and this occurs only when $a = b = 1$. Thus $c_p < 1$ for all $p < 1$. Also, clearly a_p and b_p both approach 1 as p tends to 1, so we have

$$\lim_{p \to 1^-} c_p = 1. \tag{2.3}$$

Applying Theorem 2.2 to the lattice $(0,t)\mathbb{Z} + (t^{-1}, -\alpha t)\mathbb{Z}$, we find that for every $t \geq 1$ there exists a rational r/s with $s > 0$ such that

$$F_t(s, r - s\alpha) \leq 2^{\frac{1}{p} - \frac{1}{2}} c_p. \tag{2.4}$$

Define

$$\beta_p = 2^{\frac{2}{p} - 1} c_p^2. \tag{2.5}$$

Following Mordell, if we let $j_p = b_p/a_p < 1$ then

$$a_p = \sqrt{\frac{2}{j_p^2 + 1}} \quad \text{and} \quad b_p = j_p \sqrt{\frac{2}{j_p^2 + 1}}$$

and from this we find that

$$\beta_p = \frac{(1 + j_p^p)^{2/p}}{1 + j_p^2}. \tag{2.6}$$

Applying the inequality between arithmetic and geometric means and (2.4)–(2.6), we find that for each $t \geq 1$ there exists a rational r/s with $s > 0$ such that

$$s\,|r - s\alpha| \leq 4^{-\frac{1}{p}} \left[F_t(s, r - s\alpha)\right]^2 \leq 4^{-\frac{1}{p}} \beta_p,$$

where equality holds for the first inequality if and only if $s = t^2\,|r - s\alpha|$. Letting $t \to \infty$ we find that

$$s\,|r - s\alpha| < 4^{-\frac{1}{p}} \beta_p \tag{2.7}$$

for infinitely many rational approximations r/s. The proposition below describes the values of $4^{-1/p}\beta_p$ for $p \in (0,1)$. See Figure 2.1.

Proposition 2.3. *Let $p \in (0,1)$. Then*

(1) $\lim_{p \to 1^-} 4^{-1/p}\beta_p = \frac{1}{2}$,
(2) $\lim_{p \to 0^+} 4^{-1/p}\beta_p = \frac{1}{\sqrt{5}}$, *and*
(3) the function $p \mapsto 4^{-1/p}\beta_p$ is increasing on $(0,1)$.

In order to study the behavior of β_p it is apparent that we must understand the behavior of j_p. The next lemma follows from Section 2 of [10]; its proof is elementary but quite tedious.

Lemma 2.4. *The function $p \mapsto j_p$ is differentiable and strictly increasing on $(0,1)$ and*

$$\lim_{p \to 0^+} j_p = \frac{-1 + \sqrt{5}}{2}. \tag{2.8}$$

Using Lemma 2.4 the proof of Proposition 2.3 is relatively straightforward.

Proof of Proposition 2.3. (1) By (2.5) and (2.3) we have

$$\lim_{p \to 1^-} 4^{-\frac{1}{p}} \beta_p = \lim_{p \to 1^-} \frac{1}{2} c_p^2 = \frac{1}{2}.$$

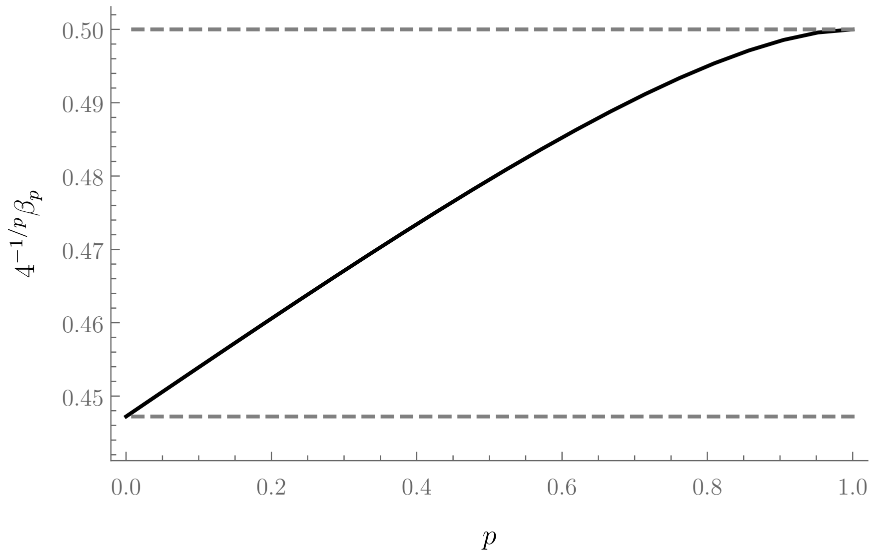

FIGURE 2.1. A plot of $4^{-1/p}\beta_p$ for $p \in (0,1)$, with the limiting values of $\frac{1}{2}$ and $\frac{1}{\sqrt{5}}$ shown as dashed lines.

(2) Let $x \in [\frac{1}{2}, 1]$. For p close to zero we have the Taylor expansion

$$\log(1+x^p) = \log 2 + \tfrac{1}{2}(x^p - 1) + O\left((x^p-1)^2\right).$$

Furthermore,

$$x^p - 1 = \sum_{n=1}^{\infty} \frac{(p\log x)^n}{n!} = p\log x + O(p^2).$$

It follows that

$$4^{-\frac{1}{p}}(1+x^p)^{2/p} = \exp\left(\tfrac{2}{p}\log(1+x^p) - \tfrac{2}{p}\log 2\right) = \exp\left(\log x + O(p)\right) = x + O(p).$$

Thus

$$4^{-\frac{1}{p}}\frac{(1+x^p)^{2/p}}{1+x^2} = \frac{1}{x^{-1}+x} + O(p)$$

uniformly for $x \in [\frac{1}{2}, 1]$. By (2.8) we compute

$$\lim_{p\to 0^+} 4^{-\frac{1}{p}}\beta_p = \lim_{p\to 0^+} 4^{-\frac{1}{p}}\frac{(1+j_p^p)^{2/p}}{1+j_p^2} = \frac{1}{2\left(-1+\sqrt{5}\right)^{-1} + \frac{1}{2}\left(-1+\sqrt{5}\right)} = \frac{1}{\sqrt{5}}.$$

(3) We will prove that the derivative of the function $p \mapsto 4^{-1/p}\beta_p$ is positive. Since $\beta_p > 0$ it is enough to prove that $f'(p) > 0$, where

$$f(p) = \log(4^{-1/p}\beta_p) = -\frac{2\log 2}{p} + \frac{2\log(1+j_p^p)}{p} - \log(1+j_p^2).$$

A straightforward computation yields

$$f'(p) = \frac{2}{p^2}\left(\log 2 - \log(1+j_p^p) - \frac{j_p^p\log(1/j_p^p)}{1+j_p^p}\right) + 2j_p'\left(\frac{j_p^{p-1}}{1+j_p^p} - \frac{j_p}{1+j_p^2}\right).$$

By Lemma 2.4 we have $j_p' > 0$. Since $j_p \in (0,1)$ and $p \in (0,1)$ we have $j_p^{p-2} > 1$ so

$$1 + j_p^p < j_p^{p-2} + j_p^p = j_p^{p-2}(1 + j_p^2),$$

which implies

$$\frac{j_p^{p-1}}{1 + j_p^p} - \frac{j_p}{1 + j_p^2} > 0.$$

Thus to show that $f'(p) > 0$ it suffices to prove that $g(j_p^p) < \log 2$, where

$$g(x) := \log(1 + x) + \frac{x \log(1/x)}{1 + x}.$$

Indeed, we have $g(x) < \log 2$ for any $x \in (0,1)$ because

$$g'(x) = \frac{\log(1/x)}{(1 + x)^2} > 0$$

and $\lim_{x \to 1^-} g(x) = \log 2$. □

3. The p-continued fraction for $p \in (0,1)$

In this section we prove Theorem 1.1. Fix $p \in (0,1)$. We will need the following analogue of Minkowski's first convex body theorem for the non-convex bodies

$$\mathcal{B}_t(P) = \mathcal{B}_t^{\langle p \rangle}(P) = \left\{ P' \in \mathbb{R}^2 : F_t(P') < F_t(P) \right\},$$

where $P \in \mathbb{R}^2$ and $t > 0$. We say that a lattice is admissible for a set S if the only lattice point inside S is the origin.

Proposition 3.1. *Fix $p \in (0,1)$, $t \geq 1$, and $P \in \mathbb{R}^2$. If L is an admissible lattice for $\mathcal{B}_t(P)$ then*

$$\operatorname{area} \mathcal{B}_t(P) \leq C_p \det L,$$

where

$$C_p = 2^{\frac{2}{p}+1} \frac{\Gamma\left(1 + \frac{1}{p}\right)^2}{\Gamma\left(1 + \frac{2}{p}\right)} c_p^2$$

and c_p is given in (2.1).

Proof. Let $R = F_t(P)$. By Theorem 2.2, there is a point $(x, y) \in L$ that satisfies (2.2). But this point is not in $\mathcal{B}_t(P)$ since L is admissible for $\mathcal{B}_t(P)$. Thus

$$R \leq F(x, y) \leq 2^{\frac{1}{p} - \frac{1}{2}} c_p (\det L)^{\frac{1}{2}}. \tag{3.1}$$

The area of $\mathcal{B}_t(P)$ is given by

$$\operatorname{area} \mathcal{B}_t(P) = 4 \int_0^R (R^p - x^p)^{\frac{1}{p}}\, dx = 4R^2 p^{-1} \int_0^1 t^{\frac{1}{p}-1} (1 - t)^{\frac{1}{p}}\, dt = 4R^2 \frac{\Gamma\left(1 + \frac{1}{p}\right)^2}{\Gamma\left(1 + \frac{2}{p}\right)},$$

where we have used [2, (5.12.1)] in the last equality. With (3.1) we obtain

$$\operatorname{area} \mathcal{B}_t(P) \leq 2^{\frac{2}{p}+1} \frac{\Gamma\left(1 + \frac{1}{p}\right)^2}{\Gamma\left(1 + \frac{2}{p}\right)} c_p^2 (\det L),$$

as desired. □

We will generate the convergents of the p-continued fraction using a recursive algorithm, described in Lemma 3.4, that yields a sequence of points in the lattice

$$L_\alpha = (0,1)\mathbb{Z} + (1,-\alpha)\mathbb{Z},$$

where a point $(s, r - s\alpha) \in L_\alpha$ corresponds to a rational r/s. Of course we can assume without loss of generality that $\alpha \in (-\frac{1}{2}, \frac{1}{2})$. The properties of the algorithm will show that the sequence of convergents is, in fact, convergent and is precisely the sequence of best approximations with respect to F, ordered by increasing denominator.

We will use this elementary lemma several times.

Lemma 3.2. *Fix $p \in (0,1)$. If $|x'| \le |x|$ and $|y'| \le |y|$, then $F(x', y') \le F(x, y)$.*

We also have the following property of $\mathcal{B}_t(P)$ as t varies. Roughly speaking, this lemma shows that in the first quadrant, $\mathcal{B}_t(P)$ contracts to the left of P and expands to the right of P as t increases.

Lemma 3.3. *Fix $p \in (0,1)$ and $P = (x_P, y_P) \in \mathbb{R}^2$ with $x_P \ge 0$, and let $t_0, t_1 \in \mathbb{R}$ with $t_1 > t_0 > 0$. Let $x > 0$ and for $j \in \{0,1\}$ define $y_j > 0$ by $F_{t_j}(x, y_j) = F_{t_j}(P)$, when a solution to this equation exists. We have that*

(1) if $x < x_P$, then $y_0 > y_1$, and
(2) if $x > x_P$, then $y_0 < y_1$.

Proof. From the implicit definition of y_j, we find

$$y_j = t_j^{-2}\left(x_P^p - x^p + t_j^{2p}|y_P|^p\right)^{\frac{1}{p}}$$

as an explicit formula for y_j.

For algebraic ease, we consider the quantity $y_0^p - y_1^p$. Since $p > 0$, this quantity will have the same sign as $y_0 - y_1$. Thence

$$\begin{aligned} y_0^p - y_1^p &= t_0^{-2p}\left(x_P^p - x^p + t_0^{2p}|y_P|^p\right) - t_1^{-2p}\left(x_P^p - x^p + t_1^{2p}|y_P|^p\right) \\ &= t_0^{-2p}(x_P^p - x^p) + |y_P|^p - t_1^{-2p}(x_P^p - x^p) - |y_P|^p \\ &= (t_0^{-2p} - t_1^{-2p})(x_P^p - x^p). \end{aligned}$$

Note that the first term in the above expression is positive, since $t_1 > t_0$. If $x < x_P$, then the second term in the above expression is positive, meaning $y_0^p - y_1^p$ is positive and $y_0 > y_1$. If instead $x > x_P$, the argument reverses and $y_0 < y_1$. ∎

The lemma below describes the aforementioned recursive algorithm.

Lemma 3.4. *Fix an irrational $\alpha \in (-\frac{1}{2}, \frac{1}{2})$ and set $P_{-1} = (0,1)$ and $t_{-1} = 1$. For each $m \in \mathbb{Z}$ with $m \ge 0$, there exists a $P_m = (x_m, y_m) \in L_\alpha$ and a $t_m > 1$ with the following properties:*

(1) $x_m > x_{m-1}$ and $|y_m| < |y_{m-1}|$,
(2) L_α is admissible for $\mathcal{B}_{t_m}(P_m)$,
(3) $F_{t_m}(P_{m-1}) = F_{t_m}(P_m)$, and
(4) for any $t \in (t_{m-1}, t_m)$, there is no $P' = (x', y') \in L_\alpha$ different from P_{m-1} with $x' > 0$ and $F_t(P') \le F_t(P_{m-1})$.

Proof. We construct the points P_m inductively. By the construction of L_α, the only points on the unit ball $\mathcal{B} = \mathcal{B}_{t_{-1}}(P_{-1})$ are $\pm P_{-1}$. For any $P = (x, y)$ with $y \neq 0$ we have

$$F_t(P) = F(t^{-1}x, ty) \geq F(0, ty) = t\,|y|\,,$$

so $F_t(P) \to \infty$ as $t \to \infty$. Note that the area of $\mathcal{B}_t(P)$ is given by

$$\text{area } \mathcal{B}_t(P) = F_t^2(P)\text{area } \mathcal{B}.$$

Thus by Proposition 3.1, there must be a maximum t for which L_α is admissible for $\mathcal{B}_t(P_{-1})$. Call this maximum t_0. Among the finitely many $P' = (x', y') \in L_\alpha$ with $F_{t_0}(P') = F_{t_0}(P_{-1})$, there is a unique P' with maximal x' because α is irrational. Let $P_0 = P'$. Note that $x_0 > x_{-1}$. We repeat this process starting with P_0, always choosing the new point with maximal x-coordinate. Let $m \geq 0$. We construct P_{m+1} from P_m by increasing t, starting at t_m, until L_α is no longer admissible for $\mathcal{B}_t(P_m)$, and set t_{m+1} as the maximum t for which L_α is admissible for $\mathcal{B}_t(P_m)$. We call the point on the boundary of $\mathcal{B}_{t_{m+1}}(P_m)$ with maximal x-coordinate P_{m+1}. We show that the sequence of points constructed this way satisfies properties (1)–(4) of the lemma.

(1) For $t > 0$ and $P = (x, y) \in \mathbb{R}^2$, let $\mathcal{S}_t(P)$ denote the inner part of the ball $\mathcal{B}_t(P)$ given by

$$\mathcal{S}_t(P) = \{P' = (x', y') \in \mathbb{R}^2 : |x'| \leq |x| \text{ and } F_t(P') < F_t(P)\}.$$

By Lemmas 3.2 and 3.3 we have $\mathcal{S}_t(P_{m-1}) \subseteq \mathcal{S}_{t_{m-1}}(P_{m-1})$ for all $t \geq t_{m-1}$. Thus the new point P_m has $x_m > x_{m-1}$. Using Lemma 3.3 again, we find that $|y_m| < |y_{m-1}|$.

(2)–(3) These follow from our choice of t_m as the maximal t for which L_α is admissible for $\mathcal{B}_t(P_{m-1})$ and our choice of P_m on the boundary of $\mathcal{B}_{t_m}(P_{m-1})$.

(4) For any $t \in (t_{m-1}, t_m)$, the lattice L_α is admissible for $B_t(P_{m-1})$, so there are no points $P' \in L_\alpha$ with $F_t(P') < F_t(P_{m-1})$. Suppose there is a $t \in (t_{m-1}, t_m)$ and a point $P' \in L_\alpha$ for which $F_t(P') = F_t(P_{m-1})$, i.e. for which P' is on the boundary of $\mathcal{B}_t(P_{m-1})$. Then, by Lemma 3.3, for all $t' > t$ we have $P' \in \mathcal{B}_{t'}(P_{m-1})$, contradicting the definition of t_m. ☐

Figure 3.1 shows a few steps of the algorithm described above.

We claim that the sequence $\{t_m\}$ from Lemma 3.4 tends to infinity. Indeed, the area of the ball $\mathcal{B}_t(P_m)$ is

$$\text{area } \mathcal{B}_t(P_m) = F_t^2(P_m) \text{ area } \mathcal{B},$$

where $\mathcal{B} = \mathcal{B}_1((0,1))$ is the unit ball, so we observe that

$$x_m t_m^{-1} = F(x_m t_m^{-1}, 0) \leq F(x_m t_m^{-1}, y_m t_m) = F_{t_m}(P_m) \leq C_p^{\frac{1}{2}} \text{area } \mathcal{B}^{-\frac{1}{2}},$$

by Lemma 3.2 and Proposition 3.1. Since C_p is fixed for any fixed p and $x_m \to \infty$, we must have that $t_m \to \infty$.

To any sequence $\{r_n/s_n\}$ of rational numbers with $r_n, s_n \in \mathbb{Z}$ and $\gcd(r_n, s_n) = 1$ we can associate a generalized continued fraction with rational coefficients; that is, an expression of the form

$$a_0 + \frac{\varepsilon_1}{a_1+}\,\frac{\varepsilon_2}{a_2+}\,\frac{\varepsilon_3}{a_3+}\cdots := a_0 + \cfrac{\varepsilon_1}{a_1 + \cfrac{\varepsilon_2}{a_2 + \cfrac{\varepsilon_3}{a_3 + \cdots}}}$$

with $a_j, \varepsilon_j \in \mathbb{Q} \setminus \{0\}$ and

$$\frac{r_n}{s_n} = a_0 + \frac{\varepsilon_1}{a_1+}\,\frac{\varepsilon_2}{a_2+}\cdots\frac{\varepsilon_n}{a_n}.$$

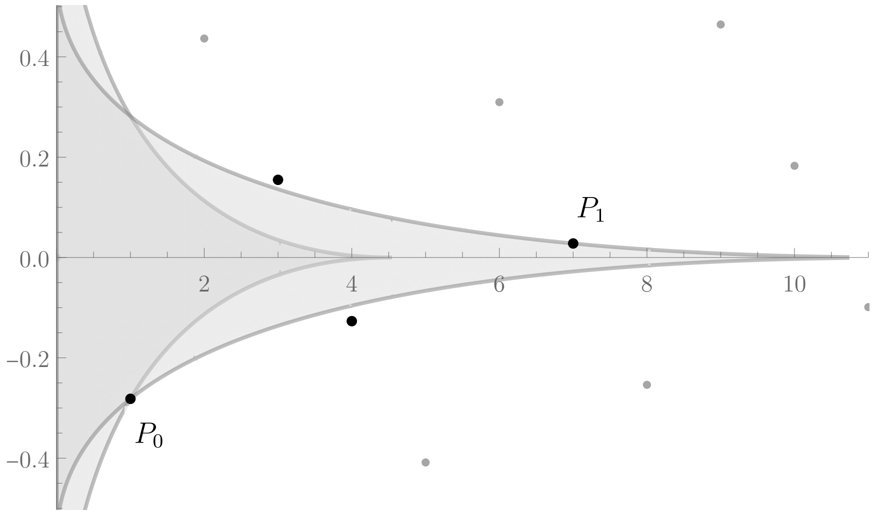

FIGURE 3.1. The lattice L_α for $\alpha = 3 - e$ with $p = 0.5$. Dark lattice points correspond to the regular convergents. The points $P_0 = (1, -\alpha)$ and $P_1 = (7, 2 - 7\alpha)$ give best approximations for α relative to $F^{\langle p \rangle}$.

The ε_j and a_j are defined recursively in terms of r_n and s_n, as we describe below. If the sequence $\{r_n/s_n\}$ converges (resp. diverges), we say that the continued fraction is convergent (resp. divergent). For any sequence $\{\rho_n\}$ of nonzero reals we have the transformation

$$a_0 + \frac{\varepsilon_1}{a_1+}\,\frac{\varepsilon_2}{a_2+}\,\frac{\varepsilon_3}{a_3+}\cdots = a_0 + \frac{\rho_1\varepsilon_1}{\rho_1 a_1+}\,\frac{\rho_1\rho_2\varepsilon_2}{\rho_2 a_2+}\,\frac{\rho_2\rho_3\varepsilon_3}{\rho_3 a_3+}\cdots \tag{3.2}$$

which preserves the convergents, as can be shown by induction. In particular, a continued fraction with rational coefficients can be easily transformed into one with integral coefficients (except possibly a_0) by clearing denominators. Of course, for a generic sequence, this process of clearing denominators will produce arbitrarily large values of $|\varepsilon_j|$ and $|a_j|$.

Fix an irrational $\alpha \in (-\frac{1}{2}, \frac{1}{2})$ and let $P_m = (x_m, y_m)$ as in Lemma 3.4. For $m \geq 1$, let

$$g_m = \begin{pmatrix} x_m & y_m \\ x_{m-1} & y_{m-1} \end{pmatrix}$$

and set

$$g_0 = \begin{pmatrix} x_0 & y_0 \\ 0 & x_0 \end{pmatrix}.$$

Define $\widetilde{a}_m$ and $\widetilde{\varepsilon}_m$ by

$$g_m = \begin{pmatrix} \widetilde{a}_m & \widetilde{\varepsilon}_m \\ 1 & 0 \end{pmatrix} g_{m-1} \tag{3.3}$$

for $m \geq 1$. We further define $s_m = x_m$ and $r_m = y_m + x_m\alpha$. Then r_m and s_m are integers because the point (x_m, y_m) has coordinates (r_m, s_m) in the $\mathbb{Z}$-basis $\{(0,1), (1,-\alpha)\}$. Since

$$g_m = \begin{pmatrix} s_m & r_m - s_m\alpha \\ s_{m-1} & r_{m-1} - s_{m-1}\alpha \end{pmatrix}$$

we see that $\det g_m \in \mathbb{Z} \setminus \{0\}$ for all m and that

$$\begin{pmatrix} r_m & s_m \\ r_{m-1} & s_{m-1} \end{pmatrix} = g_m \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.4}$$

Furthermore, $\widetilde{\varepsilon}_m = -(\det g_m)/(\det g_{m-1}) \in \mathbb{Q}$ and $\widetilde{a}_m \in \mathbb{Q}$. Substituting for g_m in (3.4), we find that

$$\begin{pmatrix} r_m & s_m \\ r_{m-1} & s_{m-1} \end{pmatrix} = \begin{pmatrix} \widetilde{a}_m & \widetilde{\varepsilon}_m \\ 1 & 0 \end{pmatrix} \begin{pmatrix} s_{m-1} & r_{m-1} - s_{m-1}\alpha \\ s_{m-2} & r_{m-2} - s_{m-2}\alpha \end{pmatrix} \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}.$$

From this, we get the recurrence relations

$$\begin{aligned} r_m &= \widetilde{a}_m r_{m-1} + \widetilde{\varepsilon}_m r_{m-2}, & r_0 &= y_0 + x_0\alpha, & r_{-1} &= x_0, \\ s_m &= \widetilde{a}_m s_{m-1} + \widetilde{\varepsilon}_m s_{m-2}, & s_0 &= x_0, & s_{-1} &= 0. \end{aligned}$$

Following the analogous proof from the theory of regular continued fractions, we obtain

$$\frac{r_m}{s_m} = a_0 + \frac{\widetilde{\varepsilon}_1}{\widetilde{a}_1+} \frac{\widetilde{\varepsilon}_2}{\widetilde{a}_2+} \cdots \frac{\widetilde{\varepsilon}_m}{\widetilde{a}_m},$$

where $a_0 = r_0/s_0$. Note that $s_0 \geq 1$, so $a_0 \in \mathbb{Q}$ but we may not have $a_0 \in \mathbb{Z}$.

We prefer to have integer coefficients. While we cannot change that $a_0 \in \mathbb{Q}$, we can use (3.2) to replace $\widetilde{a}_m$ and $\widetilde{\varepsilon}_m$, $m \geq 1$, with integers. To that end, we define

$$a_m = (\det g_{m-1})\widetilde{a}_m$$

for $m \geq 1$. For $m \geq 2$, we define

$$\varepsilon_m = (\det g_{m-2})(\det g_{m-1})\widetilde{\varepsilon}_m$$

and set $\varepsilon_1 = (\det g_0)\widetilde{\varepsilon_1}$. We note that the denominator of $\widetilde{\varepsilon}_m$ is a divisor of $\det g_{m-1}$ by the construction in (3.3). Using the recurrence relations and the fact that $\gcd(r_m, s_m) = 1$ for all m, we find that the same is true for the denominator of $\widetilde{a}_m$. Thus a_m and ε_m are always integral. We then have a generalized continued fraction for r_m/s_m, with only a_0 not strictly integral. We note that this transformation is one of the form given in (3.2), which preserves the convergents to the continued fraction. Thus,

$$\frac{r_m}{s_m} = a_0 + \frac{\varepsilon_1}{a_1+} \frac{\varepsilon_2}{a_2+} \cdots \frac{\varepsilon_m}{a_m}.$$

We are not guaranteed that $\gcd(\varepsilon_m, a_m) = 1$ or even that $\gcd(\varepsilon_m, a_m, \varepsilon_{m+1}) = 1$. However, this generalized continued fraction is unique up to transformation.

Proof of Theorem 1.1. Let $\delta > 0$. By Proposition 2.3, there exists a $p = p_\delta$ such that

$$4^{-\frac{1}{p}}\beta_p \leq \frac{1}{\sqrt{5}} + \delta.$$

(1) The p-convergents are best approximations with respect to F because the associated lattice points satisfy part (4) of Lemma 3.4.

(2) Since every p-convergent is a best approximation with respect to F, (2.7) is satisfied. Thus, every r_m/s_m satisfies

$$s_m\,|r_m - s_m\alpha| < \frac{1}{\sqrt{5}} + \delta.$$

(3) Since $\widetilde{\varepsilon}_m = -(\det g_m)/(\det g_{m-1})$, we have for $m \geq 2$ the bound

$$|\varepsilon_m| \leq |(\det g_{m-2})(\det g_m)|\,.$$

For $m = 1$, we see that $\varepsilon_1 \leq |\det g_m|$. We therefore only need to prove an upper bound on $|\det g_m|$ to obtain a bound for $|\varepsilon_m|$. This is accomplished in the next proposition, which proves that the determinant is bounded and the bound depends only on p. By applying a transformation as in (3.2), we could obtain an equivalent continued fraction where

$\gcd(\varepsilon_m, a_m, \varepsilon_{m+1}) = 1$, but the values of each new $|\varepsilon_m|$ would only be smaller. These new partial numerators would still be bounded by a constant dependent only on p. ☐

Proposition 3.5. *Fix $p \in (0,1)$ and let $\alpha \in (-\frac{1}{2}, \frac{1}{2})$ be an irrational number. With g_m as above, for each $m \geq 1$ we have*

$$|\det g_m| \leq 2^{4/p-1} \frac{\Gamma\left(1+\frac{1}{p}\right)^2}{\Gamma\left(1+\frac{2}{p}\right)} c_p^2. \tag{3.5}$$

Proof. Without loss of generality we may assume that $y_{m-1} > 0$. We divide into two cases depending on the sign of y_m.

Suppose first that $y_m > 0$. Then the points $P_m = (x_m, y_m)$ and $P_{m-1} = (x_{m-1}, y_{m-1})$ are in the first quadrant, with P_{m-1} to the left of P_m. By Lemma 3.4 there exists a $t > 0$ such that $F_t(P_m) = F_t(P_{m-1})$. Let $Q_1 = (t^{-1}x_m, ty_m) = (u_1, v_1)$ and $Q_2 = (t^{-1}x_{m-1}, ty_{m-1}) = (u_2, v_2)$. Then Q_1 and Q_2 both lie on the boundary of $R\mathcal{B}$, where $R = F_t(P_m)$ and $\mathcal{B} = \{P \in \mathbb{R}^2 : F(P) < 1\}$ is the unit ball. Furthermore,

$$|\det g_m| = u_1 v_2 - u_2 v_1.$$

We begin by computing the area of the region D bounded by the line segments OQ_1 and OQ_2 and the portion of the curve $u^p + v^p = R^p$ in the first quadrant. Let $\theta_1 = \arctan(v_1/u_1)$ and $\theta_2 = \arctan(v_2/u_2)$ denote the angles that the line segments OQ_1 and OQ_2 make with the positive x-axis. In polar coordinates (r, θ) the region D is described as

$$\theta_1 \leq \theta \leq \theta_2, \qquad 0 \leq r \leq r(\theta) = \frac{R}{((\cos\theta)^p + (\sin\theta)^p)^{1/p}}.$$

Thus we have

$$\operatorname{area}(D) = \int_{\theta_1}^{\theta_2} \int_0^{r(\theta)} r\, dr d\theta = \frac{R^2}{2} \int_{\theta_1}^{\theta_2} \frac{d\theta}{((\cos\theta)^p + (\sin\theta)^p)^{2/p}}.$$

Making the change of variable $u = \tan\theta$ we find that

$$\operatorname{area}(D) = \frac{R^2}{2} \int_{v_1/u_1}^{v_2/u_2} \frac{du}{(1+u^p)^{2/p}} = \frac{R^2}{2} \int_{v_1/u_1}^{v_2/u_2} \left(\frac{(1+u)^2}{(1+u^p)^{2/p}} \right) \frac{du}{(1+u)^2}.$$

The minimum of the function $u \mapsto \frac{(1+u)^2}{(1+u^p)^{2/p}}$, for positive u, is $2^{2-2/p}$; it occurs at $u = 1$. Thus we have

$$\operatorname{area}(D) \geq 2^{1-2/p} R^2 \int_{v_1/u_1}^{v_2/u_2} \frac{du}{(1+u)^2} = 2^{1-2/p} R^2 \frac{u_1 v_2 - v_1 u_2}{(u_1+v_1)(u_2+v_2)}.$$

The maximum value of $(u, v) \mapsto u + v$ subject to the constraint $u^p + v^p = R^p$ is R; it occurs when $(u, v) = (R, 0)$. Thus we have

$$\operatorname{area}(D) \geq 2^{1-2/p} |\det g_m|.$$

There are two copies of D inside $R\mathcal{B} = \mathcal{B}_1(Q_1)$, so we conclude by Proposition 3.1 that

$$|\det g_m| \leq 2^{2/p-2}(2\operatorname{area}(D)) \leq 2^{2/p-2}\operatorname{area}(\mathcal{B}_1(Q_1)) < 2^{4/p-1} \frac{\Gamma\left(1+\frac{1}{p}\right)^2}{\Gamma\left(1+\frac{2}{p}\right)} c_p^2.$$

If $y_m < 0$, define Q_1 and Q_2 as before and define $Q_1' = (u_1, -v_1)$. Let D denote the region bounded by the line segments OQ_1' and OQ_2 and the portion of the curve $u^p + v^p = R^p$ in the first quadrant. Then we have

$$\operatorname{area}(D) \geq 2^{1-2/p}(u_1v_2 + v_1u_2).$$

Let D' denote the triangle OQ_1Q_1'. Then

$$\operatorname{area}(D') = u_1|v_1| \geq u_2|v_1| \geq 2^{2-2/p}u_2|v_1| = -2^{2-2/p}u_2v_1,$$

so

$$\operatorname{area}(D \cup D') \geq 2^{1-2/p}(u_1v_2 - v_1u_2) = 2^{1-2/p}|\det g_m|.$$

The remainder of the proof proceeds as in the first case. □

4. Regular convergents skipped by the p-continued fraction

In this section we prove Theorem 1.2 which concerns the question of how many of the regular convergents of a given irrational α are skipped by the p-continued fraction. We first list several well-known results about the regular continued fraction. Fix an irrational α. For $p \in (0,1)$, let p_n/q_n denote the regular convergents of α and r_n/s_n the p-convergents. First, p_n and q_n are defined recursively by

$$\begin{aligned} p_n &= b_np_{n-1} + p_{n-2}, & p_{-2} &= 0, & p_{-1} &= 1, \\ q_n &= b_nq_{n-1} + q_{n-2}, & q_{-2} &= 1, & q_{-1} &= 0, \end{aligned}$$

where b_n is the n-th regular partial quotient. Additionally, we have that for $n \geq -1$,

$$q_np_{n-1} - p_nq_{n-1} = (-1)^n, \tag{4.1}$$

and for $n \geq 0$,

$$q_np_{n-2} - p_nq_{n-2} = (-1)^{n-1}b_n. \tag{4.2}$$

These results and their proofs can be found in Section 3 of [9, Ch. I].

We claim that each convergent of the p-continued fraction of α is also a regular convergent. Indeed, suppose that r/s with $s > 0$ is a convergent of the p-continued fraction of α. Then there exists a $t > 1$ such that for any $r'/s' \neq r/s$ with $s' > 0$, we have that

$$F_t(s, r - s\alpha) < F_t(s', r' - s'\alpha).$$

If $s' \leq s$ then by Lemma 3.2 we have that $|r - s\alpha| < |r' - s'\alpha|$. Lagrange [4] showed that all such r/s are regular convergents of α.

The number of consecutive regular convergents p_n/q_n that can be skipped by the p-continued fraction is closely related to the determinants that appear in the following proposition.

Proposition 4.1. *Notation as above, let b_n denote the n-th regular partial quotient for α and for $\ell \geq 0$ define*

$$D_\ell(n) = (-1)^{n-1}\left(q_{n-1}p_{n+\ell} - p_{n-1}q_{n+\ell}\right).$$

Then $D_0(n) = 1$, $D_1(n) = b_{n+1}$, and for $\ell \geq 2$, we have

$$D_\ell(n) = b_{n+\ell}D_{\ell-1}(n) + D_{\ell-2}(n).$$

Proof. For the preliminary cases $\ell \in \{0,1\}$ we have $D_0(n) = 1$ and $D_1(n) = b_{n+1}$ by (4.1) and (4.2). For $\ell \geq 2$, using the recursive definitions of $p_{n+\ell}$ and $q_{n+\ell}$ we have that

$$\begin{aligned} D_\ell(n) &= (-1)^{n-1}\left(q_{n-1}(b_{n+\ell}p_{n+\ell-1} + p_{n+\ell-2}) - p_{n-1}(b_{n+\ell}q_{n+\ell-1} + q_{n+\ell-2})\right) \\ &= b_{n+\ell}(-1)^{n-1}(q_{n-1}p_{n+\ell-1} - p_{n-1}q_{n+\ell-1}) + (-1)^{n-1}(q_{n-1}p_{n+\ell-2} - p_{n-1}q_{n+\ell-2}) \\ &= b_{n+\ell}D_{\ell-1}(n) + D_{\ell-2}(n), \end{aligned}$$

as desired. □

Proof of Theorem 1.2. By Proposition 4.1, using that the regular partial quotients of α satisfy $b_n \geq 1$ for all n, we have the rough lower bound

$$D_\ell(n) \geq F_{\ell+1} \geq \frac{\varphi^\ell}{\sqrt{5}},$$

where F_ℓ is the ℓ-th Fibonacci number and $\varphi = \frac{1}{2}(1+\sqrt{5})$. On the other hand, $D_\ell(n) = |\det g_m|$ for some m, so by Proposition 3.5 we have $D_\ell(n) \leq C_p$ where C_p is the expression on the right hand side in (3.5). Thus

$$\ell \leq \log_\varphi(\sqrt{5}C_p).$$

□

5. Proof of Theorem 1.3

We begin by stating several lemmas regarding 1-periodic continued fractions. For the remainder of this section we assume that $a \in \mathbb{Z}$ with $a \geq 1$ and that

$$\alpha = a + \frac{1}{a+}\,\frac{1}{a+}\,\frac{1}{a+}\cdots = \frac{a+\sqrt{a^2+4}}{2}. \tag{5.1}$$

Lemma 5.1. *Suppose that α satisfies (5.1) and define a sequence R_n by $R_0 = 0$, $R_1 = 1$, and*

$$R_n = aR_{n-1} + R_{n-2} \qquad \text{for } n \geq 2. \tag{5.2}$$

Then for all $n \geq 0$ we have

$$R_n = \frac{\alpha^n - (-\alpha)^{-n}}{\alpha + \alpha^{-1}}. \tag{5.3}$$

Proof. We can rewrite (5.2) using matrices, letting us solve for any R_{n+1} and R_n for $n \geq 1$ by writing

$$\begin{pmatrix} R_{n+1} \\ R_n \end{pmatrix} = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The eigenvalues of $\left(\begin{smallmatrix} a & 1 \\ 1 & 0 \end{smallmatrix}\right)$ are α and $-\alpha^{-1}$, where α is given by (5.1) and we have used that $\alpha = a + \alpha^{-1}$. Diagonalizing, we obtain

$$\begin{pmatrix} R_{n+1} \\ R_n \end{pmatrix} = \frac{1}{\alpha+\alpha^{-1}} \begin{pmatrix} 1 & 1 \\ \alpha^{-1} & -\alpha \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha^{-1} \end{pmatrix}^n \begin{pmatrix} \alpha & 1 \\ \alpha^{-1} & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Further simplification yields

$$\begin{pmatrix} R_{n+1} \\ R_n \end{pmatrix} = \frac{1}{\alpha+\alpha^{-1}} \begin{pmatrix} \alpha^{n+1} - (-\alpha)^{-n-1} & \alpha^n - (-\alpha)^{-n} \\ \alpha^n - (-\alpha)^{-n} & \alpha^{n-1} - (-\alpha)^{-n+1} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

This completes the proof. □

Lemma 5.2. *If α satisfies (5.1) then $p_{n-1} = q_n$ for all $n \geq 1$.*

Proof. Since α satisfies (5.1), every partial quotient is a. Then because $q_0 = 1$, p_n and q_n each satisfy with offset the recursive definition for R_n given in Lemma 5.1. In particular, $p_n = R_{n+2}$ and $q_n = R_{n+1}$. Thus $p_{n-1} = R_{n+1} = q_n$. $\square$

The formula (5.3) can also be written as

$$R_n = \frac{\alpha^n - \cos(\pi n)\alpha^{-n}}{\alpha + \alpha^{-1}}, \tag{5.4}$$

which extends naturally to a smooth function $n \mapsto R_n$ for real n. This point of view will be particularly useful later.

Lemma 5.3. *Suppose that α satisfies (5.1) and define R_n by (5.4). Then*

$$R_{n+1} - R_n\alpha = \cos(\pi n)\alpha^{-n}.$$

Proof. By (5.4) we have

$$\begin{aligned} R_{n+1} - R_n\alpha &= \frac{\alpha^{n+1} + \cos(\pi n)\alpha^{-n-1} - \alpha\left(\alpha^n - \cos(\pi n)\alpha^{-n}\right)}{\alpha + \alpha^{-1}} \\ &= \cos(\pi n)\frac{\alpha^{-n-1} + \alpha^{-n+1}}{\alpha + \alpha^{-1}} = \cos(\pi n)\alpha^{-n}, \end{aligned}$$

as desired. $\square$

We next define a curve and examine its intersection with the lattice points corresponding to the regular convergents of α. Let n be a nonnegative integer and let $Q_n = (R_n, |R_{n+1} - R_n\alpha|)$. This point is either the lattice point corresponding to a regular convergent or the reflection of said point across the x-axis because of Lemma 5.2 and our choice of α. Fix an $m \in \mathbb{N}$. We define $\mathcal{C}_m^{\langle p\rangle}$ to be the portion of the boundary of the ball $\mathcal{B}_t^{\langle p\rangle}((0,1))$ in the first quadrant, where t is such that Q_m is on the boundary of $\mathcal{B}_t^{\langle p\rangle}((0,1))$. See Figure 5.1. We can describe the curve $\mathcal{C}_m^{\langle p\rangle}$ algebraically by the equation

$$t^p = (xt^{-1})^p + (yt)^p. \tag{5.5}$$

Since Q_m is on $\mathcal{C}_m^{\langle p\rangle}$, we can write

$$t = \left(\frac{R_m^p}{1 - |R_{m+1} - R_m\alpha|^p}\right)^{\frac{1}{2p}}. \tag{5.6}$$

We claim there exists a unique $p = p(m)$ such that Q_{m+1} is also on $\mathcal{C}_m^{\langle p\rangle}$. Equivalently,

$$\left(\frac{R_m^p}{1 - |R_{m+1} - R_m\alpha|^p}\right)^{\frac{1}{2p}} = t = \left(\frac{R_{m+1}^p}{1 - |R_{m+2} - R_{m+1}\alpha|^p}\right)^{\frac{1}{2p}}.$$

Simplification yields

$$R_m^p - R_m^p\,|R_{m+2} - R_{m+1}\alpha|^p = R_{m+1}^p - R_{m+1}^p\,|R_{m+1} - R_m\alpha|^p,$$

or, after Lemma 5.3,

$$\left(\frac{R_{m+1}}{R_m}\right)^p = \frac{1 - \alpha^{-mp-p}}{1 - \alpha^{-mp}}. \tag{5.7}$$

While we are primarily interested in solving (5.7) when m is an integer, it will be useful to show that a unique solution $p = p(m)$ exists for all sufficiently large real m.

Lemma 5.4. *If $m \geq 3$ then there exists a unique $p = p(m) \in (0,1)$ satisfying (5.7).*

Proof. Let $m \geq 3$ and define

$$f(p) = \frac{1-\alpha^{-mp-p}}{1-\alpha^{-mp}}.$$

We claim that $f(p)$ is strictly decreasing for $p \in (0,\infty)$. Assuming this for now, and using that $f(p) > 1$ we find that

$$\frac{d}{dp}\left[f(p)\right]^{\frac{1}{p}} = \left(\frac{f'(p)}{pf(p)} - \frac{\log f(p)}{p^2}\right)\left[f(p)\right]^{\frac{1}{p}} < 0.$$

By L'Hôpital's rule, $f(p) \to 1+\frac{1}{m}$ as $p \to 0$, so $\lim_{p\to 0^+}[f(p)]^{1/p} = \infty$. Since $\lim_{p\to\infty}[f(p)]^{1/p} = 1$, we see that $p \mapsto [f(p)]^{1/p}$ is a bijection from $(0,\infty)$ to $(1,\infty)$. Thus there is a unique p for which $[f(p)]^{1/p} = R_{m+1}/R_m$. We claim that this p must be in $(0,1)$. Indeed, when $p \geq 1$ we have

$$[f(p)]^{\frac{1}{p}} \leq f(1) = \frac{1-\alpha^{-m-1}}{1-\alpha^{-m}} = 1 + \alpha^{-m}\frac{1-\alpha^{-1}}{1-\alpha^{-m}} < 1 + \frac{1}{\alpha^3}.$$

On the other hand, by Lemma 5.3 we have

$$\frac{R_{m+1}}{R_m} \geq \alpha - \frac{1}{\alpha^m R_m} \geq \alpha - \frac{1}{\alpha^3}.$$

Since $x > 1 + 2/x^3$ for all $x > 1.6$ and since $\alpha > 1.618$, we see that $[f(p)]^{1/p} < R_{m+1}/R_m$, so the p satisfying (5.7) must be in $(0,1)$. Thus it remains to show that $f(p)$ is decreasing.

We have

$$f'(p) = -\frac{\left[m(1-\alpha^{-p}) - \alpha^{-p}(1-\alpha^{-mp})\right]\alpha^{-mp}\log\alpha}{(1-\alpha^{-mp})^2},$$

so it is enough to show that

$$m(1-\alpha^{-p}) > \alpha^{-p}(1-\alpha^{-mp}).$$

For any $m \geq 3$ and any fixed $\beta \in (0,1)$ we have $1-\beta^m < m(1-\beta)$ since the two sides agree when $m = 1$ and the left side grows slower than the right as m increases. □

We claim that for $p = p(m)$ as above with sufficiently large m, L_α is admissible for the ball $\mathcal{B}_t^{\langle p\rangle}((0,1))$. This is equivalent to Q_{m+k} being above $\mathcal{C}_m^{\langle p\rangle}$ for any $k \in \mathbb{Z}\setminus\{0,1\}$ such that $k > -m$. By solving (5.5) for y, and comparing with the y-coordinate of Q_{m+k}, we find that we can write the condition of Q_{m+k} being above $\mathcal{C}_m^{\langle p\rangle}$ as

$$|R_{m+k+1} - R_{m+k}\alpha| > \left(1 - R_{m+k}^p t^{-2p}\right)^{\frac{1}{p}}.$$

After some simplification and substituting the expression in (5.6) for t, we see that for a fixed m, Q_{m+k} remains above $\mathcal{C}_m^{\langle p\rangle}$ if and only if

$$R_{m+k}^p - R_{m+k}^p\,|R_{m+1} - R_m\alpha|^p > R_m^p - R_m^p\,|R_{m+k+1} - R_{m+k}\alpha|^p.$$

By Lemma 5.3, this can be simplified to

$$\left(\frac{R_{m+k}}{R_m}\right)^p > \frac{1-\alpha^{-mp-kp}}{1-\alpha^{-mp}}. \tag{5.8}$$

Lemma 5.5. *For each real number $m \geq 3$, define $p = p(m) \in (0,1)$ by (5.7). Then $p \sim \frac{\log 2}{m\log\alpha}$ as $m \to \infty$. Furthermore, for all sufficiently large m, for any $k \in \mathbb{Z}\setminus\{0,1\}$ with $k > -m$, the inequality (5.8) holds.*

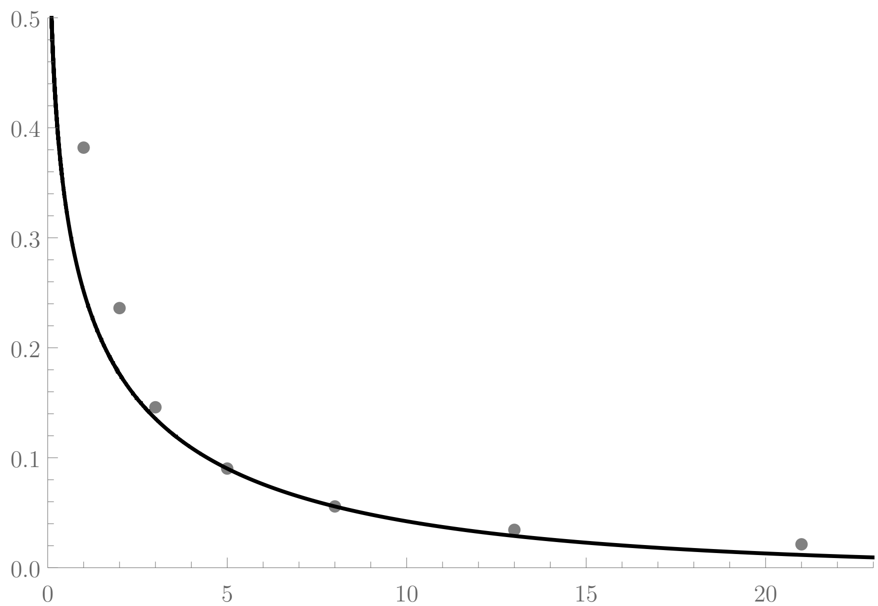

FIGURE 5.1. The edge of the ball, $\mathcal{C}_m^{\langle p \rangle}$, for $p \approx .27$, with $\alpha = \varphi = \frac{1+\sqrt{5}}{2}$ and $m = 5$. The points Q_{m+k} for $k \in \{-3, -2, -1, 0, 1, 2, 3\}$ are presented here.

Proof. Let

$$F(m,p) = \left(\frac{R_{m+1}}{R_m}\right)^p - \frac{1-\alpha^{-mp-p}}{1-\alpha^{-mp}}.$$

Then $F(m,p)$ is (at least) thrice continuously differentiable in the region $m > 0$ and $p > 0$. So by the implicit function theorem the function $p(m)$ is thrice continuously differentiable. We claim that $p(m) \to 0$ as $m \to \infty$. Indeed, if $\limsup_{m\to\infty} p(m) = L > 0$ then since $R_{m+1}/R_m \to \alpha$ as $m \to \infty$ we would have $\alpha^L = 1$, a contradiction. Since $p(m) \in (0,1)$ we must have that $p(m) \to 0$.

Let $k \in \mathbb{Z} \setminus \{0\}$. Using (5.4) we have

$$\begin{aligned} \frac{R_{m+k}}{R_m} &= \alpha^k \frac{1-(-1)^k \cos(\pi m)\alpha^{-2m-2k}}{1-\cos(\pi m)\alpha^{-2m}} \\ &= \alpha^k \left(1 + \cos(\pi m)\alpha^{-2m} \frac{1-(-1)^k\alpha^{-2k}}{1-\cos(\pi m)\alpha^{-2m}}\right) \\ &= \alpha^k \left(1 + O(\alpha^{-2m})\right). \end{aligned}$$

Thus we have

$$\log\left(\frac{R_{m+k}}{R_m}\right) = k\log\alpha + O\left(\alpha^{-2m}\right). \tag{5.9}$$

For large m, we have the asymptotic expansion

$$p = p(m) = \frac{p_0}{m} + \frac{p_1}{m^2} + O\left(\frac{1}{m^3}\right)$$

for some constants p_0, p_1. To determine p_0, we use that $p = p_0/m + O(1/m^2)$ to get

$$\alpha^{-mp} = \alpha^{-p_0} + O(1/m), \quad \frac{1}{1-\alpha^{-mp}} = \frac{\alpha^{p_0}}{\alpha^{p_0}-1} + O(1/m),$$
$$1-\alpha^{-kp} = \frac{kp_0\log\alpha}{m} + O(1/m^2),$$

from which it follows that

$$\log\left(\frac{1-\alpha^{-mp-kp}}{1-\alpha^{-mp}}\right) = \log\left(1+\alpha^{-mp}\frac{1-\alpha^{-kp}}{1-\alpha^{-mp}}\right) = \log\left(1+\frac{kp_0\log\alpha}{(\alpha^{p_0}-1)m}+O\left(\frac{1}{m^2}\right)\right)$$
$$= \frac{kp_0\log\alpha}{(\alpha^{p_0}-1)m} + O\left(\frac{1}{m^2}\right). \tag{5.10}$$

Hence, by the logarithm of (5.7), (5.9), and (5.10) with $k=1$ we must have

$$\frac{p_0\log\alpha}{m} = \frac{p_0\log\alpha}{(\alpha^{p_0}-1)m}.$$

Thus $\alpha^{p_0}=2$. This information significantly simplifies the determination of p_1. Using that $p = \log 2/(m\log\alpha) + p_1/m^2 + O(1/m^3)$ we find that

$$\alpha^{-mp} = \frac{1}{2} - \frac{p_1\log\alpha}{2m} + O(1/m^2), \quad \frac{1}{1-\alpha^{-mp}} = 2 - \frac{2p_1\log\alpha}{m} + O(1/m^2),$$
$$1-\alpha^{-kp} = \frac{k\log 2}{m} + \frac{2kp_1\log\alpha - k^2(\log 2)^2}{2m^2} + O(1/m^3).$$

Thus

$$\log\left(\frac{1-\alpha^{-mp-kp}}{1-\alpha^{-mp}}\right) = \frac{k\log 2}{m} - \frac{k^2(\log 2)^2 - kp_1\log\alpha + kp_1\log 4\log\alpha}{m^2} + O\left(\frac{1}{m^3}\right)$$

Again, by (5.7), (5.9), and (5.10) we find that

$$\frac{p_1\log\alpha}{m^2} = -\frac{(\log 2)^2 - p_1\log\alpha + p_1\log 4\log\alpha}{m^2}.$$

It follows that $p_1 = -\log 2/(2\log\alpha)$. Thus

$$p = \frac{\log 2}{m\log\alpha}\left(1-\frac{1}{2m}\right) + O\left(\frac{1}{m^3}\right). \tag{5.11}$$

Now we assume that $k \neq 1$. Using (5.9), (5.10), and (5.11), the logarithm of the inequality (5.8) becomes

$$\frac{k\log 2}{m} - \frac{k\log 2}{2m^2} + O\left(\frac{1}{m^3}\right) > \frac{k\log 2}{m} - \frac{k\log 2}{2m^2}(2k\log 2 + 1 - 2\log 2) + O\left(\frac{1}{m^3}\right),$$

which, for sufficiently large m, is true for any $k \in \mathbb{Z}\setminus\{0,1\}$. □

For $p\in(0,1)$ and $t\geq 1$ the closure of the ball $\mathcal{B}_t^{\langle p\rangle}((0,1))$ is the set of points $(x,y)\in\mathbb{R}^2$ satisfying $(|x|/t)^p + (|y|t)^p \leq t^p$. This inequality is only satisfied for $|y|\leq 1$, so it is equivalent to the condition

$$0 \leq \frac{|x|^p}{1-|y|^p} \leq t^{2p}.$$

Thus $\mathcal{B}_t^{\langle p\rangle}((0,1)) \subseteq \mathcal{B}_{t'}^{\langle p\rangle}((0,1))$ for $t'\geq t$. Let m be sufficiently large and let $p=p(m)$. Then the previous lemma shows that the first two lattice points that touch the boundary of $\mathcal{B}_t^{\langle p \rangle}((0,1))$ as t increases from 1 must be the lattice points corresponding to Q_m and Q_{m+1}, and furthermore the lattice must be admissible for this ball. This then gives us that $r_0/s_0 = R_{m+2}/R_{m+1} = p_m/q_m$. □

References

[1] Nickolas Andersen and William Duke. On a theorem of Davenport and Schmidt. *Acta Arith.*, 198(1):37–75, 2021.

[2] *NIST Digital Library of Mathematical Functions.* http://dlmf.nist.gov/, Release 1.1.4 of 2022-01-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[3] P. M. Gruber and C. G. Lekkerkerker. *Geometry of numbers*, volume 37 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1987.

[4] J. L. Lagrange. *Additions aux élements d'algebra d'Euler*, volume VII of *Oeuvres*.

[5] Hermann Minkowski. Généralisation de la théorie des fractions continues. *Ann. Sci. École Norm. Sup. (3)*, 13:41–60, 1896.

[6] Hermann Minkowski. Ueber die Annäherung an eine reelle Grösse durch rationale Zahlen. *Math. Ann.*, 54(1-2):91–124, 1900.

[7] Hermann Minkowski. *Geometrie der Zahlen.* Bibliotheca Mathematica Teubneriana, Band 40. Johnson Reprint Corp., New York-London, 1968.

[8] L. J. Mordell. On the geometry of numbers in some non-convex regions. *Proc. London Math. Soc. (2)*, 48:339–390, 1945.

[9] Wolfgang M. Schmidt. *Diophantine approximation*, volume 785 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

[10] G. N. Watson. A problem in elementary analysis proposed by Mordell. *Proc. London Math. Soc. (2)*, 48:391–400, 1945.

Brigham Young University Mathematics Department, Provo UT 84602
Email address: nick@math.byu.edu

UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095-1555
Email address: wdduke@ucla.edu

Brigham Young University Mathematics Department, Provo UT 84602
Email address: zmhacking@gmail.com

Brigham Young University Mathematics Department, Provo UT 84602
Email address: amy.woodall713@gmail.com