THE EMBEDDING CAPACITY OF 4-DIMENSIONAL SYMPLECTIC ELLIPSOIDS

DUSA MCDUFF AND FELIX SCHLENK

ABSTRACT. This paper calculates the function $c(a)$ whose value at a is the infimum of the size of a ball that contains a symplectic image of the ellipsoid $E(1, a)$. (Here $a \geq 1$ is the ratio of the area of the large axis to that of the smaller axis.) The structure of the graph of $c(a)$ is surprisingly rich. The volume constraint implies that $c(a)$ is always greater than or equal to the square root of a, and it is not hard to see that this is equality for large a. However, for a less than the fourth power τ^4 of the golden ratio, $c(a)$ is piecewise linear, with graph that alternately lies on a line through the origin and is horizontal. We prove this by showing that there are exceptional curves in blow ups of the complex projective plane whose homology classes are given by the continued fraction expansions of ratios of Fibonacci numbers. On the interval $\left[\tau^4, 7\right]$ we find $c(a) = \frac{a+1}{3}$. For $a \geq 7$, the function $c(a)$ coincides with the square root except on a finite number of intervals where it is again piecewise linear. The embedding constraints coming from embedded contact homology give rise to another capacity function c_{ECH} which may be computed by counting lattice points in appropriate right angled triangles. According to Hutchings and Taubes, the functorial properties of embedded contact homology imply that $c_{ECH}(a) \leq c(a)$ for all a. We show here that $c_{ECH}(a) \geq c(a)$ for all a.

CONTENTS

Date: April 15, 2011.
2000 *Mathematics Subject Classification.* 53D05, 14B05, 32S05, 11A55.
Key words and phrases. symplectic embeddings, Fibonacci numbers.
partially supported by NSF grant DMS 0604769.
partially supported by SNF grant 200021-125352/1.

1. Introduction

1.1. **Statement of results.** As has been known since the time of Gromov's Nonsqueezing Theorem, questions about symplectic embeddings lie at the heart of symplectic geometry. To date, most results have concerned the embeddings of balls or of products of balls since these are most amenable to analysis. (See Cieliebak, Hofer, Schlenk and Latschev [5] for a comprehensive survey of embedding problems.) However, ellipsoids are another very natural class of examples. As pointed out by Hofer, the simplicity of the characteristic flow on their boundary makes them a natural test case for understanding the role of variational properties in symplectic geometry. One would like to understand the extent to which obstructions coming from periodic orbits capture all symplectic invariants. Judging from the evidence of the current work, it seems one cannot take a naive approach. As pointed out in McDuff [16], the Ekeland–Hofer capacities of [6] (which are purely variational) do not give all obstructions. Instead one must use invariants coming from the Hutchings–Taubes [12] embedded contact homology, which has an unavoidably geometric flavor. Indeed Taubes [25] has recently shown it equals a version of Seiberg–Witten Floer homology and so is a gauge theory.

In view of the work of Guth [7] on higher dimensional symplectic embedding questions, there has been renewed interest in this kind of question. However, we restrict consideration to four dimensions since the methods and results in this case are very different from those in higher dimensions; cf. Remark 1.1.4. For relevant background and a survey of the results of the current paper see [17].

Given a real number $a \geq 1$ denote by $E(1, a)$ the closed ellipsoid

$$E(1,a) \;:=\; \Big\{ x_1^2 + x_2^2 + \frac{x_3^2 + x_4^2}{a} \leq 1 \Big\} \;\subset\; \mathbb{R}^4.$$

This paper studies the function $c\colon [1,\infty) \to \mathbb{R}$ defined by

$$c(a) \;:=\; \inf\Big\{\mu : E(1,a) \overset{s}{\hookrightarrow} B(\mu)\Big\} \tag{1.1.1}$$

where $B(\mu) := \Big\{\sum x_i^2 \le \mu\Big\}$ is the ball of radius $\sqrt{\mu}$, and $A \overset{s}{\hookrightarrow} B$ means that A embeds symplectically in B. This is one of a range of symplectic capacity functions defined by Cieliebak, Hofer, Latschev and Schlenk in [5], and the first to be calculated.

Since $E(1,a)$ has volume $a\pi^2/2$, we must have $c(a) \ge \sqrt{a}$. Here is another elementary result.

Lemma 1.1.1. *The function c is nondecreasing and continuous. Further, it has the following scaling property:*

$$\frac{c(\lambda a)}{\lambda a} \le \frac{c(a)}{a} \quad \textit{when } \lambda > 1. \tag{1.1.2}$$

Proof. The first statement is clear. The second holds because $E(1,\lambda a) \subset \sqrt{\lambda}\, E(1,a)$ when $\lambda > 1$ and also $E(1,a) \overset{s}{\hookrightarrow} B(\mu)$ if and only if $\sqrt{\lambda}\, E(1,a) \overset{s}{\hookrightarrow} \sqrt{\lambda}\, B(\mu) = B(\lambda\mu)$. □

The function $c(a)$ was calculated[1] in [16] for integral a as follows:

$$c(a) = \sqrt{a} \;\text{ if } a \in \mathbb{N} \text{ is } 1, 4 \text{ or } \ge 9, \tag{1.1.3}$$
$$c(2) = c(3) = c(4) = 2,\;\; c(5) = c(6) = \tfrac{5}{2},\;\; c(7) = \tfrac{8}{3},\;\; c(8) = \tfrac{17}{6}.$$

Its monotonicity and scaling property are then enough to determine all its values for $a \le 6$: it is constant on the intervals $[2,4]$ and $[5,6]$ and otherwise linear, with graph along appropriate lines through the origin. (See Figure 1.1 and Corollaries 1.2.4 and 1.2.8 below.)

It turns out that the two steps of $c(a)$ that we described above extend to an infinite stairs for $a \in [1,\tau^4]$ where $\tau^4 = \frac{7+3\sqrt{5}}{2}$ is the fourth power of the golden ratio $\tau := \frac{1+\sqrt{5}}{2}$. We call this Fibonacci stairs. Denote by $g_n := f_{2n-1}$, $n \ge 1$, the terms in the odd places of the Fibonacci sequence f_n. (For short, we call these the "odd Fibonacci numbers".) Thus the sequence g_n starts with $1,\ 2,\ 5,\ 13,\ 34,\ \ldots.$ Set $g_0 = 1$ and for each $n \ge 0$ define

$$a_n \;=\; \Big(\tfrac{g_{n+1}}{g_n}\Big)^2 \qquad\text{and}\qquad b_n \;=\; \tfrac{g_{n+2}}{g_n}.$$

Then

$$a_0 = 1 \;<\; b_0 = \tfrac{2}{1} = 2 \;<\; a_1 = (\tfrac{2}{1})^2 = 4 \;<\; b_1 = \tfrac{5}{1} = 5 \;<$$
$$a_2 = (\tfrac{5}{2})^2 = 6\tfrac{1}{4} \;<\; b_2 = \tfrac{13}{2} = 6\tfrac{1}{2} \;<\; a_3 = \big(\tfrac{13}{5}\big)^2 = 6\tfrac{19}{25} \;<\; b_3 = \tfrac{34}{5} = 6\tfrac{4}{5} \;<\; \ldots.$$

More generally,

$$\ldots < a_n < b_n < a_{n+1} < b_{n+1} < \ldots, \qquad \text{and } \lim a_n = \lim b_n = \tau^4 \approx 6.854.$$

[1]The first nontrivial result here, that $c(4) = 2$, was proved earlier by Opshtein in [21]. See also Theorem 4 in Opshtein [22].

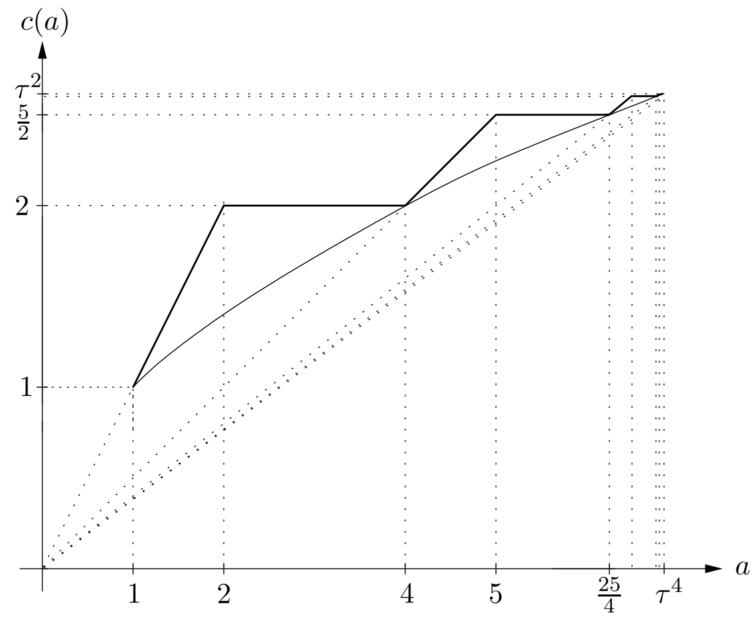

FIGURE 1.1. The Fibonacci stairs: The graph of $c(a)$ on $\left[1, \tau^4\right]$.

Theorem 1.1.2. (i) *For each $n \geq 0$, $c(a) = \frac{a}{\sqrt{a_n}}$ for $a \in [a_n, b_n]$, and c is constant with value $\sqrt{a_{n+1}}$ on the interval $[b_n, a_{n+1}]$.*

(ii) $c(a) = \frac{a+1}{3}$ *on* $[\tau^4, 7]$.

(iii) *There are a finite number of closed disjoint intervals $I_j \subset [7, 8\frac{1}{36}]$ such that $c(a) = \sqrt{a}$ for all $a > 7$, $a \notin I_j$. Moreover, c is piecewise linear in each I_j, with one non-smooth point in the interior of I_j.*

(iv) $c(a) = \sqrt{a}$ *for* $a \geq 8\frac{1}{36}$.

The argument proving part (i) hinges on the existence of an unexpected relation between the function $c(a)$ and the Fibonacci numbers. We shall see that this relation persists for a just larger than τ^4, and so we also deal with the interval $[\tau^4, 7]$ by largely arithmetic means. However, the analysis of $c(a)$ for $a > \tau^4$ gets easier the larger a is. As we show in Corollary 1.2.4, it is almost trivial to see that $c(a) = \sqrt{a}$ when $a \geq 9$, and it is not much harder to see that $c(a) = \sqrt{a}$ when $a \geq 8\frac{1}{36}$. The method used also shows that there are finitely many obstructions when $a \geq 7$. The full analysis of $c(a)$ on $[7, 8\frac{1}{36}]$ takes more effort. The intervals I_j contain rational numbers with small denominators; for example the three longest contain $7, 7\frac{1}{2}$, and 8, cf. Figure 1.2.
We refer to Theorem 5.2.3 for a full description of $c(a)$ on the interval $[7, 8\frac{1}{36}]$. One point to note here is that although all the flatter portions of the graph of c are horizontal when $a < \tau^4$, this is not true when $a \in [7, 8]$; for example the two parts of the graph of c centered at $a = 7\frac{1}{8}$ both have positive slope.

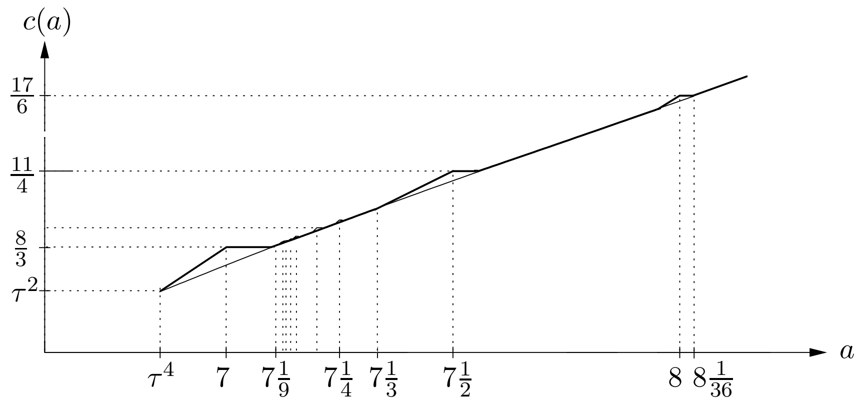

FIGURE 1.2. The graph of $c(a)$ on $\left[\tau^4, 8\frac{1}{36}\right]$.

Connection with counting lattice points

As we explain in §1.3 below, the obstructions to embeddings $E(1,a) \overset{s}{\hookrightarrow} B(\mu)$ that we consider come from exceptional spheres in blow ups of $\mathbb{C}P^2$. Hofer[2] suggested that one should also be able to obtain a complete set of obstructions from the embedded contact homology theory recently developed by Hutchings and Taubes [12]. The embedded contact homology $ECH_*\big(E(a,b)\big)$ of a 4-dimensional ellipsoid has one generator in each even degree with action of the form $ma + nb$; $m, n \geq 0$. Since the action is a nondecreasing function of degree, the actions of the generators arranged in the order of increasing degree form the sequence $N(a,b)$ obtained by arranging all numbers of the form $ma+nb$; $m,n \geq 0$, in nondecreasing order (with multiplicities). We will say that $N(a,b) \preccurlyeq N(a',b')$ if each term in $N(a,b)$ is no greater than the corresponding term in $N(a',b')$. Using work by Taubes, Hutchings shows in [10] that the sequence $N(a,b)$ is a monotone invariant of $E(a,b)$:

$$E(a,b) \overset{s}{\hookrightarrow} E(a',b') \implies N(a,b) \preccurlyeq N(a',b'). \tag{1.1.4}$$

Even before this was proven, Hofer suggested that

$$E(a,b) \overset{s}{\hookrightarrow} E(a',b') \iff N(a,b) \preccurlyeq N(a',b'). \tag{C}$$

One might be able to prove this directly by showing that the embedded curves that provide the morphism $ECH_*\big(E(a,b)\big) \to ECH_*\big(E(a',b')\big)$ correspond precisely to the exceptional spheres that give our obstructions. However, so far all approaches have been more indirect.

For $a \geq 1$ define

$$c_{ECH}(a) := \inf \{\mu > 0 \mid N(1,a) \preccurlyeq N(\mu,\mu)\}.$$

As we show in § 2.4, the function c_{ECH} can be understood in terms of counting lattice points in triangles. Then Hutchings's result (1.1.4) for the case that the target ellipsoid

[2]Private communication.

is a ball becomes

$$c_{ECH}(a) \le c(a) \quad \text{for all } a \ge 1.$$

In this paper we prove the converse.

Theorem 1.1.3. $c_{ECH}(a) \ge c(a)$ *for all* $a \ge 1$.

Thus Hofer's conjecture (C) holds if the target is a ball: $c_{ECH}(a) = c(a)$ for all a.

Remark 1.1.4. (i) The methods used to analyze the embedding of a 4-dimensional ellipsoid into a ball work equally well when one considers embeddings from one ellipsoid to another. In other words, Theorem 1.2.2 below has an analog that is applicable to this setting; see [16, Theorem 1.5]. One can also use much the same method to analyze embeddings of an ellipsoid into $S^2 \times S^2$; see [20].

(ii) One might wonder if these results can be extended to higher dimensions. For example, in dimension 6 is there a symplectic embedding $E(a,b,c) \overset{s}{\hookrightarrow} E(a',b',c')$ if and only if $N(a,b,c) \preccurlyeq N(a',b',c')$? Guth's construction in [7] of an embedding $E(1,R,R) \overset{s}{\hookrightarrow} E(2,10,2R^2)$ for *all* $R > 1$ shows that the answer is no. It is not at present clear what the correct condition should be in this case. (See Hind–Kerman [9] for a more precise version of Guth's result, and Buse–Hind [4] for some further results.) Note that embedded contact homology is a specifically 4-dimensional theory, as are the results stated in Theorem 1.2.2 and Proposition 1.2.12 below on which our calculation of c is based.

(iii) There are two early papers by Biran with constructions that are somewhat similar to ours. In [2] he uses an iterated ball packing construction to obtain information on the Kähler cone of blow ups of $\mathbb{C}P^2$. Continued fractions are relevant here, but Biran does not use them in the way we do. The survey article [3] mentions how an understanding of embeddings of ellipsoids might help calculate this Kähler cone, and hence suggests a potential application of our work. However, for this one would need to understand which embeddings of ellipsoids give rise to Kähler forms, a question that we do not consider.

(iv) Meanwhile, the full Hofer conjecture (C) has been proven by McDuff [18], without using embedded contact homology; see also Hutchings [11]. ◇

1.2. **Method of proof.** The first author showed in [16] that if $a \ge 1$ is rational there is a finite sequence $\boldsymbol{w}(a) := (w_1, \dots, w_M)$ of rational numbers such that the ellipsoid $E(1,a)$ embeds symplectically in the ball $B(\mu)$ exactly if the corresponding collection $\sqcup_i B(w_i)$ of M disjoint balls embeds symplectically in $B(\mu)$. This ball embedding problem was reduced in McDuff–Polterovich [19] to the question of understanding the symplectic cone of the M-fold blow up of $\mathbb{C}P^2$. After further work by McDuff [15] and Biran [1], the structure of this cone was finally elucidated in Li–Liu [14] and Li–Li [13].

The key to understanding this cone is the following set $\mathcal{E}_M$.

Definition 1.2.1. *Denote by* X_M *the* M*-fold blow up of* $\mathbb{C}P^2$ *with any symplectic structure* ω_M *obtained by blow-up from the standard structure on* $\mathbb{C}P^2$*. Let* $L, E_1, \dots, E_M \in$

$H_2(X_M)$ be the homology classes of the line and the M exceptional divisors. We define $\mathcal{E}_M$ to be the set consisting of $(0;-1,0,\dots,0)$ and of all tuples $(d;\boldsymbol{m})$ of nonnegative integers $(d;m_1,\dots,m_M)$ with $m_1\ge\dots\ge m_M$ and such that the class $E_{(d;\boldsymbol{m})}:=dL-\sum_i m_iE_i$ is represented in (X_M,ω_M) by a symplectically embedded sphere of self-intersection -1. If there is no danger of confusion, we will write $\mathcal{E}$ instead of $\mathcal{E}_M$. Clearly, $\mathcal{E}_M\subset\mathcal{E}_{M'}$ whenever $M\le M'$.

Since these classes $E_{(d;\boldsymbol{m})},(d;\boldsymbol{m})\in\mathcal{E}_M$, have nontrivial Gromov invariant, they have symplectically embedded representatives for all choices of the blow-up form ω_M. Therefore the above definition does not depend on this choice.

Denote by $-K:=3L-\sum E_i$ the standard anti-canonical divisor in X_M, and consider the corresponding symplectic cone $\mathcal{C}_K$, consisting of all classes on X_M that may be represented by a symplectic form with first Chern class Poincaré dual to $-K$. Then Li–Li show in [13] that

$$\mathcal{C}_K=\left\{\alpha\in H^2(X_M)\mid\alpha^2>0,\ \ \alpha(E)>0\ \text{ for all } E\in\mathcal{E}_M\right\}.$$

Proposition 1.2.12 below gives necessary and sufficient conditions for an element $(d;\boldsymbol{m})$ to belong to $\mathcal{E}_M$. Before discussing this, we explain the relevance of $\mathcal{E}_M$ to our problem. The following result is proved in [16]. We will denote by $\ell,e_i\in H^2(X_M)$ the Poincaré duals to L,E_i and by $\boldsymbol{m}\cdot\boldsymbol{w}=\sum_{i=1}^M m_iw_i$ the Euclidean scalar product in $\mathbb{R}^M$.

Theorem 1.2.2. *For each rational $a\ge1$ there is a finite weight expansion $\boldsymbol{w}(a)=(w_1,\dots,w_M)$ such that $E(1,a)$ embeds symplectically in the interior of $B^4(\mu)$ if and only if $\mu\ell-\sum w_ie_i\in\mathcal{C}_K$. Moreover $w_i\le1$ for all i and $\sum_i w_i^2=a$.*

Corollary 1.2.3. *If the rational number $a\ge1$ has weight expansion $\boldsymbol{w}(a)=\boldsymbol{w}=(w_i)$, then*

$$c(a)\;=\;\sup\Big(\sqrt{a},\ \mu(d;\boldsymbol{m})(a)\mid(d;\boldsymbol{m})\in\mathcal{E}\Big),$$

where $\mu(d;\boldsymbol{m})(a):=\frac{\boldsymbol{m}\cdot\boldsymbol{w}(a)}{d}$.

Proof. The above description of $\mathcal{C}_K$ shows that $E(1,a)$ embeds into the interior of $B^4(\mu)$ if and only if the tuple $(\mu,\boldsymbol{w})$ satisfies the conditions

(i) $\mu^2>\boldsymbol{w}\cdot\boldsymbol{w}=:\sum w_i^2$,

(ii) $d\mu>\boldsymbol{m}\cdot\boldsymbol{w}=:\sum m_iw_i$ for all $(d;\boldsymbol{m})\in\mathcal{E}_M$.

The corollary now follows because $\boldsymbol{w}\cdot\boldsymbol{w}=a$. □

Biran showed in [1] that $\mu(d;\boldsymbol{m})(k)\le\sqrt{k}$ for all $(d;\boldsymbol{m})\in\mathcal{E}$ for all integers $k\ge9$. His argument extends to all $a\ge9$ and shows:

Corollary 1.2.4. *$c(a)=\sqrt{a}$ when $a\ge9$.*

Proof. Since c is continuous by Lemma 1.1.1, it suffices to check this for rational a. Fix $(d;\boldsymbol{m})\in\mathcal{E}$. The corresponding symplectically embedded (-1)-sphere E has $c_1(E)=1$, and so $3d-1=\sum_i m_i$. Therefore, $\sum_i m_iw_i\le\sum_i m_i=3d-1$. For $a\ge9$ we thus find

$$\mu(d;\boldsymbol{m})(a)\;:=\;\frac{\boldsymbol{m}\cdot\boldsymbol{w}}{d}\;<\;3\;\le\;\sqrt{a}.$$

Now use Corollary 1.2.3. □

In view of Corollary 1.2.3, our task is two-fold; first to understand the weight expansions and then to understand the restrictions placed on embeddings by the elements of $\mathcal{E}_M$. The description that we now give for $\boldsymbol{w}(a)$ is convenient for calculations but is somewhat different from that in [16]. The equivalence of the two definitions is established in Corollary A.10.

Definition 1.2.5. *Let $a = p/q \in \mathbb{Q}$ written in lowest terms. The* **weight expansion** $\boldsymbol{w} := (w_i) := (w_1, \dots, w_M)$ *of $a \geq 1$ is defined recursively as follows:*

- $w_1 = 1$, *and* $w_n \geq w_{n+1} > 0$ *for all* n;
- *if* $w_i > w_{i+1} = \dots = w_n$ *(where we set* $w_0 := a$*), then*

$$w_{n+1} = \begin{cases} w_n & \text{if } w_{i+1} + \dots + w_{n+1} = (n-i+1)w_{i+1} \leq w_i \\ w_i - (n-i)w_{i+1} & \text{otherwise;} \end{cases}$$

- *the sequence stops at* w_n *if the above formula gives* $w_{n+1} = 0$.

The number M of entries in $\boldsymbol{w}(a)$ is called the **length** $\ell(a)$ *of* a.

For example, $a = 25/9$ has weight expansion $\boldsymbol{w}(a) = (1, 1, \frac{7}{9}, \frac{2}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9})$, which we will abbreviate as $(1^{\times 2}, \frac{7}{9}, \frac{2}{9}^{\times 3}, \frac{1}{9}^{\times 2})$.

We may also think of this expansion (w_i) as consisting of $N+1$ blocks of length ℓ_s of the (decreasing) numbers x_s where $x_0 = 1$; viz:

$$\begin{aligned} \boldsymbol{w}(a) \;&:=\; \big(\underbrace{1, \dots, 1}_{\ell_0}, \underbrace{x_1, \dots, x_1}_{\ell_1}, \dots, \underbrace{x_N, \dots, x_N}_{\ell_N}\big) \\ &= \; \big(1^{\times \ell_0}, x_1^{\times \ell_1}, \dots, x_N^{\times \ell_N}\big). \end{aligned} \tag{1.2.1}$$

Then $x_1 = a - \ell_0 < 1$, $x_2 = 1 - \ell_1 x_1 < x_1$, and so on. In this form the sequence can be generated as follows. If $a = \frac{p}{q}$, first draw a rectangle of length p and height q, then mark off as many (say ℓ_0) squares of side length q as possible, then in the remaining rectangle of size $q \times (p - \ell_0 q)$ mark off as many (say ℓ_1) squares of side length $(p - \ell_0 q)$ as possible, continuing in this way until the rectangle is completely filled. Then $q x_j$ is the side length of the $(j+1)$st set of squares, while the ℓ_j are the multiplicities: see Figure 1.3. As is well known, the multiplicities $\ell_j, 0 \leq j \leq N$, give the continued fraction expansion $[\ell_0; \ell_1, \dots, \ell_N]$ of p/q. For example, $25/9 = [2; 1, 3, 2]$ and

$$\frac{25}{9} = 2 + \frac{1}{1 + \frac{1}{3 + \frac{1}{2}}}.$$

Notice also that $25/9 = 2 \cdot 1^2 + (7/9)^2 + 3(2/9)^2 + 2(1/9)^2$.

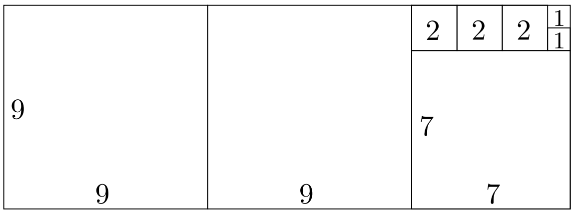

FIGURE 1.3. The expansion for $a = 25/9$.

Lemma 1.2.6. *Let $\boldsymbol{w} := (w_1, \dots, w_M)$ be the weight expansion of $a = \frac{p}{q} \geq 1$. Then*

$$w_M = \tfrac{1}{q},$$

$$\boldsymbol{w} \cdot \boldsymbol{w} := \sum_{i=1}^{M} w_i^2 = a, \tag{1.2.2}$$

$$\sum_{i=1}^{M} w_i = a + 1 - \tfrac{1}{q}. \tag{1.2.3}$$

Proof. Equation (1.2.2) holds because the total area of all the squares is pq. To understand the sum, suppose that there are $N+1$ sets of squares in the expansion (1.2.1) so that $x_{N+1} = 0$ and write

$$\begin{aligned}\sum_i w_i &= 1 + \Big(\underbrace{1 + \dots + 1}_{\ell_0 - 1} + x_1\Big) + \Big(\underbrace{x_1 + \dots + x_1}_{\ell_1 - 1} + x_2\Big) + \dots \\ &\qquad + \Big(\underbrace{x_N + \dots + x_N}_{\ell_N - 1} + x_{N+1}\Big) \\ &= 1 + (a-1) + (1 - x_1) + \dots + (x_{N-1} - x_N) \\ &= 1 + a - x_N.\end{aligned}$$

It remains to note that $x_N = w_M = 1/q$. This is obvious from the geometric construction. For, $qx_N = qw_M$ is the side length of the smallest square in the decomposition of the rectangle. If this length were divisible by s, then the side lengths of all the squares would be divisible by s. Hence both p and q would be divisible by s. But they are mutually prime by hypothesis. □

We next describe the sets $\mathcal{E}_M$. The first lemma is well known, and can be easily deduced from Proposition 1.2.12 below.

Lemma 1.2.7. *The set $\mathcal{E}_M$ is finite for $M \leq 8$ with elements $(d; m_1, \dots, m_M)$ equal to:*

$$(0; -1),\ \ (1; 1, 1),\ \ (2; 1^{\times 5}),\ \ (3; 2, 1^{\times 6}),$$
$$(4; 2^{\times 3}, 1^{\times 5}),\ \ (5; 2^{\times 6}, 1, 1),\ \ (6; 3, 2^{\times 7}).$$

From this one can immediately calculate $c(a)$ for those a whose weight expansion has $M \le 8$.

Corollary 1.2.8. *The function c takes the following values:*

$$c(2) = c(3) = c(4) = 2, \quad c(5) = c(6) = \tfrac{5}{2},$$
$$c\big(\tfrac{13}{2}\big) = \tfrac{13}{5}, \quad c(7) = \tfrac{8}{3}, \quad c(8) = \tfrac{17}{6}.$$

Moreover, its graph is linear on each subinterval $[1, 2]$, $[2, 4]$, $[4, 5]$, $[5, 6]$.

Proof. The values of $c(a)$ for integers $a \in [1, 8]$ were calculated in [16, Cor 1.2]. One can similarly calculate $c(\frac{13}{2})$ since the length of $\boldsymbol{w}\big(\frac{13}{2}\big)$ is < 9. The second statement then follows from Lemma 1.1.1. □

From Lemma 1.2.7 we can also compute c near 7.

Proposition 1.2.9. *For $a \in [6\frac{11}{12}, 7]$ we have $c(a) = \frac{1}{3}(a+1)$. Also $c(a) = \frac{8}{3}$ for $a \in [7, 7\frac{1}{9}]$.*

Proof. Since c is continuous, it suffices to prove these identities for $a \in \mathbb{Q}$. First assume that $a < 7$ and write $a = 6 + x$. Then

$$\boldsymbol{w}(a) \;=\; \big(1^{\times 6}, x, w_8, \dots, w_M\big)\,,$$

where $0 < w_i < 1 - x$ for $i \ge 8$. The element $(3; 2, 1^{\times 6}) \in \mathcal{E}_7$ gives the constraint $c(a) \ge \mu_0 = \frac{1}{3}(a+1)$.

Since $1 - x \le x/9$, at least the first 9 of the weights $w_8, w_9, \dots$ are equal. Hence, by Corollary 1.2.4, we can fully pack all but the first 7 balls into one ball of size λ where $a = 6 + x^2 + \lambda^2$. It remains to show that the 8 balls of sizes

$$W = (1, \dots, 1, x, \lambda)$$

fit into $B(\mu_0)$, that is, $W \cdot \boldsymbol{m} \le \frac{d}{3}(7 + x)$ for all $(d; \boldsymbol{m}) \in \mathcal{E}$.

This is clear for classes in $\mathcal{E}_7$. The classes in $\mathcal{E}_8 \smallsetminus \mathcal{E}_7$ are $(4; 2^{\times 3}, 1^{\times 5})$, $(5; 2^{\times 6}, 1, 1)$, $(6; 3, 2^{\times 7})$. The strongest constraint comes from $(5; 2^{\times 6}, 1, 1)$ and equals

$$\mu_1 \;=\; \tfrac{1}{5}\big(12 + x + \lambda\big).$$

The desired inequality $\mu_1 \le \mu_0$ is equivalent to $1 + 3\lambda \le 2x$. Since $\lambda^2 = x(1-x)$ we need $13x^2 - 13x + 1 \ge 0$, which is satisfied when $x \ge \frac{11}{12}$.

We know $c(7) = \frac{8}{3}$ by Corollary 1.2.8. Therefore it suffices to show that $c(7\frac{1}{9}) = \frac{8}{3}$. As above, since the nine balls $B(\frac{1}{9})$ fully fill $B(\frac{1}{3})$, we just need to check that the finite number of elements in $\mathcal{E}_8$ give no obstruction to embedding 8 balls, seven of size 1 and one of size $\frac{1}{3}$, into $B(\frac{8}{3})$. □

Remark 1.2.10. Similarly, there is an obstruction at $a = 8$ given by the class $(d; \boldsymbol{m}) = (6; 3, 2^{\times 7})$. For $a < 8$ with $\boldsymbol{w}(a) = (1^{\times 7}, a-7, \dots)$ this gives the constraint $\mu(a) = \frac{1+2a}{6}$, while for $a \ge 8$ we get $\mu(a) = \frac{17}{6}$. Therefore $c(a) \ge \mu(a) \ge \sqrt{a}$ for $\frac{8+3\sqrt{7}}{2} \le a \le 8\frac{1}{36}$. However, unfortunately, one cannot argue as above to show that $c = \mu$ on some interval $(8 - \varepsilon, 8\frac{1}{36}]$ because the auxiliary packings would involve 9 balls and $\mathcal{E}_9$ is infinite. We shall prove that $c = \mu$ on $\big[\frac{8+3\sqrt{7}}{2}, 8\frac{1}{36}\big]$ in Sections 5.2 and 5.3 by different methods. ◇

Now consider $\mathcal{E}_M$, $M \geq 9$. We say that a tuple of integers $(d; \boldsymbol{m}) = (d; m_1, \dots, m_M)$ is *ordered* if $m_i \geq m_{i+1}$ when $m_i \neq 0$, $m_{i+1} \neq 0$, and if the m_i with $m_i = 0$ are at the end. For instance, the elements of $\mathcal{E}_M$ are ordered in view of Definition 1.2.1. To characterize $\mathcal{E}_M$ when $M \geq 9$ we need the following definition.

Definition 1.2.11. *The* **Cremona transform** *of an ordered tuple* $(d; \boldsymbol{m})$ *is*

$$(2d - m_1 - m_2 - m_3;\ d - m_2 - m_3,\ d - m_1 - m_3,\ d - m_1 - m_2,\ m_4,\ m_5,\ \dots).$$

A **standard Cremona move** $\mathcal{C}r$ *takes an ordered tuple* $(d; \boldsymbol{m})$ *to the tuple obtained by ordering the Cremona transform of* $(d; \boldsymbol{m})$*. More generally, a* **Cremona move** *is the composite of a Cremona transform with any permutation of* $\boldsymbol{m}$*.*

Standard Cremona moves preserve $\mathcal{E}_M$ because they are achieved by Cremona transformations, which (modulo permutations of the E_i) are just reflections $A \mapsto A + (A \cdot C)\, C$ in the (-2)-sphere in the class $C := L - E_1 - E_2 - E_3$; cf. [13]. [3] In particular these moves preserve the intersection product and the first Chern class $c_1(M) := 3L - \sum_i E_i$.

Proposition 1.2.12. (i) *The following identities hold for all* $(d; \boldsymbol{m}) \in \mathcal{E}_M$*.*

$$\sum_i m_i = 3d - 1, \qquad \boldsymbol{m} \cdot \boldsymbol{m} := \sum_i m_i^2 = d^2 + 1. \tag{1.2.4}$$

(ii) *For all pairs* $(d; \boldsymbol{m}), (d'; \boldsymbol{m}')$ *of distinct elements of* $\mathcal{E}_M$ *we have*

$$\boldsymbol{m} \cdot \boldsymbol{m}' := \sum_i m_i\, m_i' \ \leq\ d\, d'.$$

(iii) *A tuple* $(d; \boldsymbol{m})$ *satisfying the Diophantine conditions in Equation* (1.2.4) *belongs to* $\mathcal{E}_M$ *exactly if it may be reduced to* $(0; -1, 0, \dots, 0)$ *by repeated standard Cremona moves.*

Proof. The two equations in (i) express the fact that any symplectically embedded (-1)-sphere E has $c_1(E) = 1$ and $E \cdot E = -1$. Since the elements in $\mathcal{E}_M$ are all represented by embedded J-holomorphic spheres for generic J, part (ii) holds by positivity of intersections.

Part (iii) for a class $(d; \boldsymbol{m})$ with $d = 0$ is clear. Part (iii) for non-negative tuples $(d; \boldsymbol{m})$ may be deduced from Li–Li's arguments in [13, Lemma 3.4]. In this paper the authors work in a more general context than ours, considering *all* symplectic forms on X_M, while we consider only those symplectic forms with the standard first Chern class (or anticanonical class) $-K := 3L - \sum E_i$. They also introduce many ideas, such as the symplectic genus. However, the concept relevant here is that of a *reduced* class. This is a class $A := dL - \sum m_i E_i$ with

$$d > 0, \quad m_1 \geq m_2 \geq \dots \geq 0, \quad d \geq m_1 + m_2 + m_3.$$

We can clearly assume that $M \geq 3$. In this case they show that for every class A with $A^2 = -1$ and $c_1(A) > 0$ there is a combination of Cremona moves and reflections

[3] If ω is a symplectic form on X_M for which the class C is represented by a Lagrangian sphere S_L, then the Cremona transformation can be realized by the Dehn twist in S_L; cf. Seidel [23].

$E_i \mapsto -E_i$ that transform A either into E_1 (which corresponds to $(0; -1, 0, \dots, 0)$) or into a reduced class A'. Each step consists of a Cremona transform, followed by an adjustment of signs to make the coefficients of the $-E_i$ non-negative and then a permutation to reorder the m_i. Since the Cremona transform takes $(d; \boldsymbol{m})$ to

$$(2d - m_1 - m_2 - m_3;\ d - m_2 - m_3,\ d - m_1 - m_3,\ d - m_1 - m_2,\ m_4,\ \dots)$$

this decreases d unless A is reduced. Li–Li show by a simple algebraic computation that, because we start with a class with $A \cdot A \geq -1$ and the reduction process preserves the intersection form, the coefficient d cannot become negative. Thus one stops the process when the coefficient d is at its minimum. If $d = 0$ then the final class A_0 is $(0; -1, 0 \dots, 0)$. On the other hand, if A_0 is reduced, another essentially algebraic argument (part 3 of their Lemma 3.4) shows that $A_0 \cdot E \geq 0$ for all solutions E to the equations (1.2.4). Hence $A_0 \notin \mathcal{E}_M$ since $A_0 \cdot A_0 = -1$.

It is easy to adapt the results of this lemma to our situation. We are interested here only in symplectic forms with the standard canonical class K, and therefore cannot change the signs of the E_i. However, if there is a sequence σ of standard Cremona moves that takes a tuple $(d; \boldsymbol{m})$ that satisfies (1.2.4) to a tuple $(d'; \boldsymbol{m}')$ with some $m_i' < 0$, then either $d' = 0$ and $(d'; \boldsymbol{m}') = (0; -1, 0, \dots, 0)$, or $d' \neq 0$. But in the latter case $(d; \boldsymbol{m})$ cannot be in $\mathcal{E}_M$ since we would have

$$E_{(d;\boldsymbol{m})} \cdot E \;=\; E_{(d';\boldsymbol{m}')} \cdot E_i \;=\; m_i' \;<\; 0,$$

where E is the image of E_i under the reverse sequence σ^{-1} of Cremona moves. Therefore (iii) must hold. □

Definition 1.2.13. *A class $E = (d; \boldsymbol{m}) \in \mathcal{E}$ is called* **obstructive** *if $\mu(d; \boldsymbol{m})(z) > \sqrt{z}$ on some nonempty interval I. Further we say that E is* **obstructive at** *a if $\mu(d; \boldsymbol{m})(a) > \sqrt{a}$.*

Thus our task is to understand enough about the obstructive classes to figure out the supremum of the corresponding constraint functions $\mu(d; \boldsymbol{m})$.

Remark 1.2.14. (i) Later we will expend considerable effort to show that certain classes $E = (d; \boldsymbol{m})$ that satisfy the identities (1.2.4) do in fact lie in $\mathcal{E}_M$. In some cases, the corresponding constraints $\mu(d; \boldsymbol{m})(a)$ contribute to $c(a)$. However, in many other cases (for example the classes $E\big(b_k(i)\big)$ of Proposition 4.2.2 for $i \geq 3$, see Lemma 4.3.1) the constraint $\mu(d; \boldsymbol{m})(a)$ does not contribute to $c(a)$; rather E influences $c(a)$ because $E \cdot E' \geq 0$ for all $E' \in \mathcal{E}_M \smallsetminus E$. For this positivity of intersections to hold, it is not necessary that $E \in \mathcal{E}_M$. As explained in the proof of Proposition 1.2.12, it suffices that when we apply standard Cremona moves to E we do not arrive at a class with $d > 0$ and some $m_i < 0$, but instead end up at a reduced class. (By [13, Lemma 3.6], the class E then must have positive symplectic genus, and therefore cannot be represented by a smoothly embedded sphere.) However, it is just as difficult to check this condition as it is to check whether $E \in \mathcal{E}_M$, and, in fact, it turns out that $E \in \mathcal{E}_M$ in all cases of interest to us here.

(ii) As we show in Proposition 5.2.1, there are only finitely many tuples $(d;\boldsymbol{m}) \in \mathcal{E}_M$ that are obstructive at some $a \geq 7$. In fact, there are precisely 13 such classes; the class $\left(3;2,1^{\times 6}\right)$ centered at $a=7$, another 8 classes that contribute to $c(a)$ as described in Theorem 5.2.3, and 4 more classes listed in Lemma 5.2.5, that are "hidden" in the sense that they contribute nothing new to $c(a)$. Although we work mostly by hand, we do use the computer programs of Appendix B to prove Corollary 5.2.10, which states that there are no other relevant classes.

In contrast, there are infinitely many classes that are obstructive somewhere on the interval $[1,7]$, and we do not try to compute them all. As shown in Example 2.3.1, the part of the graph of an obstruction $\mu(d;\boldsymbol{m})(a)$ that lies above $\sqrt{a}$ can be quite complicated and need not have the scaling or positivity properties of c that are described in Lemma 1.1.1. Further, even though Corollary 2.1.4 states that at each point a where $c(a) > \sqrt{a}$ there are only finitely many obstructive classes with $\mu(d;\boldsymbol{m})(a) = c(a)$, we do not know if for some a_0 there are infinitely many classes with $\mu(d;\boldsymbol{m})(a_0) > \sqrt{a_0}$. By Remark 5.2.2 this cannot happen when $a_0 > \tau^4$. When $a_0 = \tau^4$, Proposition 4.3.2 shows that there are infinitely many classes that are obstructive on an interval whose closure contains a_0. However, because $c(\tau^4) = \tau^2$ no class is obstructive at τ^4 itself. We have no relevant results when $a_0 < \tau^4$.

(iii) We compute $c(a)$ for $a \leq 7$ by looking not only at classes with $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ but also at some other classes that influence $c(a)$ indirectly. The most interesting of these are the classes described in Proposition 4.3.2 that are made from the even terms of the Fibonacci sequence. They play a dual role. Though obstructive, they contribute nothing new to $c(a)$ and so the corresponding graph is called the *ghost stairs*. Their importance is rather that they allow one to calculate $c(a)$ at a series of points e_k where they are not obstructive; cf. the proof of Corollary 4.2.4. $\diamond$

1.3. **Outline of paper.** Corollary 1.2.3 gives a formula for $c(a)$ that we can interpret using the description of $\mathcal{E}$ contained in Proposition 1.2.12. However, this formula is not at all explicit, and the methods needed to understand it depend on the size of a. One can compute $c(a)$ by direct methods when $a < \tau^4$, but for larger a our arguments require a deeper understanding of the constraint functions $\mu(d;\boldsymbol{m})(a)$. Therefore we begin in Section 2 by developing some tools to distinguish the obstructive classes in $\mathcal{E}$.

First, we show in Proposition 2.1.1 that $(d;\boldsymbol{m})$ is obstructive at a only if the vector $\boldsymbol{m}$ is almost parallel to $\boldsymbol{w}(a)$. If $\boldsymbol{m}$ is parallel to $\boldsymbol{w}(a)$ for some a, then we call $(d;\boldsymbol{m})$ a *perfect* obstruction at a. Lemma 2.1.5 shows that these elements determine the function $c(z)$ for z near a, while Corollary 3.1.3 shows that the only perfect obstructions occur at the numbers b_n of the Fibonacci stairs. Second, we show in Lemma 2.1.3 that if $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ on the interval I, then I has a unique central point a_0 distinguished by the fact that $\ell(a_0) = \ell(\boldsymbol{m})$ while $\ell(a) > \ell(\boldsymbol{m})$ for all other $a \in I$. (Here, $\ell(\boldsymbol{m})$ denotes the number of positive entries in $\boldsymbol{m}$.) Lemmas 2.1.7 and 2.1.8 describe other useful properties of obstructive classes.

These are the basic results needed to determine $c(a)$ when $a \geq 7$. (Since by Proposition 5.2.1 there are only finitely many obstructive $(d;\boldsymbol{m})$ for $a \geq 7$, we can analyze these on a case by case basis, without using more general results.) However, we continue in Section 2 with a deeper analysis of the functions $\mu(d;\boldsymbol{m})$, so that we can explain the relation of $c(a)$ to the lattice point counting problem. This analysis is based on Proposition 2.2.6 which derives surprising identities satisfied by weight expansions. (These are quadratic identities involving the weight expansions of a and its "mirror" $\overleftarrow{a}$.) This proposition also turns out to be helpful in understanding $c(a)$ on $[\tau^4, 7]$, where there are infinitely many obstructive classes.

Our next main result is Proposition 2.3.2 which shows that the central point of I is the break point of $\mu(d;\boldsymbol{m})$ in the sense that $\mu(d;\boldsymbol{m})$ is linear on each component of $I \setminus \{a_0\}$. Moreover, one can apply Proposition 2.2.6 to show that the coefficients of these linear functions are remarkably close to those of the linear functions that occur in the counting problem. In §2.4, we explain this connection and prove Theorem 1.1.3 (which states that $c_{ECH} \geq c$).

In Section 3 we calculate $c(a)$ for $a \in [1, \tau^4]$ by direct methods. Theorem 3.1.1 states that there are classes $E(a_n)$ and $E(b_n)$ in $\mathcal{E}$ given by tuples $(d;\boldsymbol{m})$ constructed from the weight expansions of ratios of odd Fibonacci numbers. As we see in Corollary 3.1.2, because these classes are perfect, their very existence together with the scaling property of c is enough to calculate $c(a)$ in this range. The difficulty here is to prove that these classes really do belong to $\mathcal{E}$. In particular, describing what happens to these classes under Cremona moves involves establishing many quadratic identities for Fibonacci numbers. Therefore in Section 3.2 we develop an inductive method to prove such identities; cf. Proposition 3.2.3. The proof of Theorem 3.1.1 is completed in §3.3. This section is essentially independent of Section 2.

We next compute $c(a)$ on $[\tau^4, 7]$. The obstruction $(3; 2, 1^{\times 6})$ centered at 7 gives the lower bound $c(a) \geq \frac{a+1}{3}$ on this interval, and our task in Section 4 is to show that no obstruction exceeds this one. One difficulty is that the quantity $y(a) := a + 1 - 3\sqrt{a}$ tends to 0 as a approaches τ^4, permitting the existence of infinitely many obstructive classes; cf. Proposition 4.3.2. Another is that the line $\frac{a+1}{3}$ does not pass through the origin. Therefore we can no longer use the scaling property of c, which in the case of the interval $[1, \tau^4]$ allowed us to restrict attention to the points a_n, b_n. Nevertheless, by using the results of Section 2 we show in Proposition 4.1.6 that there is only a double sequence of relevant points.

One could then attempt a direct calculation of $c(a)$ at these points, combining more elaborate versions of the estimation techniques used in Section 5 with arithmetic results based on Corollary 2.2.7. This is possible. However it is very complicated and it turns out that there is a much easier proof. The sequence $E(b_n)$ of perfect classes that determine the Fibonacci stairs really consists of two subsequences $E_k(0)$ and $E_k(1)$ that are the first two members of an infinite family $E_k(i)$, $i \geq 0$, of sequences of "nearly perfect" classes in $\mathcal{E}$. The classes $E_k(2)$, $k \geq 1$, form the ghost stairs discussed in Remark 1.2.14(iii), while the classes $E_k(i)$, $i > 2$, are not obstructive by Lemma 4.3.1. Nevertheless, as we show in Lemma 4.1.10, the fact that they are nearly perfect puts constraints on the possible obstructive classes. The desired conclusion follows by combining this result with Proposition 4.1.6.

Section 5 carries out a detailed analysis of the obstructions in the interval $[7, 9]$. The argument is based on the equality in Proposition 2.1.1 (iv). This estimates the "error" (the difference between $\boldsymbol{m}$ and a suitable multiple of $\boldsymbol{w}(a)$) at a rational point $a = p/q$ in terms of the quantity $y(a)-1/q$, where again $y(a) = a+1-3\sqrt{a}$. Since $y(\tau^4) = 0$, this estimate gets better the further a is from τ^4 and the larger q is. One easy consequence is Proposition 5.2.1, stating that there are only finitely many obstructive classes $(d; \boldsymbol{m})$ for $a \geq 7$. The proof (in §5.2) that $c(a) = \sqrt{a}$ for $a \geq 8\frac{1}{36}$ is also easy.

To work out exactly what the constraints are requires some computation. It would be possible, though very tedious, to do this entirely by hand. We have aimed to use the computer as little as possible, and so have developed quite a few techniques for estimating the error. Since $y(a) - \frac{1}{q}$ is negative when $a = 7\frac{1}{k}$, we must treat these points separately, by purely arithmetic means. Thus in this case we simply look for suitable solutions $(d; \boldsymbol{m})$ of the Diophantine equations (1.2.4) with centers at these points, using Lemma 2.1.7 to limit possibilities. This computation (in Lemma 5.2.5) finds several classes that contribute to $c(a)$ as well as some interesting "hidden" classes for which $\mu(d; \boldsymbol{m})(a) = c(a)$ at just one point, namely the center. There are some other obstructive classes centered at points of the form $a = 7\frac{2}{2k+1}$. (The table in Theorem 5.2.3 lists all classes that contribute to $c(a)$.) We show that there are no other obstructive classes in §5.2 using estimates developed in §5.1 as well as two computer programs that are described in Appendix B.

Finally, Appendix A explains the connection between our current definition of the weight expansion of a in terms of the continued fraction expansion of a and the definition used in [16], which came from a blow up construction. The results here are no doubt well known; we included them for the sake of completeness.

Acknowledgments. We wish to thank Dylan Thurston for making some very helpful observations at the beginning of this project (he was the first person to point out a connection with Fibonacci numbers); Michael Hutchings for explaining the obstructions coming from embedded contact homology; and Helmut Hofer and Peter Sarnak for their inspiration and encouragement.

2. Foundations

2.1. **Basic observations.** Given a with weight expansion $\boldsymbol{w}(a)$ of length $\ell(a) = M$ and $(d; \boldsymbol{m}) \in \mathcal{E}_M$, we define $\varepsilon := \varepsilon(a) = (\varepsilon_1, \dots, \varepsilon_M)$ by setting

$$\boldsymbol{m} = \frac{d}{\sqrt{a}}\boldsymbol{w}(a) + \varepsilon. \tag{2.1.1}$$

We will refer to the vector ε as the *error*, and to quantities such as $\sum \varepsilon_i^2$ as the *squared error*. We need to understand the function $\mu(d; \boldsymbol{m})(a) = \boldsymbol{m} \cdot \boldsymbol{w}(a)/d$ defined in Corollary 1.2.3.

Our arguments will be based on the following observations.

Proposition 2.1.1. *For all* $(d; \boldsymbol{m}) \in \mathcal{E}$ *and* a*, we have*

(i) $\mu(d;\boldsymbol{m}) := \mu(d;\boldsymbol{m})(a) \;\le\; \sqrt{a}\sqrt{1+1/d^2}$;

(ii) $\mu(d;\boldsymbol{m}) > \sqrt{a} \iff \varepsilon\cdot\boldsymbol{w} > 0$;

(iii) *If* $\mu(d;\boldsymbol{m}) > \sqrt{a}$, *then* $E := \varepsilon\cdot\varepsilon = \sum \varepsilon_i^2 < 1$;

(iv) *Let* $y(a) := a + 1 - 3\sqrt{a}$ *where* $a = p/q$. *Then*

$$-\sum \varepsilon_i = 1 + \tfrac{d}{\sqrt{a}}\big(y(a) - \tfrac{1}{q}\big). \tag{2.1.2}$$

Proof. Lemma 1.2.6 and Proposition 1.2.12 imply that

$$\mu(d;\boldsymbol{m})\, d \;=\; \boldsymbol{w}\cdot\boldsymbol{m} \;\le\; \|\boldsymbol{w}\|\,\|\boldsymbol{m}\| \;=\; \sqrt{a}\sqrt{d^2+1}.$$

This proves (i). (ii) is immediate, while (iii) follows from (ii) because

$$d^2+1 \;=\; \boldsymbol{m}\cdot\boldsymbol{m} \;=\; \big(\tfrac{d}{\sqrt{a}}\boldsymbol{w}(a)+\varepsilon\big)\cdot\big(\tfrac{d}{\sqrt{a}}\boldsymbol{w}(a)+\varepsilon\big) \;=\; d^2 + 2\tfrac{d}{\sqrt{a}}\boldsymbol{w}(a)\cdot\varepsilon + \varepsilon\cdot\varepsilon.$$

Finally, to prove (iv) observe that

$$3d-1 \;=\; \sum m_i \;=\; \tfrac{d}{\sqrt{a}}\big(a+1-\tfrac{1}{q}\big) + \sum \varepsilon_i.$$

Hence

$$\tfrac{d}{\sqrt{a}}\big(a+1-3\sqrt{a}\big) - \tfrac{d}{\sqrt{a}\,q} + 1 + \sum\varepsilon_i = 0.$$

This completes the proof. □

Remark 2.1.2. (i) Proposition 2.1.1 (iii) implies that an element $(d;\boldsymbol{m})\in\mathcal{E}$ gives an obstruction at a (i.e. has $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$) only if the vector $\boldsymbol{m}$ is “almost parallel” to the vector $\boldsymbol{w}(a)$. In particular, if we are interested in solutions that provide obstructions when $a = k+x$ for $x\in(0,1)$ we need the first k entries to be equal within the allowable error. As we will see in Lemma 2.1.7 below, this means that the first k entries of $\boldsymbol{m}$ must lie in the set $\{m_1, m_1-1\}$, with at most one entry different from the others. Some elements of $\mathcal{E}$ with $d\le 9$ that satisfy these conditions for $a\in[6,8]$ are

$$(2;1^{\times 5}),\;\; (3;2,1^{\times 6}),\;\; (5;2^{\times 6},1^{\times 2}),\;\; (8;3^{\times 7},1^{\times 2}).$$

It turns out that these elements all do give obstructions.

(ii) Another noteworthy point is that $y(a) = 0$ when $a=\tau^4$. Therefore (iv) gives most information when $a-\tau^4$ is quite large, e.g. if $a>7$; see §5.1. ◇

The next result explains the basic structure of the constraints. Throughout we write $\ell(\boldsymbol{m})$ for the number of positive entries in $\boldsymbol{m}$, and $\ell(a)$ for the length of the weight sequence $\boldsymbol{w}(a)$.

Lemma 2.1.3. *Let* $(d;\boldsymbol{m})\in\mathcal{E}$, *and suppose that* I *is a maximal nonempty open interval such that* $\sqrt{a} < \mu(d;\boldsymbol{m})(a)$ *for all* $a\in I$. *Then there is a unique* $a_0\in I$ *such that* $\ell(a_0) = \ell(\boldsymbol{m})$. *Moreover* $\ell(a)\ge\ell(\boldsymbol{m})$ *for all* $a\in I$.

Proof. Denote by $w_i(a)$ the ith weight of a considered as a function of a. Then it is piecewise linear, and is linear on any open interval that does not contain an element a' with length $\ell(a') \le i$. That is, the formula[4] for $w_i(a)$ can change only if it or one of the earlier weights becomes zero.

Therefore if $\ell(a) > \ell(\boldsymbol{m})$ for all $a \in I$, the function $\mu(d;\boldsymbol{m})(a)$ is linear in I. But this is impossible since the function $\sqrt{a}$ is concave and $I \subset (1,9)$ is bounded.

Thus there is $a_0 \in I$ with $\ell(a_0) \le \ell(\boldsymbol{m})$. On the other hand, if $\ell(a) < \ell(\boldsymbol{m})$, then $\sum_{i\le \ell(a)} m_i^2 < d^2+1$, so that

$$|\boldsymbol{w}\cdot\boldsymbol{m}| \;\le\; \|\boldsymbol{w}\|\sqrt{\sum_{i\le\ell(a)} m_i^2} \;\le\; d\|\boldsymbol{w}\| \;=\; d\sqrt{a}.$$

Hence $\mu(d;\boldsymbol{m})(a) \le \sqrt{a}$, i.e. $a \notin I$.

The uniqueness follows from the properties of continued fractions. We claim that if $b > a$ and $\ell(b) = \ell(a)$ then there must be some number $y \in (a,b)$ with $\ell(y) < \ell(a)$. Since such y cannot be in I, this gives the required uniqueness. To prove the claim, let a have continued fraction expansion $[\ell_0;\ell_1,\dots,\ell_N]$ and consider the functions $x_j(a)$ as in equation (1.2.1). If N is even, the function $x_N(z)$ decreases as z increases. Hence ℓ_N increases and so the next number $z > a$ with length $\le \ell(a)$ is $[\ell_0;\ell_1,\dots,\ell_{N-1}]$ which has length $< \ell(a)$. Similarly, if we look for numbers $z < a$ with $\ell(z) \le \ell(a)$, then the first one is $[\ell_0;\dots,\ell_N - 1]$. Similar arguments apply if N is odd. □

Corollary 2.1.4. *Suppose that $c(a) > \sqrt{a}$. Then*

(i) *There are (possibly equal) elements $(d^\pm;\boldsymbol{m}^\pm) \in \mathcal{E}$ and $\varepsilon > 0$ such that*

$$c(z) \;=\; \begin{cases} \mu(d^-;\boldsymbol{m}^-)(z) & \text{for all } z \in (a-\varepsilon, a], \\ \mu(d^+;\boldsymbol{m}^+)(z) & \text{for all } z \in [a, a+\varepsilon). \end{cases}$$

(ii) *On each of the intervals in (i) there are rational numbers $\alpha,\beta \ge 0$ such that $c(a) = \alpha + \beta a$.*

(iii) *The set of $(d;\boldsymbol{m})$ such that $c(a) = \mu(d;\boldsymbol{m})(a)$ is finite.*

Proof. Since $c(a) > \sqrt{a}$, there exists $D \in \mathbb{N}$ with $\sqrt{1+1/D^2} < c(a)/\sqrt{a}$. If $(d;\boldsymbol{m}) \in \mathcal{E}$ is such that $\mu(d;\boldsymbol{m})(a) = c(a) > \sqrt{a}$, then $d \le D$ by Proposition 2.1.1 (i). But there are only finitely many elements $(d;\boldsymbol{m}) \in \mathcal{E}$ with $d \le D$. Since $c(a)$ is continuous by Lemma 1.1.1, we must have $\sqrt{1+1/D^2} < c(z)/\sqrt{z}$ for all z sufficiently close to a. Further, as we have seen in the proof of Lemma 2.1.3, each function $\mu(d;\boldsymbol{m})(z)$ is piecewise linear, and it has rational coefficients because the weight functions $w_i(z)$ do. Hence, $c(z)$ is the supremum of a finite number of rational linear functions. This proves (i) and (iii). Moreover, if near a we write $c(z) = \alpha + \beta z$ for some rational numbers α,β, then $\beta \ge 0$ since c is nondecreasing, while $\alpha \ge 0$ because of the scaling property in equation (1.1.2). □

[4]See Lemma 2.2.1 for an explicit expression.

Let us call an element $(d;\boldsymbol{m}) \in \mathcal{E}$ *perfect* if $\boldsymbol{m}$ is a multiple of the weight vector $\boldsymbol{w}(b)$ of some $b>1$. The next lemma combined with Corollary 2.1.4 shows that these elements determine $c(a)$ for a near b.

Lemma 2.1.5. *Suppose that $(d;\boldsymbol{m}) \in \mathcal{E}$ is perfect: $\boldsymbol{m} = \kappa\,\boldsymbol{w}(b)$ for some $b>1$. Then*

(i) *$\mu(d;\boldsymbol{m})(b) = c(b) > \sqrt{b}$, and $(d;\boldsymbol{m})$ is the only class with $\mu(d;\boldsymbol{m})(b) = c(b)$.*

(ii) *$\boldsymbol{m} = q\,\boldsymbol{w}(b)$ where $b = p/q$ in lowest terms, and $b < \tau^4$.*

Proof. (i) Since $(d;\boldsymbol{m}) \in \mathcal{E}$ and $\boldsymbol{m} = \kappa\,\boldsymbol{w}(b)$, we have $d^2 < d^2+1 = \boldsymbol{m}\cdot\boldsymbol{m} = \kappa^2\boldsymbol{w}(b)\cdot\boldsymbol{w}(b) = \kappa^2 b$, whence $d < \kappa\sqrt{b}$. Therefore,

$$\mu(d;\boldsymbol{m})(b) \;:=\; \frac{\boldsymbol{m}\cdot\boldsymbol{w}(b)}{d} \;=\; \frac{\kappa b}{d} \;>\; \frac{\kappa b}{\kappa\sqrt{b}} \;=\; \sqrt{b}.$$

Let $(d';\boldsymbol{m}') \in \mathcal{E}$ be another solution. Since $(d';\boldsymbol{m}') \neq (d;\boldsymbol{m})$, positivity of intersections (part (ii) of Proposition 1.2.12) shows that $dd' \geq \boldsymbol{m}\cdot\boldsymbol{m}' = \kappa\,\boldsymbol{m}'\cdot\boldsymbol{w}(b)$. This and $d^2 < d^2+1 = \boldsymbol{m}\cdot\boldsymbol{m} = \kappa\,\boldsymbol{m}\cdot\boldsymbol{w}(b)$ yield

$$\mu(d';\boldsymbol{m}')(b) \;=\; \frac{\boldsymbol{m}'\cdot\boldsymbol{w}(b)}{d'} \;\leq\; \frac{d}{\kappa} \;<\; \frac{\boldsymbol{m}\cdot\boldsymbol{w}(b)}{d} \;=\; \mu(d;\boldsymbol{m})(b).$$

(ii) Let $\boldsymbol{w}(b) = (w_1,\dots,w_M)$. Since $w_M = \frac{1}{q}$ and $m_M \in \mathbb{Z}$, we have $\kappa = sq$ for some integer s. Equations (1.2.4) and Lemma 1.2.6 give

$$\begin{aligned} 3d-1 \;&=\; \sum m_i \;=\; \kappa\sum w_i \;=\; sq\big(b+1-\tfrac{1}{q}\big),\\ d^2+1 \;&=\; \sum m_i^2 \;=\; \kappa^2\sum w_i^2 \;=\; (sq)^2 b. \end{aligned}$$

If $s=1$, then $3d = q(b+1)$ so that

$$1+b-3\sqrt{b} = 3\tfrac{d}{q} - 3\tfrac{\sqrt{d^2+1}}{q} < 0.$$

Thus $\sqrt{b} < \tau^2$.

Otherwise, adding the two above equations gives $s|d(d+3)$. But the first equation shows that s,d are mutually prime. Therefore $s|d{+}3$ and $s|3d{-}1$; hence $s|10$. Therefore $s = 2$, 5 or 10.

The identity $(d+3)^2 = (d^2+1)+2(3d-1)+10$ shows that $s^2|2(3d-1)+10$. If $2|s$ this means that d is even which is impossible since $3d-1$ is even.

If $s=5$ then $3d+4 = 5q(b+1)$ so that

$$y(b) \;:=\; 1+b-3\sqrt{b} = \tfrac{1}{5q}\big(3d+4-3\sqrt{d^2+1}\big) > 0 \tag{2.1.3}$$

Thus $\sqrt{b} > \tau^2$. But this is impossible: For $(d;\boldsymbol{m}) \in \mathcal{E}$ must have nonnegative intersection with the class $(3;2,1^{\times 6}) \in \mathcal{E}$. Therefore

$$2m_1 + m_2 + \cdots + m_7 \leq 3d.$$

If $b \in [\tau^4, 7]$, then $\boldsymbol{m} = 5q(1^{\times 6}, b-6, \dots)$, and we obtain

$$5q(1+b) \leq 3d,$$

a contradiction. Assume now that $b = 7\frac{r}{q} > 7$. Assume first that $b = 7\frac{1}{q}$. Then $3d - 1 = 40q$ and $d^2 + 1 = 25q^2(7 + \frac{1}{q})$. Solving the first equation for d and inserting the result into the second equation, we get the equation

$$5q^2 - 29q + 2 \;=\; 0$$

whose solutions are not integral. Thus $b = 7\frac{r}{q}$ with $r \geq 2$. By (2.1.3),

$$q \;=\; \frac{3d + 4 - 3\sqrt{d^2 + 1}}{5y(b)} \;<\; \frac{4}{5y(b)} \;<\; \frac{4}{5y(7)} \;<\; 13,$$

whence $q \leq 12$. Thus $b \geq 7\frac{2}{12}$, and so $q < \frac{4}{5y(7\frac{1}{6})} < 6$, whence $q \leq 5$. Thus $b \geq 7\frac{2}{5}$, and so $q < \frac{4}{5y(7\frac{2}{5})} < 4$, whence $q \leq 3$. Thus $b \geq 7\frac{2}{3}$, and so $q < \frac{4}{5y(7\frac{2}{3})} < 3$, whence $q \leq 2$. Thus $b \geq 8$, and so $q < \frac{4}{5y(8)} < 2$, whence $q = 1$. Thus $b \geq 9$, and so $q < \frac{4}{5y(9)} < 1$, which is impossible. □

Remark 2.1.6. We show later that the only perfect elements are those at the numbers b_n of the Fibonacci stairs; see Corollary 3.1.3. However there are many nearly perfect elements that are relevant to the problem such as the classes $E(a_n)$ of Theorem 3.1.1 and the classes $E\big(b_k(i)\big)$ of Proposition 4.2.2. These elements are perfect except for some adjustments on the last block. ◇

The next lemma expands on the first part of Remark 2.1.2.

Lemma 2.1.7. *Assume that $(d; \boldsymbol{m}) \in \mathcal{E}$ is such that $\mu(d; \boldsymbol{m})(a) > \sqrt{a}$. Let $J := \{k, \dots, k+s-1\}$ be a block of $s \geq 2$ consecutive integers for which $w(a_i)$, $i \in J$, is constant. Then*

(i) *One of the following holds:*

$$\begin{aligned} &m_k = \cdots = m_{k+s-1} \quad \textit{or} \\ &m_k = \cdots = m_{k+s-2} = m_{k+s-1} + 1 \quad \textit{or} \\ &m_k - 1 = m_{k+1} = \cdots = m_{k+s-1}. \end{aligned}$$

(ii) *There is at most one block of length $s \geq 2$ on which the m_i are not all equal.*

(iii) *If there is a block J of length s on which the m_i are not all equal then $\sum_{i \in J} \varepsilon_i^2 \geq \frac{s-1}{s}$.*

Proof. Let $w_i(a) = x$ for $i \in J$. By Proposition 2.1.1 (iii) we have

$$\sum_{i=k}^{k+s-1} \left| \tfrac{dx}{\sqrt{a}} - m_i \right|^2 \;=\; \sum_{i=k}^{k+s-1} \varepsilon_i^2 < 1.$$

Thus $\{m_k, \dots, m_{k+s-1}\}$ can contain at most two different integers, which must be neighbors if they are different, say m, $m+1$. We can also clearly assume that $m < \frac{dx}{\sqrt{a}} < m + 1$. Therefore (i) holds when $s < 4$.

So suppose that $s \geq 4$ and assume that $m+1$ occurs t times. Set $v = \frac{dx}{\sqrt{a}} - n \in [0,1)$, where $n \in \mathbb{Z}$. Then the squared error on this block is

$$\begin{aligned}\sum_{i=k}^{k+s-1} \left|\frac{dx}{\sqrt{a}} - m_i\right|^2 &= \sum_{i=k}^{k+t-1} |v-1|^2 + \sum_{i=k+t}^{k+s-1} |v|^2 \\ &= t(v-1)^2 + (s-t)v^2 \\ &\geq \tfrac{2(s-2)}{s} \text{ for all } v \in (0,1) \text{ if } t \in \{2, \dots, s-2\}.\end{aligned}$$

But $2(s-2)/s \geq 1$ when $s \geq 4$. This proves (i). Parts (ii) and (iii) follow from the fact that the minimum squared error on a block of length s on which the m_i are not all equal is $1 - \frac{1}{s} \geq \frac{1}{2}$. □

The following lemma will also be important for detecting potentially obstructive solutions $(d; \boldsymbol{m})$.

Lemma 2.1.8. *Let $(d;\boldsymbol{m}) \in \mathcal{E}$ be such that $\mu(d;\boldsymbol{m}) > \sqrt{a}$ for some a with $\ell(a) = \ell(\boldsymbol{m}) = M$. Let $w_{k+1}, \dots, w_{k+s}$ be a block, but not the first block, of $\boldsymbol{w}(a)$.*

(i) *If this block is not the last block, then*

$$|m_k - (m_{k+1} + \cdots + m_{k+s} + m_{k+s+1})| < \sqrt{s+2}.$$

If this block is the last block, then

$$|m_k - (m_{k+1} + \cdots + m_{k+s})| < \sqrt{s+1}.$$

(ii) *Always,*

$$m_k - \sum_{i=k+1}^{M} m_i < \sqrt{M-k+1}.$$

Proof. (i) We prove the first claim, the second claim is proven in the same way. By definition (2.1.1) of the errors, $m_i = \frac{d}{\sqrt{a}} w_i + \varepsilon_i$ for all i. Since $w_{k+1} = \cdots = w_{k+s}$ and $w_{k+s+1} = w_k - s\, w_{k+1}$, we have that

$$m_k - (m_{k+1} + \cdots + m_{k+s} + m_{k+s+1}) = \varepsilon_k - (\varepsilon_{k+1} + \cdots + \varepsilon_{k+s} + \varepsilon_{k+s+1}),$$

and so

$$|m_k - (m_{k+1} + \cdots + m_{k+s} + m_{k+s+1})| \leq |\varepsilon_k| + |\varepsilon_{k+1}| + \cdots + |\varepsilon_{k+s}| + |\varepsilon_{k+s+1}|.$$

Since $\varepsilon \cdot \varepsilon = \sum \varepsilon_i^2 < 1$ by Proposition 2.1.1 (iii), the latter sum is $< \sqrt{s+2}$.

(ii) If the block $w_{k+1}, \dots, w_{k+s}$ is the last block, then $M = k+s$, and so the stated estimate is the same as in (i). So assume that $w_{k+1}, \dots, w_{k+s}$ is not the last block. Since $w_k = s\, w_{k+1} + w_{k+s+1}$, we then have $w_k < \sum_{i=k+1}^M w_i$. Therefore,

$$m_k < -\varepsilon_k + \sum_{i=k+1}^{M} (m_i + \varepsilon_i) \leq \sum_{i=k+1}^{M} m_i + \sum_{i=k}^{M} |\varepsilon_i| < \sum_{i=k+1}^{M} m_i + \sqrt{M-k+1},$$

as claimed. □

2.2. **Some identities for weight expansions.** This subsection establishes some rather surprising identities for weight expansions.

Let $a > 1$ be a rational number and consider its continued fraction expansion

$$a := [\ell_0; \ell_1, \ldots, \ell_N] = \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \ldots}}. \tag{2.2.1}$$

Usually, to avoid ambiguity, we assume that $\ell_N \geq 2$, but in this section only it is convenient to permit the case $\ell_N = 1$ as well. We define the sequence $\alpha^a := (\alpha^a_j)_{j=0}^{N+1}$ by setting $\alpha^a_0 = 1$, $\alpha^a_1 = -\ell_0$, and

$$\alpha^a_j = \alpha^a_{j-2} - \ell_{j-1}\alpha^a_{j-1}, \quad j = 2, \ldots, N+1. \tag{2.2.2}$$

Similarly, define $\beta^a = (\beta^a_j)_{j=0}^{N+1}$ by the same recursive formula, but starting with $\beta^a_0 = 0$, $\beta^a_1 = 1$. Thus β^a does not depend on ℓ_0. Both sequences alternate in sign.

We chose the notation α^a, β^a for these sequences because, as we now see, they are the coefficients of the linear functions $w_i(z)$ for z lying on the appropriate side of a. Write the weight expansion $\boldsymbol{w}(a)$ of a as

$$\boldsymbol{w}(a) = \left(1^{\times \ell_0}, (x_1(a))^{\times \ell_1}, \ldots, (x_N(a))^{\times \ell_N}\right) = \big(\underbrace{1, \ldots, 1}_{\ell_0}, \underbrace{x_1, \ldots, x_1}_{\ell_1}, \ldots\big).$$

Then the weights $x_j := x_j(a)$ satisfy the recursive formula (2.2.2) with $x_0 = 1$, and $x_1 = a - \ell_0$, so that $x_j = \alpha^a_j + a\beta^a_j$ for $j \leq N$. If N is odd, this formula for $x_j(z)$, $j \leq N$, continues to hold for all $z < a$ that are so close to a that the $x_j(z)$ are positive. In this case, $z = [\ell_0; \ell_1, \ldots, \ell_N, h, \ldots]$ where $h \geq 1$. Similarly, if N is even, it holds for $z > a$ and sufficiently close to a. This proves the following result.

Lemma 2.2.1. *Let a and N be as above. Then if N is odd there is $\varepsilon > 0$ and $h \geq 1$ such that for $z \in (a - \varepsilon, a)$ we have*

$$\boldsymbol{w}(z) = \left(1^{\times \ell_0}, \big(x_1(z)\big)^{\times \ell_1}, \ldots, \big(x_N(z)\big)^{\times \ell_N}, \big(x_{N+1}(z)\big)^{\times h}, \ldots\right)$$

where $x_j(z) = \alpha^a_j + z\beta^a_j$ for $j \leq N+1$. If N is even, the same statement holds for $z \in (a, a+\varepsilon)$. Moreover, in both cases $x_j(z)$ is an increasing function of z for j odd, and a decreasing function for j even.

We now give a second description for the sequences $\boldsymbol{w}(a)$, α^a, and β^a in terms of the convergents of a related number $\overleftarrow{a}$ (the mirror of a) which helps to explain their symmetry properties.

Definition 2.2.2. *Let $\ell_j \geq 1$ for $0 \leq j \leq N$, where $N \geq 1$. We define*

$$\overrightarrow{a} := [\ell_0; \ell_1, \ldots, \ell_N], \quad \overleftarrow{a} := [\ell_N; \ell_{N-1}, \ldots, \ell_0],$$

and call $\overleftarrow{a}$ *the* **mirror** *of* $\vec{a}$. *The convergents* $(\vec{a})_k$, $0 \le k \le N$, *to* $\vec{a}$ *are defined by setting*

$$(\vec{a})_k := [\ell_0; \ell_1, \dots, \ell_k] =: \frac{p_k(\vec{a})}{q_k(\vec{a})} =: \frac{p_k}{q_k},$$

where $p_k(\vec{a})$, $q_k(\vec{a})$ *are the numerator and denominator of the rational number represented by* $(\vec{a})_k$. *In particular,* $(\vec{a})_N = p_N/q_N = a$, *and for short we write* $a = \vec{a}$.

We define the **normalized weight sequence** *of* $\vec{a}$ *as*

$$W(\vec{a}) := q_N \boldsymbol{w}(\vec{a}) = \Big(\big(X_0(\vec{a})\big)^{\times \ell_0}, \dots, \big(X_N(\vec{a})\big)^{\times \ell_N} \Big).$$

In particular, $X_N = 1$ *always. Finally, we define the* **(signed) mirror** *of a sequence* $W := \big(X_0^{\times \ell_0}, X_1^{\times \ell_1}, \dots, X_N^{\times \ell_N}\big)$ *to be*

$$\widehat{W} := \Big(X_N^{\times \ell_N}, \big(-X_{N-1}\big)^{\times \ell_{N-1}}, \dots, \big((-1)^N X_0\big)^{\times \ell_0} \Big)$$

where we reverse the order and change signs.

Note that the sequence of weights $\boldsymbol{w}(a)$ is independent of the ambiguity in the ℓ_j, but its block description and the X_j do depend on this choice.

The first part of the next lemma is well known. We then show that the normalized weights of $a = \vec{a}$ are equal to the numerators of the convergents of $\overleftarrow{a}$. Further the coefficients α^a (resp. β^a) are the numerators (resp. denominators) of the convergents of $a = \vec{a}$ shifted by 1. Define $p_{-1}(a) = 1$, $q_{-1}(a) = 0$.

Lemma 2.2.3. *Let* $a = \vec{a}$ *be as above. Then:*

(i) $\overleftarrow{a} = p_N(\vec{a})/p_{N-1}(\vec{a})$; *in particular,* $p_N(\vec{a}) = p_N(\overleftarrow{a})$.

(ii) *For* $0 \le j \le N$, *we have*

$$X_j(\vec{a}) = |\alpha^{\overleftarrow{a}}_{N-j}|, \qquad X_{N-j}(\overleftarrow{a}) = |\alpha^{\vec{a}}_j|.$$

Further $|\alpha^a_j| = p_{j-1}(a)$ *for all* a *and* $0 \le j \le N+1$, *so that* $X_j(\vec{a}) = p_{N-j-1}(\overleftarrow{a})$ *for* $j = 0, \dots, N$.

(iii) *Define* $u = \vec{u} := [\ell_1; \dots, \ell_N]$. *Then* $\overleftarrow{u} = (\overleftarrow{a})_{N-1}$, *and for* $1 \le j \le N+1$ *we have*

$$|\beta^a_j| = |\alpha^u_{j-1}| = q_{j-1}(a).$$

Proof. The following matrix identity holds by induction on k:

$$\begin{pmatrix} \ell_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} \ell_k & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_k(\vec{a}) & p_{k-1}(\vec{a}) \\ q_k(\vec{a}) & q_{k-1}(\vec{a}) \end{pmatrix}. \tag{2.2.3}$$

(i) follows by considering its transpose.

To prove (ii), observe that the elements $X_j := X_j(a)$ are decreasing positive integers with $X_N = 1$, $X_{N+1} = 0$, and by Definition 1.2.5 may be defined backwards by the iterative relation

$$X_{N-j-1} = X_{N-j+1} + \ell_{N-j}\, X_{N-j}, \qquad j \ge 1.$$

Similarly, $\alpha_0^a = 1$ and, if we set $\alpha_{-1} = 0$, the identity (2.2.2) implies that the $|\alpha_j^a|$ are increasing positive integers satisfying

$$|\alpha_{j+1}^a| = |\alpha_{j-1}^a| + \ell_j\, |\alpha_j^a|, \qquad j \geq 0.$$

But $\overleftarrow{a} = [\ell_0'; \dots, \ell_N']$ where $\ell_j' = \ell_{N-j}$. Therefore $X_j(\vec{a}) = |\alpha_{N-j}^{\overleftarrow{a}}|$ for $j = 0, \dots, N$. Since $\overleftarrow{\overleftarrow{a}} = \vec{a}$, we also have that $X_j(\overleftarrow{a}) = |\alpha_{N-j}^{\vec{a}}|$ for $j = 0, \dots, N$.

To understand the relation between α_j^a and $p_{j-1} := p_{j-1}(a)$, note first that $p_0 = \ell_0 = |\alpha_1^a|$. Further, (2.2.3) implies that $p_j = p_{j-2} + \ell_j p_{j-1}$. Therefore $|\alpha_j^a| = p_{j-1}(a)$ for $j = 0, \dots, N+1$, as claimed. This proves (ii).

The first claim in (iii) is obvious. To prove the second, note that the identity $|\beta_j^a| = |\alpha_{j-1}^u|$ follows immediately from the definitions of α and β. Therefore, by (ii), $|\beta_j^a| = |\alpha_{j-1}^u| = p_{j-2}(u)$. But if $v := [0; \ell_1, \dots, \ell_N]$, then $\frac{1}{u} = v = a - \lfloor a \rfloor$. Hence $p_0(u) = q_1(v) = q_1(a)$, and more generally, $p_k(u) = q_{k+1}(a)$. Hence $|\beta_j^a| = q_{k-1}(a)$. □

Corollary 2.2.4.

$$\begin{aligned} \widehat{W}(\overleftarrow{a}) \;&:=\; \Big(\big(X_N(\overleftarrow{a})\big)^{\times \ell_0}, \big(-X_{N-1}(\overleftarrow{a})\big)^{\times \ell_1}, \dots, \big((-1)^N X_0(\overleftarrow{a})\big)^{\times \ell_N} \Big) \\ &=\; \Big(\big(\alpha_0^a\big)^{\times \ell_0}, \big(\alpha_1^a\big)^{\times \ell_1}, \dots, \big(\alpha_N^a\big)^{\times \ell_N} \Big) \end{aligned}$$

We now show that there is a mirror version of the quadratic relation $\boldsymbol{w}(a) \cdot \boldsymbol{w}(a) = a$ proven in Lemma 1.2.6. Since $W(\vec{a}) = q_N(\vec{a})\, \boldsymbol{w}(a)$, we may also write this identity as

$$W(\vec{a}) \cdot W(\vec{a}) \;=\; p_N(\vec{a})\, q_N(\vec{a}).$$

One can prove the mirror formula directly, using the geometric definition of the weights. We now present a more elegant proof formulated in terms of the tridiagonal matrices first introduced by Sylvester in [24]:[5]

$$\mathrm{Tri}\,(\vec{a}) \;:=\; \begin{pmatrix} \ell_0 & 1 & 0 & 0 & \dots \\ -1 & \ell_1 & 1 & 0 & \dots \\ 0 & -1 & \ell_2 & 1 & \dots \\ 0 & 0 & -1 & \ell_3 & \dots \\ \dots & & & & \end{pmatrix}.$$

In the following we number the rows and columns of the $(N+1) \times (N+1)$ matrix $A := \mathrm{Tri}\,(\vec{a})$ by the labels $i, j \in \{0, \dots, N\}$. Thus $A := (a_{ij})_{0 \leq i,j \leq N}$ where $a_{ii} = \ell_i$. Its determinant is denoted $|A|$.

Lemma 2.2.5. *Let $\vec{a} = [\ell_0; \dots, \ell_N] = p_N(\vec{a})/q_N(\vec{a})$, and denote by A_{ij} the determinant of the i, j minor of $A := \mathrm{Tri}\,(\vec{a})$, where $0 \leq i, j \leq N$. Then:*

[5]We are indebted to Andrew Ranicki for telling us about Sylvester's work. There are corresponding results for the Hirzebruch–Jung continued fractions coming from tridiagonal matrices with subdiagonal entries equal to 1.

(i) *for all* $\vec{a}$ *we have* $p_N(\vec{a}) = |\mathrm{Tri}\,(\vec{a})|$,

(ii) *for all* i *we have* $A_{ii} = p_{i-1}(\vec{a})\, p_{N-i-1}(\overleftarrow{a}) = X_{N-i}(\overleftarrow{a})\, X_i(\vec{a})$.

Proof. Part (i) follows by induction. The first equality in (ii) is immediate, and the second holds because Lemma 2.2.3 (ii) implies that $p_{i-1}(\vec{a}) = X_{N-i}(\overleftarrow{a})$ and $p_{N-i-1}(\overleftarrow{a}) = X_i(\vec{a})$. □

Proposition 2.2.6. *Let* $\vec{a} = [\ell_0; \dots, \ell_N]$. *Then* $W(\vec{a}) \cdot \widehat{W}(\overleftarrow{a}) = p_N(\vec{a})$ *if* N *is even, and* $= 0$ *if* N *is odd.*

Proof. Consider the $(N+1)\times(N+1)$ matrix $A := \mathrm{Tri}\,(\vec{a}) = (a_{ij})_{0\le i,j\le N}$. First observe that

$$\sum_{i,j=0}^{N} (-1)^j a_{ij} A_{ij} \;=\; \begin{cases} |A| & \text{if } N \text{ is even,} \\ 0 & \text{if } N \text{ is odd.} \end{cases}$$

This holds because the sum for each fixed i is $(-1)^i\,|A|$. Next observe that $a_{ij}A_{ij} = 0$ unless $|i-j| \le 1$. Further, if $i = j-1$ then $A_{ij} = -A_{ji}$ so that $a_{ij}A_{ij} = a_{ji}A_{ji}$. It follows that

$$\sum_{ij} (-1)^j a_{ij} A_{ij} \;=\; \sum_i (-1)^i \ell_i A_{ii} \;=\; \sum_i (-1)^i\, \ell_i\, X_i(\vec{a})\, X_{N-i}(\overleftarrow{a}),$$

where the last equality uses Lemma 2.2.5 (ii). Finally notice that

$$W(\vec{a}) \cdot \widehat{W}(\overleftarrow{a}) \;=\; \sum_i (-1)^i\, \ell_i\, X_i(\vec{a})\, X_{N-i}(\overleftarrow{a}),$$

and that $|A| = p_N(\vec{a})$ by Lemma 2.2.5 (i). □

Corollary 2.2.7. *Let* $x_j := x_j(a)$, *where* $a = [\ell_0; \ell_1, \dots, \ell_N] > 1$ *and define* α^a_j, β^a_j *as in equation* (2.2.2). *Then:*

(i) *If* N *is even,* $\sum_{j=0}^N \ell_j\, x_j\, \alpha^a_j = a$ *and* $\sum_{j=0}^N \ell_j\, x_j\, \beta^a_j = 0$;

(ii) *If* N *is odd,* $\sum_{j=0}^N \ell_j\, x_j\, \alpha^a_j = 0$ *and* $\sum_{j=0}^N \ell_j\, x_j\, \beta^a_j = 1$.

Proof. The sums involving α have the stated value by Proposition 2.2.6 and Corollary 2.2.4. To prove the claims involving β^a, write

$$\vec{u} \;=\; [\ell_1; \ell_2, \dots, \ell_N] \;=\; [\ell'_0; \ell'_1, \dots, \ell'_{N'}],$$

where $N' = N-1$. Note that $X_i(u) = X_{i+1}(a)$, for $0 \le i \le N'$. Thus, with $\beta^a_0 := 0$, and since $\beta^a_j = \alpha^u_{j-1}$ by Lemma 2.2.3 (iii), we find that

$$\begin{aligned} q_N(a) \sum_{j=0}^{N} \ell_j\, x_j(a)\, \beta^a_j \;&=\; \sum_{j=1}^{N} \ell_j\, X_j(a)\, \alpha^u_{j-1} \\ &=\; \sum_{i=0}^{N'} \ell'_i\, X_i(u)\, \alpha^u_i. \end{aligned}$$

By what we have already shown, this sum is 0 when N' is odd (i.e. N is even) and equals the numerator $p_{N'}(u)$ of u when N' is even (i.e. N is odd). But $p_{N'}(u)$ is the denominator of $\frac{1}{u} = [0; \ell_1, \dots, \ell_N] = a - \lfloor a \rfloor$, and so equals $q_N(a)$. The result follows. □

2.3. **The nature of the obstructions.** We saw in Corollary 2.1.4 that near each point a where $c(a) > \sqrt{a}$, the function c is the supremum of a finite number of piecewise linear functions $\mu(d;\boldsymbol{m})$, and that each linear segment of c has the form $z \mapsto \alpha+\beta z$ with rational and nonnegative coefficients. The next example shows that the coefficients of the functions $\mu(d;\boldsymbol{m})$, though rational, are not restricted in this way even if we suppose that $\mu(d;\boldsymbol{m})(z) > \sqrt{z}$.

Example 2.3.1. Consider the class $(d;\boldsymbol{m}) = \big(10; 4^{\times 6}, 1^{\times 5}\big)$ in $\mathcal{E}$. (Under the name $E(a_2)$, this class will play a role in Section 3.) Abbreviate $\mu(z) = \mu(d;\boldsymbol{m})(z)$. We compute $d\,\mu(z) = 10\,\mu(z)$ on the interval $\big[6, 6\frac{1}{2}\big]$.

$$
\begin{array}{llll}
\text{on } I_1 = \big[6, 6\frac14\big]: & -6+5z \text{ on } \big[6\ \ , 6\frac15\big], & 25 & \text{on } \big[6\frac15, 6\frac14\big]; \\
\text{on } I_2 = \big[6\frac14, 6\frac13\big]: & 4z \text{ on } \big[6\frac14, 6\frac27\big], & 44-3z & \text{on } \big[6\frac27, 6\frac13\big]; \\
\text{on } I_3 = \big[6\frac13, 6\frac25\big]: & -13+6z \text{ on } \big[6\frac13, 6\frac38\big], & 38-2z & \text{on } \big[6\frac38, 6\frac25\big]; \\
\text{on } I_4 = \big[6\frac25, 6\frac12\big]: & 6+3z \text{ on } \big[6\frac25, 6\frac37\big], & 51-4z & \text{on } \big[6\frac37, 6\frac12\big];
\end{array}
$$

see Figure 2.1. The figure also shows the graph of $\sqrt{z}$ and of $c(z)$ (dashed) on $\big[6, 6\frac12\big]$, which by Theorem 1.1.2 (i) is

$$c(z) = \tfrac52 \text{ on } \big[6, 6\tfrac14\big], \qquad c(z) = \tfrac25 z \text{ on } \big[6\tfrac14, 6\tfrac12\big].$$

Note that $\ell(\boldsymbol{m}) = \ell(a) = 11$ at $6\frac15$, $6\frac27$, $6\frac38$, $6\frac37$. Also note that at $a = 6\frac38$ we have $\mu(a) > \sqrt{a}$ while at $a = 6\frac37$ we have $\mu(a) < \sqrt{a}$.

Similar results hold for the functions $\mu(d;\boldsymbol{m})$ given by the classes $E(a_n)$, $n > 2$, of Theorem 3.1.1. For example, one can use Corollary 2.2.7 to show that these functions equal $c(z)$ for z near a_n. ◇

We now show that although the coefficients α, β may be negative, they are somewhat restricted.

Proposition 2.3.2. *Let $(d;\boldsymbol{m}) \in \mathcal{E}$ and $a \in \mathbb{Q}$ be such that $\ell(\boldsymbol{m}) = \ell(a)$ and $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$. Write $a =: p/q$ in lowest terms, let $m := m_M$ be the last nonzero entry in $\boldsymbol{m}$ and let I be the connected component of the set $\{z \mid \mu(d;\boldsymbol{m})(z) > \sqrt{z}\}$ that contains a. Then there are integers $A < p$ and $B < (m+1)q$ such that*

$$
d\,\mu(d;\boldsymbol{m})(z) = \begin{cases} A + Bz & \text{if } z < a,\ z \in I, \\ (A+mp) + (B-mq)z & \text{if } z > a,\ z \in I. \end{cases}
$$

We begin the proof by establishing the following lemma.

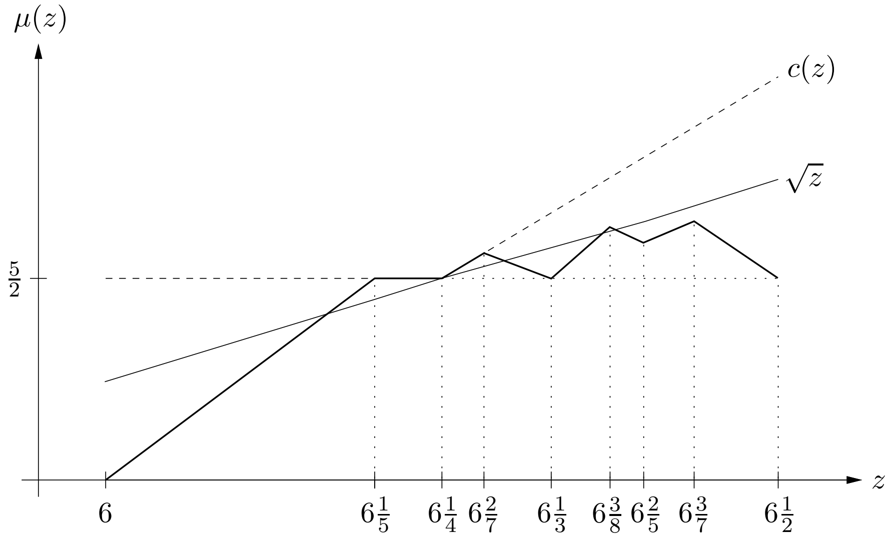

FIGURE 2.1. The graph of μ on $\left[6, 6\frac{1}{2}\right]$.

Lemma 2.3.3. *Consider $(d;\boldsymbol{m}) \in \mathcal{E}$ and $a = \frac{p}{q}$ (in lowest terms) such that $\ell(\boldsymbol{m}) = \ell(a) =: M$. Let $d\mu(d;\boldsymbol{m})(z) = A + Bz$ on a nonempty interval of the form $(a-\varepsilon, a)$. Then there is $\varepsilon' > 0$ so that for $z' \in (a, a+\varepsilon')$*

$$d\,\mu(d;\boldsymbol{m})(z') \;=\; A + Bz' + m(p-qz') \;=:\; A' + B'z'.$$

Proof. Suppose first that N is odd. Then by Lemma 2.2.1 when $z < a$, $x_j(z) = \alpha_j^a + z\beta_j^a$, for $j \le N+1$. Further,

$$x_{N+1}(z) \;=\; x_{N-1}(z) - \ell_N x_N(z) \;=\; \alpha_{N+1}^a + z\beta_{N+1}^a = p - qz \tag{2.3.1}$$

where the last equality holds by Lemma 2.2.3.

When z' is just larger than a, its Nth multiplicity is $\ell_N - 1$, and ℓ_{N+1} is very big. (Since $\ell_N \ge 2$ this still gives an allowed set of multiplicities.) Hence for such z' the formula for the linear functions $x_j(z'), j \le N$, is unchanged, but now $x'_{N+1}(z') = x_{N-1}(z') - (\ell_N - 1)x_N(z')$. (For clarity we denote by x'_{N+1} the formula that holds for $z' > a$ and by x_{N+1} the formula that holds for $z < a$.) Note that because $\ell(\boldsymbol{m}) = \ell(a)$, just one term from the $(N+1)$st block is counted in $\mu(d;\boldsymbol{m})(z')$. Hence, with $m := m_M$, we have

$$\begin{aligned} d\mu(d;\boldsymbol{m})(z') - (A + Bz') &= -m x_N(z') + m x'_{N+1}(z') \\ &= m\big(-x_N(z') + x_{N-1}(z') - (\ell_N - 1)x_N(z')\big) \\ &= m\big(x_{N-1}(z') - \ell_N x_N(z')\big) = m(p - qz'), \end{aligned}$$

where the last equality uses equation (2.3.1).

Now suppose that N is even. Then the formulas $x_j(z') := \alpha_j^a + z\beta_j^a$ give the (beginning of the) weight expansion for z' just larger than a. As above, when z is just less than a, we must modify the last multiplicities of a, reducing ℓ_N by 1, and making ℓ_{N+1} arbitrarily large. Thus as above, the formulas for the weights $x_j(z)$, $j \le N$, are unchanged but that for the $(N+1)$st weight is modified. As above we denote by x'_{N+1} the formula that holds for $z' > a$ and by x_{N+1} the formula that holds for $z < a$. Then $x'_{N+1}(z') = -p + qz' > 0$. Further, if $d\mu(d;\boldsymbol{m})(z') = A' + Bz'$ for $z' > a$, we find for $z < a$ that

$$\begin{aligned} d\mu(d;\boldsymbol{m})(z) - (A' + B'z) &= -mx_N(z) + mx_{N+1}(z) \\ &= m\big(-x_N(z) + x_{N-1}(z) - (\ell_N - 1)x_N(z)\big) \\ &= m\,x'_{N+1}(z) \;=\; -m(p - qz). \end{aligned}$$

Therefore $A + Bz = A' + B'z - m(p - qz)$, as claimed. □

To complete the proof of Proposition 2.3.2 we need to estimate the size of A, B. Here is an auxiliary lemma.

Lemma 2.3.4. *Let $\ell_0; \ell_1, \dots, \ell_N$ be any sequence of positive integers with $\ell_N \ge 2$, and let η_j, $j \ge 0$, be one of the sequences $|\alpha^a_j|$, $|\beta^a_j|$. Then $\sum_{j=0}^N \ell_j|\eta_j|^2 < \frac{1}{2}|\eta_{N+1}|^2$.*

Proof. By definition $\eta_j = \eta_{j-2} + \ell_{j-1}\eta_{j-1}$. The inequality

$$\ell_k \left(\sum_{j=0}^{k} \ell_j \eta_j^2 \right) \le \eta_{k+1}^2 \tag{2.3.2}$$

holds for $k = 0$, and may be proved for all larger k by induction. Setting $k = N$ yields the lemma. □

Proof of Proposition 2.3.2. Suppose first that N is odd, and write $m_i = \frac{d}{\sqrt{a}} w_i(a) + \varepsilon_i$ as in equation (2.1.1). For notational convenience, let us first assume that the m_i are constant on each of the blocks of length ℓ_j. Then define n_j to be this constant value on the jth block. If this assumption holds, then the ε_i are also constant on the blocks, and we denote their values by δ_j. Then $d\,\mu(\boldsymbol{m}; d)(a) = A + Ba$ where, by Lemma 2.2.1, we have

$$A = \sum \ell_j\, n_j\, \alpha^a_j, \qquad B = \sum \ell_j\, n_j\, \beta^a_j.$$

Therefore, substituting $n_j = \frac{d}{\sqrt{a}} x_j + \delta_j$, we find

$$\begin{aligned} A &= \sum \ell_j\, n_j\, \alpha^a_j \;=\; \sum \tfrac{d}{\sqrt{a}} \ell_j x_j \alpha^a_j + \sum \ell_j \delta_j \alpha^a_j \\ &= 0 + \sum \ell_j \delta_j \alpha^a_j \\ &\le \Big(\sum \ell_j \delta_j^2\Big)^{1/2} \Big(\sum \ell_j |\alpha^a_j|^2\Big)^{1/2} \;<\; \sqrt{E/2}\,|\alpha^a_{N+1}| \;<\; p. \end{aligned}$$

Here we used Corollary 2.2.7 for the third equality, and for the inequalities used the Cauchy–Schwarz inequality, $\sum \ell_j \delta_j^2 =: E < 1$ from Proposition 2.1.1, Lemma 2.3.4 and finally the fact that $|\alpha^a_{N+1}| = p$ from Lemma 2.2.3 (ii).

This argument is also valid if the m_i are not constant on the blocks. In this case, by Lemma 2.1.7 the values of n_j and δ_j may vary by 1 over the entries of one block, but that variation can be absorbed into the sum that gives $\sqrt{E}$ and will not increase it above $\sqrt{E+1} < \sqrt{2}$.

Similarly,

$$\begin{aligned} B &= \sum \ell_j\, n_j\, \beta_j^a \;=\; \sum \tfrac{d}{\sqrt{a}} \ell_j x_j \beta_j^a + \sum \ell_j \delta_j \beta_j^a \\ &= \tfrac{d}{\sqrt{a}} + \sum \ell_j \delta_j \beta_j^a \;=:\; \tfrac{d}{\sqrt{a}} + S, \end{aligned}$$

where $S := \sum_{j\le N} \ell_j \delta_j \beta_j^a$. By definition, $m = m_M = \frac{d}{\sqrt{a}}\, x_N + \delta_N$, and $x_N = \frac{1}{q}$. Therefore, assuming that the m_i are constant on the blocks we have

$$B - (m+1)q \;=\; \tfrac{d}{\sqrt{a}} - qm - q + S \;=\; -q(1+\delta_N) + S.$$

We need to show that $S < q(1+\delta_N) = |\beta^a_{N+1}|\,(1+\delta_N)$. If $\delta_N \ge 0$ we may estimate S as before by

$$S \;\le\; \sqrt{E}\Big(\sum_{j=0}^{N} \ell_j (\beta_j^a)^2\Big)^{1/2} \;<\; \tfrac{1}{\sqrt{2}} |\beta^a_{N+1}| \;<\; q.$$

Now assume that $\delta_N = -\delta$ is negative and note that $\beta_N^a > 0$ because N is odd. Therefore

$$S \;:=\; \sum_{j\le N} \ell_j \delta_j \beta_j^a \le -\ell_N \delta \beta_N^a + \sqrt{E}\Big(\sum_{j=0}^{N-1} \ell_j (\beta_j^a)^2\Big)^{1/2} \;\le\; \beta_N^a\big(\sqrt{E} - \ell_N \delta\big),$$

where we used equation (2.3.2) with $k = N-1$ and $\ell_{N-1} \ge 1$. Since $\beta_N^a < |\beta^a_{N+1}|/2 = q/2$ by the inductive formula, the desired result follows easily.

Suppose now that the m_i are not constant on the jth block. If $j < N$, then, as before, we simply need to replace E by $E+1$ in the above estimates. It is easy to check that the argument still goes through.

It remains to consider the case when the m_i are not constant on the last block. Define δ_N again by $m = m_M = \frac{d}{\sqrt{a}} x_N + \delta_N$. By Lemma 2.1.7 the last block of $\boldsymbol{m}$ is either $(m+1)^{\times \ell}, m$ with errors $(\delta_N+1)^{\times \ell}, \delta_N$; or $m+1, m^{\times \ell}$ with errors $\delta_N + 1, \delta_N^{\times \ell}$, where $\ell := \ell_N - 1$. Note that $\delta_N =: -\delta$ is negative. The sum S_N of $\delta_i \beta_i^a$ over the last block is either $\big(\ell(1-\delta) - \delta\big)\beta_N^a$ or $\big((1-\delta) - \ell\delta\big)\beta_N^a$. Since $\ell \ge 1$, in either case $S_N \le \big(\ell(1-\delta) - \delta\big)\beta_N^a$. But because $(\ell+1)\beta_N^a < \big|\beta^a_{N+1}\big| = q$, we can estimate $B - (m+1)q$ as follows:

$$\begin{aligned} B - (m+1)q &= -q(1-\delta) + \textstyle\sum_{j<N} \ell_j \delta_j \beta_j^a + S_N \\ &\le -\beta_N^a\Big((\ell+1)(1-\delta) - \sqrt{E} - \ell(1-\delta) + \delta\Big) < 0. \end{aligned}$$

This completes the proof when N is odd. The case when N is even is similar, and is left to the reader. □

2.4. **Connection to the lattice counting problem.** In this section we prove Theorem 1.1.3, stating that $c_{ECH}(a) \geq c(a)$ for all $a \geq 1$. Recall that for $a \geq 1$,

$$c_{ECH}(a) \;:=\; \inf\{\mu > 0 \mid N(1,a) \preccurlyeq N(\mu,\mu)\}\,.$$

The first step is to describe c_{ECH} in another way. As Hutchings pointed out[6], the inequalities $N(1,a) \preccurlyeq N(\mu,\mu)$ can be understood in terms of counting lattice points in triangles, as follows. Let $a \geq 1$ be irrational. For each pair of integers $A, B \geq 0$, consider the closed triangle

$$T^a_{A,B} \;:=\; \big\{(x,y) \in \mathbb{R}^2 \mid x, y \geq 0,\; x + ay \leq A + aB\big\}.$$

Thus the slant edge of $T^a_{A,B}$ has slope $-\frac{1}{a}$ and passes through the integral point (A,B). Then the number $\#\big(T^a_{A,B} \cap \mathbb{Z}^2\big)$ of integer points in the triangle $T^a_{A,B}$ is just the number of elements in $N(1,a)$ that are $\leq A + Ba$. We define

$$k_{A,B}(a) \;:=\; \tfrac{A+Ba}{d}, \tag{2.4.1}$$

where d is the smallest positive integer such that

$$\#\big(T^a_{A,B} \cap \mathbb{Z}^2\big) \leq \tfrac{1}{2}(d+1)(d+2).$$

(Note that $N(1,1) = (0,1,1,2,2,2,3,3,3,3,4,\dots)$ has precisely $\frac{1}{2}(d+1)(d+2)$ entries that are $\leq d$.) Further, set

$$K(a) \;:=\; \sup_{A,B \geq 0} \big\{k_{A,B}(a)\big\}.$$

We extend the function K to rational a by defining

$$K(a) \;:=\; \sup_{z<a,\, z \text{ irrat}} K(z). \tag{2.4.2}$$

Lemma 2.4.1. *$K(a) = c_{ECH}(a)$ for all $a \geq 1$.*

Proof. For each $\lambda > 1$ we have $N(1,\lambda a) \preccurlyeq \lambda N(1,a)$. Therefore, the conclusions of Lemma 1.1.1 hold for c_{ECH} as well as for c. In particular, c_{ECH} is continuous and nondecreasing. Therefore, (2.4.2) also holds for c_{ECH}. It hence suffices to prove the lemma for irrational a.

Fix an irrational a. If $c_{ECH}(a) < K(a)$ then one can find a rational number $\mu > c_{ECH}(a)$ and non-negative integers A, B with $\mu < k_{A,B}(a)$. Since $\mu > c_{ECH}(a)$ we have $N(1,a) \preccurlyeq N(\mu,\mu)$. This inequality implies that for all non-negative integers A, B we have

$$\#\{p \in N(1,a) \mid p \leq A + Ba\} \;\geq\; \#\{p \in N(\mu,\mu) \mid p \leq A + Ba\}\,.$$

The number on the left is $\#\big(T^a_{A,B} \cap \mathbb{Z}^2\big)$, while the number on the right is $\frac{1}{2}(D+1)(D+2)$, where $D := \lfloor \frac{A+Ba}{\mu} \rfloor$. This must be a strict inequality for some A, B. To see this, let $u = (u_1, u_2, u_3, \dots)$ be the sequence of natural numbers obtained by arranging in increasing order all the numbers on the LHS obtained by running through all pairs of integers $A, B \geq 0$. Since a is irrational, each number in $N(1,a)$ occurs with multiplicity 1. The definition of $N(1,a)$ therefore shows that $u = (1,2,3,\dots)$. On the other hand, the

[6] Private communication.

numbers in $N(\mu,\mu)$ appear with larger and larger multiplicity. The sequence obtained in this way from the RHS therefore jumps by larger and larger amounts.

Consider A, B such that this is a strict inequality. Then $k_{A,B}(a) = \frac{A+Ba}{d}$ where $d > D$. On the other hand because a is irrational and μ is rational, $D+1 > \frac{A+Ba}{\mu} > D$, so that

$$k_{A,B}(a) \;=\; \tfrac{A+Ba}{d} \;\le\; \tfrac{A+Ba}{D+1} \;<\; \mu.$$

Since this contradicts our assumptions, we conclude that $c_{ECH}(a) \ge K(a)$.

To complete the proof, it suffices to show that $K(a) \ge \mu$ for all $\mu < c_{ECH}(a)$. For such μ we have $N(1,a) \not\preccurlyeq N(\mu,\mu)$. Therefore there is A, B such that

$$\#\left\{p \in N(1,a) \mid p \le A+Ba\right\} \;<\; \#\left\{p \in N(\mu,\mu) \mid p \le A+Ba\right\}.$$

With D as before, this implies that $d \le D$, so that $k_{A,B}(a) = \frac{A+Ba}{d} \ge \frac{A+Ba}{D} \ge \mu$. Hence $K(a) = \sup k_{A,B}(a) \ge \mu$ as required. □

We are now going to prove Theorem 1.1.3 by direct calculation, showing that for each of the constraints $(d;\boldsymbol{m})$ that contributes to $c(a)$ there is a triangle that contributes to $K(a)$ in exactly the same way. Therefore we will assume the results of Theorems 1.1.2 and 5.2.3.

The key to understanding the relation between the functions $k_{A,B}$ of equation (2.4.1) and the number of lattice points in the triangles $T^a_{A,B}$ is the following lemma, that was explained to us by Hutchings.

Lemma 2.4.2. *Suppose that a is rational, abbreviate $T := T^a_{A,B}$, and suppose that*

$$\#(T \cap \mathbb{Z}^2) \;\le\; \tfrac12(d+1)(d+2) + s - 1 \;=\; \tfrac12(d^2+3d) + s,$$

where $s \ge 1$ is the number of integral points on the slant edge of T. Assume that (A,B) (resp. (A',B')) is the integral point on the slant edge with smallest (resp. largest) x-coordinate. Then there is $\varepsilon > 0$ such that

$$K(z) \ge \tfrac{A+zB}{d} \;\; \textit{if } z \in (a-\varepsilon, a], \qquad K(z) \ge \tfrac{A'+zB'}{d} \;\; \textit{if } z \in [a, a+\varepsilon).$$

Proof. Recall that c_{ECH} and hence K is continuous. To prove the statement for $z < a$ it therefore suffices to consider irrational z of the form $z = a - \varepsilon$. Then, for small enough $\varepsilon > 0$, the triangle $T^z_{A,B}$ contains $s-1$ fewer integral points than T. Therefore $k_{A,B}(z) \ge \frac{A+zB}{d}$, which proves the first statement. Similarly, the second statement holds because if $z = a+\varepsilon$ is irrational and $\varepsilon > 0$ is sufficiently small, the triangle $T^z_{A',B'}$ contains $s-1$ fewer integral points than T. □

Lemma 2.4.3. *$K(b_n) \ge \sqrt{a_{n+1}}$ for all $n \ge 1$.*

Proof. Consider the triangle $T_n \subset \mathbb{R}^2$ with vertices $(0,0)$, $(g_{n+2},0)$ and $(0,g_n)$, where g_n is the nth odd Fibonacci number. Because g_n, g_{n+2} are mutually prime and satisfy the identities

$$g_n + g_{n+2} = 3g_{n+1}, \qquad g_n g_{n+2} = g_{n+1}^2 + 1$$

(see Section 3.1), we find that

$$\#(T_n \cap \mathbb{Z}^2) = \tfrac{1}{2}(g_n+1)(g_{n+2}+1)+1 = \tfrac{1}{2}(g_{n+1}^2+3g_{n+1})+2. \tag{2.4.3}$$

Since $b_n = \frac{g_{n+2}}{g_n}$, we have $T_n = T^{b_n}_{0,g_n}$. In view of (2.4.3) we can apply Lemma 2.4.2 with $s=2$ and $d = g_{n+1}$: For some $\varepsilon > 0$ we have

$$K(z) \geq \tfrac{z g_n}{g_{n+1}} \quad \text{when} \quad z \in (b_n - \varepsilon, b_n],$$

and

$$K(z) \geq \tfrac{g_{n+2}}{g_{n+1}} \quad \text{when} \quad z \in [b_n, b_n+\varepsilon).$$

In particular, $K(b_n) \geq \frac{g_{n+2}}{g_{n+1}} = \sqrt{a_{n+1}}$. □

Corollary 2.4.4. *$K(a) \geq c(a)$ for all $a \in [1, \tau^4]$.*

Proof. First observe that c is the smallest continuous and nondecreasing function on $[1,\tau^4]$ that is $\geq \sqrt{a}$, has the scaling property of Lemma 1.1.1, and also satisfies $c(b_n) = \sqrt{a_{n+1}}$. On the other hand, we already remarked that c_{ECH} and hence K is continuous, nondecreasing and has the scaling property. It is also easy to see that $K(a) \geq \sqrt{a}$, because the number of integer points in a large triangle approximates its area. Therefore $K(b_n) \geq c(b_n)$ implies that $K(a) \geq c(a)$ over the whole interval. □

Proof of Theorem 1.1.3. By Corollary 2.4.4 we only need to show $K \geq c$ on the interval $[\tau^4, \infty)$. Since, as remarked there, $K(a) \geq \sqrt{a}$ for all a, we just need to check that $K(a) \geq \mu(d;\boldsymbol{m})(a)$ for all $(d;\boldsymbol{m})$ that contribute to c. Recall from the proof of Proposition 1.2.9 that the class $(3;2,1^{\times 6})$ gives the constraint $\frac{a+1}{3}$ on $[\tau^4, 7]$. Together with Theorems 1.1.2 (ii) and 5.2.3, we see that it suffices to check $K(a) \geq \mu(d;\boldsymbol{m})(a)$ for the nine classes in Table 2.4.4 below. Each of these classes contributes on both sides of its center point. It suffices to show that in each case there is a triangle that gives an equal constraint. Proposition 2.3.2 shows which triangles to take: if the constraint $(d;\boldsymbol{m})$ is centered at a, then one should consider $T := T^a_{A,B} = T^a_{A',B'}$ where $\mu(d;\boldsymbol{m})(z)$ equals $\frac{1}{d}(A+Bz)$ to the left of a and $\frac{1}{d}(A'+B'z)$ to the right. Because $c = \mu(d;\boldsymbol{m})$ in a neighborhood of the center point, this proposition together with Corollary 2.1.4 implies that $0 \leq A < p$ and $mq \leq B < (m+1)q$, so that the integral points (A,B) and $(A',B') = (A+mp, B-mq)$ are the first and last on the slant edge of T, as required by Lemma 2.4.2. Therefore, it suffices to check that in each case the coefficient d occurring in $(d;\boldsymbol{m})$ satisfies the condition in Lemma 2.4.2. Thus the number $N(A,B)$ of integer points in T must be $\leq N(d) := \frac{1}{2}(d+1)(d+2)+s-1$, where $s = m+1$ is the number of points on the slant edge of T. In fact, as the following table shows we find that

$N(A,B) = N(d)$ in each case.

(2.4.4)

a	$(d;\boldsymbol{m})$	(A,B)	(A',B')	$N(A,B)$	s	$N(d)$
7	$(3;2,1^{\times 6})$	$(1,1)$	$(8,0)$	11	2	11
$7\frac{1}{8}$	$(48;18^{\times 7},3,2^{\times 7})$	$(7,17)$	$(121,1)$	1227	3	1227
$7\frac{2}{15}$	$(64;24^{\times 7},3^{\times 7},1^{\times 2})$	$(14,22)$	$(121,7)$	2146	2	2146
$7\frac{1}{7}$	$(24;9^{\times 7},2,1^{\times 6})$	$(7,8)$	$(57,1)$	326	2	326
$7\frac{2}{13}$	$(40;15^{\times 7},2^{\times 6},1^{\times 2})$	$(14,13)$	$(107,0)$	862	2	862
$7\frac{1}{5}$	$(16;6^{\times 7},1^{\times 5})$	$(7,5)$	$(43,0)$	154	2	154
$7\frac{1}{4}$	$(35;13^{\times 7},4,3^{\times 3})$	$(0,13)$	$(87,1)$	669	4	669
$7\frac{1}{2}$	$(8;3^{\times 7},1^{\times 2})$	$(7,2)$	$(22,0)$	46	2	46
8	$(6;3,2^{\times 7})$	$(1,2)$	$(17,0)$	30	3	30

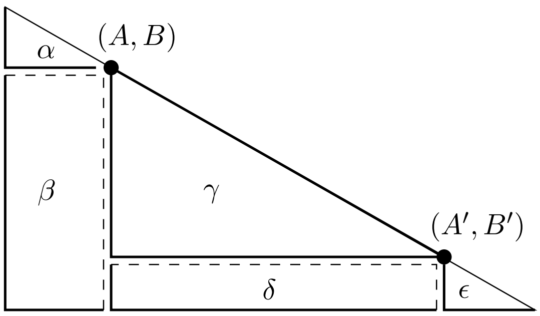

FIGURE 2.2. The subdivision of the triangle $T^a_{A,B}$.

Here we calculate $N(A,B)$ by subdividing T into five parts labeled $\alpha,\dots,\varepsilon$ as in Figure 2.2. Each part besides γ is half open, and includes the integer points on the heavy boundary edges but not those on the dashed boundary edges. For example, the rectangle β includes the integer points on the x and y-axes, but not those on the (dashed) edges shared by α, γ or δ. Thus $\#(\beta\cap\mathbb{Z}^2) = AB$. Further, because A,B (resp. (A',B')) is the integer point on the slant edge with smallest (resp. largest) x coordinate, we put all integer points on the slant edge into γ. Thus, we find that

$$\#(\gamma\cap\mathbb{Z}^2) \;=\; \tfrac{1}{2}\Big((A'-A+1)(B'-B+1)-s\Big)+s.$$

For example, in the case of the triangle $T^a_{7,17}$ with $a=7\frac{1}{8}$ and $s=3$, the numbers of integer points in $\alpha,\dots,\varepsilon$ are $7,119,979,114$ and 8, giving a total of 1227.

This completes the proof of Theorem 1.1.3. □

Remark 2.4.5. On the interval $[7,8]$, there are four other classes (described in Table 5.2.12) with the property that $\mu(d;\boldsymbol{m})(a) = c(a)$ at their center points $a = \frac{p}{q}$, but that do not contribute otherwise to $c(a)$. Let us look at their contribution to K. In each case $(A,B) = (-1, mq)$, where m is the last nonzero entry in $\boldsymbol{m}$, so that $\mu(d;\boldsymbol{m})$ does not satisfy the scaling condition to the left. In the corresponding triangles the first point on the slant edge is $(A_1, B_1) = (-1+p, (m-1)q)$ and one can check as before that in each case $s = m$ and $N(A,B) = N(d)$. Therefore to the left of each center point we obtain the inequality

$$K(a) \ge k_{A_1,B_1}(a) = \frac{p-1+(m-1)qa}{d}.$$

In each case, one can check that $k_{A_1,B_1}(a)$ is precisely $c(a)$. For example, at $7\frac{1}{8} = \frac{57}{8} =: \frac{p}{q}$, we get $\frac{56+17\cdot 8a}{384} = \frac{7+17a}{48}$ which agrees with the first line in Table 5.2.1.

(ii) Recall from the introduction that the functorial properties of embedded contact homology establish that $c_{ECH}(a) \le c(a)$ for all a. For $a \le 6\frac{1}{4}$, $a = 9$ and for $a \ge 11$ one can prove this by directly showing that for each d the closed triangle with vertices $(0,0)$, $(d\,c(a), 0)$, $(0, d\,\frac{c(a)}{a})$ contains at least $\frac{1}{2}(d+1)(d+2)$ lattice points. For extensions of such arguments see [18]. ◇

3. The Fibonacci stairs.

In this section we establish the behavior of $c(a)$ for $a \le \tau^4$.

3.1. Main results.

Recall that the Fibonacci numbers f_n for $n \ge 0$ are recursively defined by

$$f_0 = 0,\ f_1 = 1 \quad \text{and} \quad f_{n+1} = f_n + f_{n-1},\ n \ge 1. \tag{3.1.1}$$

Denote by $g_n = f_{2n-1}$, $n \ge 1$, the sequence of *odd* Fibonacci numbers. The sequence g_n starts with

$$1,\ 2,\ 5,\ 13,\ 34,\ 89,\ 233,\ 610,\ 1597,\ 4181,\ 10946,\ \dots$$

The recursion formula $f_{n+1} = f_n + f_{n-1}$ implies the recursion formula

$$g_{n+1} = 3g_n - g_{n-1}. \tag{3.1.2}$$

Using this and induction we find that

$$g_n^2 + 1 = g_{n-1}g_{n+1}. \tag{3.1.3}$$

Set

$$a_n = \left(\frac{g_{n+1}}{g_n}\right)^2 \quad \text{and} \quad b_n = \frac{g_{n+2}}{g_n}.$$

Then $\dots < a_n < b_n < a_{n+1} < b_{n+1} < \dots$. Since $\lim_{n\to\infty} \frac{f_{n+1}}{f_n} = \frac{1+\sqrt{5}}{2} =: \tau$, we have

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \tau^4 \approx 6.8541.$$

They key to establishing the Fibonacci stairs is the following result that states that there are elements in $\mathcal{E}$ corresponding to the points a_n, b_n.

Theorem 3.1.1. (i) *Let $W(b_n) = g_n\,\boldsymbol{w}(b_n)$. Then $E(b_n) := \big(g_{n+1}; W(b_n)\big) \in \mathcal{E}$.*

(ii) *Let $W'(a_n)$ be the tuple obtained from $W(a_n) := g_n^2\,\boldsymbol{w}(a_n)$ by adding an extra* 1 *at the end. Then $E(a_n) := \big(g_n g_{n+1}; W'(a_n)\big) \in \mathcal{E}$.*

Corollary 3.1.2. *Part (i) of Theorem 1.1.2 holds.*

Proof of Corollary. Since $E(b_n)$ is a perfect element, Lemma 2.1.5 (i) shows that

$$c(b_n) \;=\; \mu\big(g_{n+1}; W(b_n)\big)(b_n) \;=\; \tfrac{g_n}{g_{n+1}}\boldsymbol{w}(b_n)\cdot\boldsymbol{w}(b_n) \;=\; \tfrac{g_n}{g_{n+1}}b_n \;=\; \tfrac{g_{n+2}}{g_{n+1}} \;=\; \sqrt{a_{n+1}}.$$

Suppose that $c(a_n) > \sqrt{a_n}$ for some n. Then Corollary 1.2.3 implies that there is $(d;\boldsymbol{m}) \in \mathcal{E}$ such that

$$\boldsymbol{m}\cdot\boldsymbol{w}(a_n) \;>\; d\,\sqrt{a_n}.$$

Note that $(d;\boldsymbol{m}) \neq \big(g_n g_{n+1}; W'(a_n)\big)$ since $W'(a_n)\cdot\boldsymbol{w}(a_n) \;=\; g_n^2 a_n \;=\; g_n g_{n+1}\sqrt{a_n}$. Therefore by positivity of intersections (part (ii) of Proposition 1.2.12) we must have

$$d g_n g_{n+1} \;\geq\; \boldsymbol{m}\cdot W'(a_n) \;\geq\; g_n^2\,\boldsymbol{m}\cdot\boldsymbol{w}(a_n), \qquad \text{i.e.}\quad d\,\sqrt{a_n} \;\geq\; \boldsymbol{m}\cdot\boldsymbol{w}(a_n).$$

It follows that $c(a_n) = \sqrt{a_n}$ for all n. Thus $c(b_n) = \sqrt{a_{n+1}} = c(a_{n+1})$. Moreover,

$$\tfrac{c(b_n)}{b_n} \;=\; \tfrac{\sqrt{a_{n+1}}}{b_n} \;=\; \tfrac{1}{\sqrt{a_n}} \;=\; \tfrac{c(a_n)}{a_n}.$$

Hence c is linear on the interval $[a_n, b_n]$ by the scaling property. □

Corollary 3.1.3. *The classes $E(b_n)$ are the only perfect elements.*

Proof. On $[1,\tau^4]$, $c(a)$ is given by the Fibonacci stairs. By Lemma 2.1.5 (i), the perfect element $E(b_n)$ is the only class giving the constraint $c(b_n)$ at b_n. This and Proposition 2.3.2 show that the step of the stairs over $[a_n, a_{n+1}]$ centered at b_n is the constraint μ given by the perfect element $E(b_n)$. Lemma 2.1.5 (i) now shows that there cannot be another perfect element on $[1,\tau^4]$. By (ii) of Lemma 2.1.5 there is no perfect element on $[\tau^4,\infty)$. □

We now turn to the proof of Theorem 3.1.1. The proof of part (i) is relatively easy; it is deferred to Corollary 4.2.3 since it is a special case of Proposition 4.2.2. To prove part (ii) we first need to show that the elements $E(a_n)$ satisfy the appropriate Diophantine equations, which is accomplished in Lemma 3.1.4. Second, we must check that $E(a_n)$ reduces correctly under Cremona moves. As we see in Section 3.3, the reduction process is quite complicated (and in fact is much more complicated than for the $E(b_n)$), basically because the weight expansions $\boldsymbol{w}(a_n)$ involve quadratic rather than linear functions in the Fibonacci numbers. The intermediate Section 3.2 collects basic identities on Fibonacci numbers and explains an inductive procedure useful for checking identities on them.

Lemma 3.1.4. *The tuples $E(a_n) := \big(g_n g_{n+1}; W'(a_n)\big)$ have integer entries and satisfy equations* (1.2.4)*.*

Proof. Consider $\big(g_n g_{n+1}; W'(a_n)\big)$. Since $a_n = \frac{g_{n+1}^2}{g_n^2}$, it follows from Lemma 1.2.6 that the last entry w_M of $\boldsymbol{w}(a_n)$ is $\frac{1}{g_n^2}$. Therefore, the terms in $W(a_n) = g_n^2 \boldsymbol{w}(a_n)$ and hence in $W'(a_n)$ are all integers. Next,

$$\begin{aligned}\sum_i W_i'(a_n) &= g_n^2 \left(\sum_i w_i\right) + 1 \;=\; g_n^2\left(a_n + 1 - \tfrac{1}{g_n^2}\right) + 1\\ &= g_{n+1}^2 + g_n^2 \;=\; g_{n+1}^2 + g_{n+1}g_{n-1} - 1\\ &= g_{n+1}\left(g_{n+1} + g_{n-1}\right) - 1 \;=\; 3g_{n+1}g_n - 1.\end{aligned}$$

Finally, $W'(a_n) \cdot W'(a_n) = g_n^4\, a_n + 1 = (g_n g_{n+1})^2 + 1$. This completes the proof. □

3.2. **Identities for Fibonacci numbers.** The proof that the classes $E(a_n)$ reduce correctly involves many small calculations. To avoid having to do them explicitly, we first explain a general inductive procedure whose conclusions are summarized in Proposition 3.2.3. It is based on the following elementary result. Recall that f_k denotes the kth Fibonacci number defined in (3.1.1).

Lemma 3.2.1. *Given any three distinct numbers $s_0, s_1, s_2 \ge 0$, there are rational constants λ, μ such that $f_{s_2+j} = \lambda f_{s_0+j} + \mu f_{s_1+j}$ for all $j \ge 0$.*

Proof. The equations

$$\lambda f_{s_0} + \mu f_{s_1} = f_{s_2}, \qquad \lambda f_{s_0+1} + \mu f_{s_1+1} = f_{s_2+1}$$

have a unique solution because $\frac{f_{s_0}}{f_{s_0+1}} \ne \frac{f_{s_1}}{f_{s_1+1}}$ when $s_0 \ne s_1$. Now apply the defining relation (3.1.1). □

We will frequently use the following relations between Fibonacci numbers.

$$f_k^2 = f_{k+1}f_{k-1} - (-1)^k \tag{3.2.1}$$

$$f_{2k-1} = f_k^2 + f_{k-1}^2 \tag{3.2.2}$$

$$f_{2k} = f_{k+1}^2 - f_{k-1}^2 \tag{3.2.3}$$

Lemma 3.2.2. *For each $i \ge 0$ and $s \ge 0$, there is an identity of the form*

$$f_{s+i} f_s \;=\; \sum_{j\ge 0} a_{ij} f_{2s+j} + (-1)^s c_i,$$

with a finite number of coefficients $c_i, a_{ij} \in \mathbb{Q}$ that do not depend on s. Further, $c_i = -\sum_{j\ge 0} a_{ij} f_j$.

Proof. By (3.1.1) it suffices to prove this for $i = 0$ and $i = 2$. We claim that for all $s \ge 0$,

$$5 f_s f_s = -f_{2s} + 2f_{2s+1} - 2(-1)^s, \tag{3.2.4}$$

$$5 f_{s+2} f_s = f_{2s+1} + f_{2s+3} - 3(-1)^s. \tag{3.2.5}$$

For $s = 0$, equation (3.2.4) is true. For $s \ge 1$, equation (3.2.4) can be rewritten as

$$5 f_s^2 \;=\; f_{2s+1} + f_{2s-1} - 2(-1)^s.$$

By (3.2.2), the RHS is $f_{s+1}^2 + 2f_s^2 + f_{s-1}^2 - 2(-1)^s$, whence we need to check

$$3f_s^2 \;=\; f_{s+1}^2 + f_{s-1}^2 - 2(-1)^s. \tag{3.2.6}$$

Replacing f_{s+1} by $f_s + f_{s-1}$, this becomes $2f_s^2 = 2f_s f_{s-1} + 2f_{s-1}^2 - 2(-1)^s$, which is true since by (3.2.1), $f_s^2 = f_{s+1}f_{s-1} - (-1)^s = f_s f_{s-1} + f_{s-1}^2 - (-1)^s$.

By (3.2.1) and (3.2.2), equation (3.2.5) becomes

$$5f_{s+1}^2 + 5(-1)^{s+1} \;=\; f_{s+2}^2 + 2f_{s+1}^2 + f_s^2 - 3(-1)^s,$$

i.e.,

$$3f_{s+1}^2 \;=\; f_{s+2}^2 + f_s^2 - 2(-1)^{s+1}$$

which is true by (3.2.6). The formula for c_i holds because $f_0 = 0$. □

Proposition 3.2.3. *A quadratic identity of the form*

$$Q(s) \;:=\; \sum_{i,j\ge 0} a_{ij} f_{s+i} f_{s+j} + \sum_{j\ge 0} b_j f_{2s+j} + (-1)^s c \;=\; 0$$

holds for all $s \ge 0$ if it holds for any three distinct values of s. Moreover, if the relation is homogeneous and linear (that is, if $a_{ij} = c = 0$ for all i, j), then it suffices to check two values of s.

Proof. Suppose that $Q(s) = 0$ for $s = s_0, s_1, s_2$ where $0 \le s_0 < s_1 < s_2$. By Lemma 3.2.2 one can convert $Q(s)$ to an equivalent identity of the form

$$Q'(k) \;:=\; \sum_{j\ge 0} a_j f_{k+j} + (-1)^s c' = 0 \qquad \text{where } k = 2s.$$

We first claim that $c' = 0$. By Lemma 3.2.1 there are constants μ, λ such that

$$f_{2s_2+j} \;=\; \mu f_{2s_1+j} + \lambda f_{2s_0+j} \qquad \text{for all } j \ge 0. \tag{3.2.7}$$

Since $Q'(k) = 0$ for $k = 2s_0, 2s_1, 2s_2$, and by (3.2.7),

$$c' \;=\; \big((-1)^{s_0+s_2}\mu + (-1)^{s_1+s_2}\lambda\big)\, c'.$$

If $c' \neq 0$, we thus have $1 \in \{\pm\mu \pm \lambda\}$. This is impossible: We use the recurrence relation (3.1.1) to extend the sequence $(f_n), n \ge 0$, to negative index n. Note that $f_{-n} = (-1)^{n+1} f_n$. Then (3.2.7) holds for all $j \in \mathbb{Z}$. With $j = -2s_0$ and $j = -2s_1$ we get

$$\mu = \frac{f_{2(s_2-s_0)}}{f_{2(s_1-s_0)}}, \qquad \lambda = -\frac{f_{2(s_2-s_1)}}{f_{2(s_1-s_0)}}.$$

Therefore, $\pm\mu \pm \lambda = 1$ exactly if

$$\pm f_{2(s_2-s_0)} \;=\; \pm f_{2(s_2-s_1)} + f_{2(s_1-s_0)}.$$

The signs $++$ are impossible because $f_{m+n} = f_{m+n-1} + f_{m+n-2} > f_m + f_n$ for all even $m, n > 0$. Further, $+-$, $-+$, and $--$ are impossible because $s_2 > s_1 > s_0 > 0$.

We have shown that $Q'(k) = \sum_{j\ge 0} a_j f_{k+j}$. Recall that $Q'(k) = 0$ for $k = 2s_0, 2s_1, 2s_2$. By Lemma 3.2.1, the expression for $Q'(0)$ can be written as a linear combination of the expressions for $Q'(2s_0)$ and $Q'(2s_1)$, and the same is true for $Q'(1)$. Therefore,

$Q'(0) = 0$ and $Q'(1) = 0$. This and the defining relation (3.1.1) show that $Q'(k) = 0$ for all k. In particular, $Q'(k) = 0$ for all even k, and so $Q(s) = 0$ for all s. This proves the first statement. The second holds similarly. □

In the subsequent sections, the following abbreviations will be useful.

Definition 3.2.4. *The kth Lucas number is defined to be* $\ell_k = f_{k-1} + f_{k+1},\ k \geq 1$. *We set* $F_k := \frac{1}{3} f_{4k}$ *and* $L_k := \frac{1}{3}\ell_{4k+2}$.

Then

$$F_1 = 1,\ F_2 = 7,\ F_3 = 48,\ F_4 = 329,$$
$$L_0 = 1,\ L_1 = 6,\ L_2 = 41,\ L_3 = 281,\ L_4 = 1926.$$

Further for $k \geq 0$ define the sequence H_k by

$$H_k = \tfrac{1}{3} f_{2k} f_{2k+2}. \tag{3.2.8}$$

Then $H_0 = 0,\ H_1 = 1,\ H_2 = 8,\ H_3 = 56,\ H_4 = 385,\ H_5 = 2640, \dots$.

Lemma 3.2.5. *The following identities hold for all* $k \geq 0$*:*

(i) $F_{k+1} = L_k + F_k$;
(ii) $L_{k+1} = 5\,F_{k+1} + L_k$;
(iii) $H_{k+1} = H_k + F_{k+1} = \sum_{i=1}^{k+1} F_i$;
(iv) $L_k = 5H_k + 1$.
(v) $F_{k+1}^2 - F_k F_{k+2} = 1$.

Proof. The second identity in (iii) follows from the first identity in (iii) by induction. All other identities have the form considered in Proposition 3.2.3, and so it is enough to check each of them for at most three low values of k. □

3.3. **Reducing** $E(a_n)$**.** We begin with a general remark about the reduction process.

Remark 3.3.1. Consider a tuple $(d; \boldsymbol{m})$ that satisfies the Diophantine identities (1.2.4). Proposition 1.2.12 states that $(d; \boldsymbol{m}) \in \mathcal{E}$ exactly if it reduces to $(0; -1, 0, \dots, 0)$ under standard Cremona moves, as defined in Definition 1.2.11. In fact, it clearly suffices to reduce $(d; \boldsymbol{m})$ to a known element of $\mathcal{E}$ by any sequence of Cremona moves. Each such move consists of an application of the Cremona transformation

$$(d; \boldsymbol{m}) \mapsto (2d - m_1 - m_2 - m_3; d - m_2 - m_3, d - m_1 - m_3, d - m_1 - m_2, \dots)$$

followed by a choice of reordering. It does not matter whether this reordering restores the natural order; all that is important is that in the end, after doing many such moves, we arrive at a known element of $\mathcal{E}$. In fact, the reorderings chosen below all do restore the natural order. The point of this remark is that there is no need to *prove* this. ◇

Example 3.3.2. The first few elements $E(a_n)$ are

$$\begin{aligned}
\big(g_2 g_3; W'(a_2)\big) &= \big(10; 4^{\times 6}, 1^{\times 5}\big),\\
\big(g_3 g_4; W'(a_3)\big) &= \big(65; 25^{\times 6}, 19^{\times 1}, 6^{\times 3}, 1^{\times 7}\big),\\
\big(g_4 g_5; W'(a_4)\big) &= \big(442; 169^{\times 6}, 142^{\times 1}, 27^{\times 5}, 7^{\times 3}, 6^{\times 1}, 1^{\times 7}\big).
\end{aligned}$$

These values of n are too low for our general arguments in §3.3.1 and 3.3.2 to apply. Hence one proves that they reduce correctly by direct calculation.

The following list of Fibonacci numbers will be useful in the subsequent proofs.

n	0	1	2	3	4	5	6	7	8	10	12
f_n	0	1	1	2	3	5	8	13	21	55	144

3.3.1. *Reducing $E(a_n)$ for even n.* Throughout this subsection we will consider $n = 2m \geq 2$ to be a fixed even number. We will obtain an explicit expression for $W'(a_n)$ and then examine its reduction by Cremona moves. By Example 3.3.2 it suffices to consider the case $n \geq 6$. Hence this case of Theorem 3.1.1 follows from Propositions 3.3.6, 3.3.9 and 3.3.10.

For each fixed n, denote $k' := n - k$ and define

$$\begin{aligned} u_k \; &:= \; f^2_{2n-2k-1} + 2\,H_k \;=\; f^2_{2k'-1} + 2\,H_k \quad k = 0, \dots, m-1, \\ v_k \; &:= \; 3\,F_{n-k} - 2\,F_k = 3\,F_{k'} - 2\,F_k, \qquad k = 1, \dots, m. \end{aligned} \tag{3.3.1}$$

Note that u_k, v_k depend on n, though for simplicity the notation does not make this explicit. Also, $v_k > 0$ for all $k \leq m$ and $v_m = F_m$. However, the above formula for v_k gives a negative number when $k > m$. This is why the expansion in Proposition 3.3.3 below changes its form at the term v_m. Note also that by equations (3.2.1) and (3.2.8)

$$f^2_{2n+1} \;=\; f_{2n} f_{2n+2} + 1 \;=\; 3H_n + 1. \tag{3.3.2}$$

Hence we also have

$$u_k \;=\; 3\,H_{k'-1} + 2\,H_k + 1. \tag{3.3.3}$$

Proposition 3.3.3. *If $n = 2m$ is even, then the continued fraction expansion of a_n is*

$$[6; \underbrace{1,\, 5}_{m-1},\, 3,\, 1, \underbrace{5,\, 1}_{m-1}] \;=:\; [6; \{1,5\}^{\times(m-1)}, 3,\, 1, \{5,1\}^{\times(m-1)}],$$

and the (renormalized) weight expansion $W(a_n)$ is (A_n, B_n), where the vectors A_n, B_n are

$$\begin{aligned} A_n \;&=\; \big({u_0}^{\times 6}, {v_1}^{\times 1}, {u_1}^{\times 5}, {v_2}^{\times 1}, \dots, {u_{m-2}}^{\times 5}, {v_{m-1}}^{\times 1}, {u_{m-1}}^{\times 5}\big), \\ B_n \;&=\; \big({F_m}^{\times 3}, {L_{m-1}}^{\times 1}, {F_{m-1}}^{\times 5}, {L_{m-2}}^{\times 1}, \dots, {F_2}^{\times 5}, {L_1}^{\times 1}, {F_1}^{\times 5}, {L_0}^{\times 1}\big). \end{aligned}$$

Remark 3.3.4. Although our usual convention is that the expression $a = [\ell_0; \ell_1, \dots, \ell_N]$ always has $\ell_N \geq 2$, we relax this condition here in order to simplify the formulas. We allow ourselves another liberty at the end of these expansions: in the formula for B_n in Proposition 3.3.3 the last two weights F_1, L_0 are equal, so the ending multiplicity is in fact 6 rather than 5, 1. ◇

Proof. Recall that $a_n = \left(\frac{g_{n+1}}{g_n}\right)^2 = \frac{f^2_{2n+1}}{f^2_{2n-1}}$. Also, because $b_{n-1} < a_n < b_n$, the weight expansion $\boldsymbol{w}(a_n)$ begins as $(1^{\times 6}, (a_n - 6)^{\times 1}, \dots)$. Hence

$$W(a_n) = f^2_{2n-1} \boldsymbol{w}(a_n) = \big({f^2_{2n-1}}^{\times 6}, f^2_{2n+1} - 6\,f^2_{2n-1}, \dots\big).$$

Therefore, because $v_m = F_m$ as noted above, the expansion of $W(a_n)$ up to and including the term L_{m-1} follows from Lemma 3.3.5 below. The rest of the expansion holds by Lemma 3.2.5. □

Lemma 3.3.5. *Let $n = 2m \geq 4$. The following identities hold.*

(i) $f_{2n+1}^2 = 6\,u_0 + v_1$ *and* $v_1 < u_0$*;*
(ii) $u_k = v_{k+1} + u_{k+1}$ *and* $u_{k+1} < v_{k+1}$ *for* $k = 0, \dots, m-2$*;*
(iii) $v_k = 5\,u_k + v_{k+1}$ *and* $v_{k+1} < u_k$ *for* $k = 1, \dots, m-1$*;*
(iv) $u_{m-1} = 3\,F_m + L_{m-1}$ *and* $L_{m-1} < F_m$*.*

Proof. Statement (i) is equivalent to

$$f_{2n+1}^2 = 6\,f_{2n-1}^2 + 3F_{n-1} - 2. \tag{3.3.4}$$

Since $f_{s+1}^2 = 6f_{s-1}^2 + f_{2s-4} - 2(-1)^s$ holds true for $s = 2, 3, 4$, this equation holds true for all $s \geq 2$ by Proposition 3.2.3; in particular it holds true for all even $s \geq 2$, and so (3.3.4) holds true.

The equality in (ii) follows from the definitions of u_k, v_k by using (3.3.3) and $H_k = \sum_{i=1}^k F_i$ and by dividing the equations into two equations, one for k and one for $k' := n - k$. To prove the equation in (iii) we again divide it into two equations, one for k and one for $k' := n - k$, namely

$$-2F_k = 10H_k + 2 - 2F_{k+1}, \qquad 3F_{k'} = 5f_{2k'-1}^2 - 2 + 3F_{k'-1}.$$

The first equation is equivalent to $L_k = 5H_k + 1$, which holds by Lemma (3.2.5) (iv), and the second holds because $f_{2s} = 5f_{s-1}^2 - 2(-1)^s + f_{2s-4}$ is true for $s = 2, 3, 4$ and hence for all $s \geq 2$ by Proposition 3.2.3.

Equation (3.3.3) and Lemma 3.2.5 imply that

$$u_{m-1} = 3H_m + 2H_{m-1} + 1 = 3F_m + 5H_{m-1} + 1 = 3F_m + L_{m-1}.$$

This proves (iv).

Since u_k, v_k are positive in the given ranges, the equalities in (ii) and (iv) imply the inequalities in (i) and (iii). Similarly, the equalities in (iii) imply the inequalities in (ii). This completes the proof. □

The reduction process has three steps that are described in Propositions 3.3.6, 3.3.9 and 3.3.10. Notice, before we begin, that the weights of a_n divide into three groups, namely $(m-1)$ pairs $\{1, 5\}$, two central blocks with multiplicities $3, 1$, and finally $(m-1)$ pairs $\{5, 1\}$. In the first step we show that a set of 5 Cremona moves has the effect of moving the first $\{1, 5\}$ pair from the left to a $\{5, 1\}$ pair on the right. Moreover, doing this introduces no new weights to the right while slightly modifying the first block on the left.

Denote $V_n = E(a_n) = (g_n g_{n+1}; W(a_n), 1)$. Note that

$$g_n g_{n+1} = f_{2n-1} f_{2n+1} = f_{2n}^2 + 1 = f_{2n}^2 + f_1^2$$

by (3.2.1).

Proposition 3.3.6. *For $n=2m\ge 6$, the vector V_n is reduced by $5(m-1)$ Cremona moves to the vector $V_n^1=\Big(f^2_{2(m+1)}+f^2_{2m-1};\,A_n^1,B_n^1\Big)$, where*

$$\begin{aligned} A_n^1 :&= \big((u_{m-1}+F_{m-1})^{\times 1},\,{u_{m-1}}^{\times 5}\big),\\ B_n^1 :&= \big({F_m}^{\times 3},\,{L_{m-1}}^{\times 1},\,{F_{m-1}}^{\times 10},\,{L_{m-2}}^{\times 2},\dots,\,{F_2}^{\times 10},\,{L_1}^{\times 2},\,{F_1}^{\times 10},\,{L_0}^{\times 2},\,1^{\times 1}\big). \end{aligned}$$

Let

$$V(n,1):=\big(f^2_{2n}+1;A(n,1),B(n,1)\big):=\big(f^2_{2n}+1;A_n,B_n\big)=V_n,$$

and for $k=2,\dots,m$ define the vector $V(n,k)$ by $\Big(f^2_{2(n-k+1)}+f^2_{2k-1};A(n,k),B(n,k)\Big)$, where

$$\begin{aligned} A(n,k) &= \big((u_{k-1}+F_{k-1})^{\times 1},{u_{k-1}}^{\times 5},{v_k}^{\times 1},{u_k}^{\times 5},\dots,,{v_{m-1}}^{\times 1},{u_{m-1}}^{\times 5}\big),\\ B(n,k) &= \big({F_m}^{\times 3},{L_{m-1}}^{\times 1},\dots,{F_k}^{\times 5},{L_{k-1}}^{\times 1},{F_{k-1}}^{\times 10},{L_{k-2}}^{\times 2},\dots,\\ &\qquad {F_1}^{\times 10},{L_0}^{\times 2},1^{\times 1}\big). \end{aligned}$$

Then $V(n,1)=V_n$ and $V(n,m)=V_n^1$. Moreover, $A(n,k+1)$ is obtained from $A(n,k)$ by replacing its first seven entries by the single entry $u_k+F_k=f^2_{2(n-k)-1}+2\,H_k+F_k$; and $B(n,k+1)$ is obtained from $B(n,k)$ by inserting $\big({F_k}^{\times 5},L_{k-1}\big)$. Proposition 3.3.6 will follow if we prove:

Lemma 3.3.7. *For $1\le k\le m-1$, $V(n,k)$ reduces to $V(n,k+1)$ by 5 Cremona moves.*

We prove Lemma 3.3.7 in five steps. Throughout, we abbreviate $n-k$ to k'. Thus $k'>k$.

Remark 3.3.8. (i) Each step of the reduction involves many small calculations that can be done directly using identities such as (3.2.1), (3.2.2) and (3.2.3). However, in all cases the required identity is quadratic in the sense of Proposition 3.2.3. Hence they can be all proved by verifying them for just three low values of s, as we have already illustrated in the proof of Lemma 3.3.5. The only condition on the choice of s is that all subscripts of the f_i should be ≥ 0.

(ii) In each step of the reduction process there is no interaction between the terms in k and those in k'; we simplify each set of terms separately. ◇

Step 1: *There is a Cremona move that takes $V(n,k)$ to*

$$\begin{aligned} V_1(n,k) \;:=\; \big(& f_{4k'+2}-F_{k-1};\,{u_{k-1}}^{\times 3},\,v_k,\,f_{2k'+1}f_{2k'}-H_{k-1},\\ &(f_{2k'+1}f_{2k'}-H_{k-1}-F_{k-1})^{\times 2},\,{u_k}^{\times 5},\dots\big). \end{aligned}$$

Proof. The first component of the Cremona transform of $V(n,k)$ is $2\Big(f^2_{2(k'+1)}+f^2_{2k-1}\Big)-3\,u_{k-1}$, and we must show that it equals $f_{4k'+2}$. Since $u_{k-1}:=f^2_{2k'+1}+2H_{k-1}$, we need to see that

$$2\,f^2_{2(k'+1)}-3\,f^2_{2k'+1}-f_{4k'+2}\;=\;6\,H_{k-1}-2\,f^2_{2k-1}\;=\;2\big(f_{2k-2}f_{2k}-f^2_{2k-1}\big)\,.$$

But both sides are equal to -2. This is clear for the RHS by (3.2.1). For the LHS, note that $2f_{s+2}^2 - 3f_{s+1}^2 - f_{2s+2} = -2(-1)^s$, since this is true for $s = -1, 0, 1$ and hence for all $s \geq -1$ by Proposition 3.2.3.

The second three terms of the Cremona transform of $V(n,k)$ equal $\left(f_{2(k'+1)}^2 + f_{2k-1}^2\right) - 2\,u_{k-1}$, and one can check as above that this is $f_{2k'+1}f_{2k'} - H_{k-1}$. Therefore the given vector $V_1(n,k)$ is a reordering of the Cremona transform of $V(n,k)$. □

Step 2: *There is a Cremona move that takes* $V_1(n,k)$ *to*

$$\begin{aligned} V_2(n,k) \;&:=\; \Big(f_{4k'} + f_{2k'+1}f_{2k'} - H_k - F_{k-1};\; v_k,\; f_{2k'+1}f_{2k'} - H_{k-1}, \\ &\qquad (f_{2k'+1}f_{2k'} - H_{k-1} - F_{k-1})^{\times 2}, (f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1})^{\times 3},\; {u_k}^{\times 5}, \dots\Big). \end{aligned}$$

Proof. For the first component we need to see that

$$2\,(f_{4k'+2} - F_{k-1}) - 3\,u_{k-1} \;=\; f_{4k'} + f_{2k'+1}f_{2k'} - H_k - F_{k-1}.$$

But this is equivalent to the identity

$$2\,f_{4k'+2} - 3\,f_{2k'+1}^2 - f_{4k'} - f_{2k'+1}f_{2k'} \;=\; 6\,H_{k-1} + F_{k-1} - H_k, \tag{3.3.5}$$

and one can check that both sides here equal -1.

The Cremona transform also contains three terms of the form $f_{4k'+2} - F_{k-1} - 2\,u_{k-1}$, and we need to check that this is $f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1}$. But this holds because

$$f_{4k'+2} - 2\,f_{2k'+1}^2 - f_{2k'}f_{2k'-1} \;=\; \tfrac{4}{3}f_{2k-2}f_{2k} - f_{2k}f_{2k-1} + \tfrac{1}{3}f_{4k-4} \;=\; -1.$$

Thus $V_2(N,k)$ is a reordering of the Cremona transform of $V_1(n,k)$. □

Step 3: *There is a Cremona move that takes* $V_2(n,k)$ *to*

$$V_3(n,k) \;:=\; \big(f_{4k'} - F_{k-1};\; f_{2k'+1}f_{2k'} - H_{k-1} - F_{k-1},\; (f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1})^{\times 4},\; {u_k}^{\times 5}, \dots\big)$$

where the multiplicities of F_k *and* L_{k-1} *are each increased by one to* ${F_k}^{\times 6}$, ${L_{k-1}}^{\times 2}$.

Proof. Since $H_k = H_{k-1} + F_k$, the first term

$$2\,(f_{4k'} + f_{2k'+1}f_{2k'} - H_k - F_{k-1}) - v_k - 2\,(f_{2k'+1}f_{2k'} - H_{k-1}) + F_{k-1}$$

of the Cremona transform of $V_2(n,k)$ is equal to

$$2\,f_{4k'} - 2\,F_k - F_{k-1} - (f_{4k'} - 2\,F_k) \;=\; f_{4k'} - F_{k-1}.$$

Its second term

$$(f_{4k'} + f_{2k'+1}f_{2k'} - H_k - F_{k-1}) - 2\,(f_{2k'+1}f_{2k'} - H_{k-1}) + F_{k-1}$$

simplifies to

$$f_{4k'} - f_{2k'+1}f_{2k'} - F_k + H_{k-1} \;=\; f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1},$$

where the last equality follows from the identities

$$f_{4k'} - f_{2k'+1}f_{2k'} - f_{2k'}f_{2k'-1} = 0, \quad F_k - H_{k-1} - f_{2k}f_{2k-1} = 0. \tag{3.3.6}$$

The third term of the Cremona transform of $V_2(n,k)$ is

$$(f_{4k'} + f_{2k'+1}f_{2k'} - H_k - F_{k-1}) - v_k - (f_{2k'+1}f_{2k'} - H_{k-1} - F_{k-1}).$$

This is equal to F_k. In the same way, we find that its fourth term is $F_k - F_{k-1} = L_{k-1}$. The result follows immediately. □

Step 4: *There is a Cremona move that takes $V_3(n,k)$ to*

$$\begin{aligned} V_4(n,k) \;:=\; \big(& f_{2k'+1}f_{2k'} - H_{k-1} - F_{k-1} + 2\,F_k;\; f_{2k'}^2 + f_{2k-1}^2 + F_k, \\ & (f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1})^{\times 2},\, {u_k}^{\times 5}, \dots \big) \end{aligned}$$

where the multiplicities of F_k and L_{k-1} are now ${F_k}^{\times 8}$, ${L_{k-1}}^{\times 2}$.

Proof. By (3.3.6), the first term of the Cremona transform of $V_3(n,k)$ is

$$f_{2k'+1}f_{2k'} - H_{k-1} - F_{k-1} + 2\,F_k.$$

We claim that its second term is

$$f_{4k'} - F_{k-1} - 2\,(f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1}) \;=\; f_{2k'}^2 + f_{2k-1}^2 + F_k.$$

This follows from

$$f_{4k'} - 2\,f_{2k'}f_{2k'-1} - f_{2k'}^2 = 0, \quad F_{k-1} - 2\,f_{2k}f_{2k-1} + f_{2k-1}^2 + F_k = 0. \tag{3.3.7}$$

Finally, the third and fourth term of the Cremona transform of $V_3(n,k)$ are

$$f_{4k'} - F_{k-1} - (f_{2k'+1}f_{2k'} - H_{k-1} - F_{k-1}) - (f_{2k'}f_{2k'-1} - f_{2k}f_{2k-1}) \;=\; F_k$$

where the equality holds by (3.3.6). Hence $V_4(n,k)$ is a reordering of the Cremona transform, as claimed. □

Step 5: *There is a Cremona move that takes $V_4(n,k)$ to*

$$V(n,k+1) \;:=\; \big(f_{2k'}^2 + f_{2k+1}^2;\; u_k + F_k,\, {F_k}^{\times 2},\, {u_k}^{\times 5}, \dots \big) \tag{3.3.8}$$

where the multiplicities of F_k and L_{k-1} are ${F_k}^{\times 10}$, ${L_{k-1}}^{\times 2}$.

Proof. The above expression for the first term of the Cremona transform of $V_4(n,k)$ follows from the identities

$$2\,f_{2k'+1}f_{2k'} - 2\,f_{2k'}^2 - 2\,f_{2k'}f_{2k'-1} = 0,$$
$$2\,H_{k-1} + 2\,F_{k-1} - 3\,F_k + f_{2k-1}^2 - 2\,f_{2k}f_{2k-1} + f_{2k+1}^2 = 0.$$

(The second identity can be simplified by subtracting the second identity of (3.3.7).)

We next claim that the second term of this transform is $u_k + F_k$. Since $u_k = f_{2k'-1}^2 + 2\,H_k$ and $H_{k-1} + F_k = H_k$, this is equivalent to the identity

$$f_{2k'+1}f_{2k'} - 2\,f_{2k'}f_{2k'-1} - f_{2k'-1} \;=\; 3\,H_{k-1} + F_{k-1} + F_k - 2\,f_{2k}f_{2k-1}.$$

But both sides equal -1.

Finally, its third and fourth terms are F_k because

$$f_{2k'+1}f_{2k'} - f_{2k'}^2 - f_{2k'}f_{2k'-1} \;=\; H_{k-1} + F_{k-1} + f_{2k-1}^2 - f_{2k}f_{2k-1}.$$

Here both sides vanish. □

This completes the proof of Lemma 3.3.7 and hence of Proposition 3.3.6. The next stage of the reduction process results in a vector V_n^2 whose components are linear (rather than quadratic) functions of the f_k and do not involve the index k'.

Proposition 3.3.9. *When $n = 2m \geq 6$, the vector V_n^1 may be reduced by four Cremona moves to*

$$V_n^2 = \left(F_m{}^{\times 1}; L_{m-1}{}^{\times 1}, F_{m-1}{}^{\times 11}, L_{m-2}{}^{\times 2}, \ldots, F_2{}^{\times 10}, L_1{}^{\times 2}, F_1{}^{\times 13} \right).$$

Proof. Note that V_n^2 is obtained from B_n^1 by removing two copies of F_m and adding one F_{m-1}.

We first claim that the Cremona transform of V_n^1 is

$$\left(f_{4m+2} - F_{m-1};\ f_{4m-1} + F_{m-1},\ f_{4m-1}{}^{\times 2},\ u_{m-1}{}^{\times 3},\ B_n^1 \right),$$

which we reorder as

$$\left(f_{4m+2} - F_{m-1};\ u_{m-1}{}^{\times 3},\ f_{4m-1} + F_{m-1},\ f_{4m-1}{}^{\times 2},\ B_n^1 \right). \tag{3.3.9}$$

Here one obtains the first term of the transform from the identity

$$2\left(f_{2(m+1)}^2 + f_{2m-1}^2 \right) - 3\,u_{m-1} = f_{4m+2},$$

and the second term from

$$f_{2m+2}^2 + f_{2m-1}^2 - 2\,u_{m-1} = f_{4m-1} + F_{m-1}.$$

Next, observe that Lemma 3.3.5 (iv) implies that $2\,u_{m-1} = 6F_m + 2L_{m-1}$, which, as one can easily check, is just $f_{4m+2} - F_{m-1}$. Therefore, the second Cremona transform moves the vector (3.3.9) to

$$\left(u_{m-1};\ f_{4m-1} + F_{m-1},\ f_{4m-1}{}^{\times 2},\ B_n^1 \right). \tag{3.3.10}$$

We next claim that $u_{m-1} - 2\,f_{4m-1} = F_{m-1}$. Hence the third Cremona transform moves the vector (3.3.10) to

$$\left(f_{4m-1} + F_{m-1};\ F_{m-1},\ 0^{\times 2},\ B_n^1 \right),$$

which we reorder as

$$\left(2F_m; F_m{}^{\times 3},\ L_{m-1}{}^{\times 1},\ F_{m-1}{}^{\times 11},\ L_{m-2}{}^{\times 2},\ F_{m-2}{}^{\times 10},\ L_{m-2}{}^{\times 2}, \ldots,\ F_2{}^{\times 10},\ L_1{}^{\times 2},\ F_1{}^{\times 13} \right).$$

Therefore, another Cremona move takes the above vector to

$$V_n^2 = \left(F_m;\ L_{m-1}{}^{\times 1},\ F_{m-1}{}^{\times 11},\ L_{m-2}{}^{\times 2},\ F_{m-2}{}^{\times 10},\ L_{m-2}{}^{\times 2}, \ldots,\ F_2{}^{\times 10},\ L_1{}^{\times 2},\ F_1{}^{\times 13} \right).$$

This completes the proof of Proposition 3.3.9. □

Proposition 3.3.10. *For $n = 2m \geq 6$, the vector V_n^2 may be reduced to $(2; 1^{\times 5})$ by Cremona moves.*

Proof. One shows by direct calculation that this holds when $n = 6$. Therefore, by induction it suffices to show that the vector V_n^2 is reduced to V_{n-2}^2 by 6 Cremona moves. By using the identities in Lemma 3.2.5, it is not hard to prove directly that this reduction may be achieved by six standard Cremona moves. Alternatively, one can check numerically that this holds when $n = 8$ and 10, checking also that the reordering required is the same in both cases at each stage. Then it holds for all n by Proposition 3.2.3. (Note that we only need to check two values since all identities are homogeneous and linear.) □

3.3.2. *Reducing $E(a_n)$ for odd n.* Throughout this section we denote $n = 2m+1$, where $m \geq 1$. By Example 3.3.2 it suffices to consider the case $n \geq 5$. Hence this case of Theorem 3.1.1 follows from Propositions 3.3.13, 3.3.14 and 3.3.15.

We consider the numbers

$$u_k = f_{2n-2k-1}^2 + 2\,H_k, \quad k = 0, \dots, m, \qquad v_k = f_{4(n-k)} - 2\,F_k, \quad k = 1, \dots, m,$$

as before. Again we have $v_m > 0 > v_{m+1}$ but now $u_m = L_m$. Therefore Lemma 3.3.5 takes the following form.

Lemma 3.3.11. *If $n = 2m + 1$, the following identities hold.*

(i) $f_{2n+1}^2 = 6\,u_0 + v_1$ *and* $v_1 < u_0$*;*
(ii) $u_k = v_{k+1} + u_{k+1}$ *and* $u_{k+1} < v_{k+1}$ *for* $k = 0, \dots, m-1$*;*
(iii) $v_k = 5\,u_k + v_{k+1}$ *and* $v_{k+1} < u_k$ *for* $k = 1, \dots, m-1$*;*
(iv) $v_m = 3\,u_m + F_m = 3\,L_m + F_m$ *and* $F_m < L_m$.

Proof. The proofs of the equalities in (i), (ii) and (iii) go through as before, since these are based on equalities that do not mention m. (Note that the proof of the equality in (ii) works when $k = m - 1$, though the corresponding inequality failed for even n.) One then checks (iv). Then the inequality in (ii) follows from the equalities in (iii) and (iv), while the inequality in (iii) holds by (ii). □

As before, this lemma immediately gives the following result.

Proposition 3.3.12. *If $n = 2m + 1$ is odd, the continued fraction expansion of $a_n = \frac{f_{2n+1}^2}{f_{2n-1}^2}$ is*

$$[6;\, \{1,\, 5\}^{\times(m-1)},\, 1,\, 3,\, \{5,\, 1\}^{\times m}],$$

and the (renormalized) weight expansion $W(a_n)$ is $(\widehat{A}_n, \widehat{B}_n)$, where

$$\begin{aligned} \widehat{A}_n &:= \big(u_0{}^{\times 6},\, v_1{}^{\times 1}, u_1{}^{\times 5}, \dots, v_{m-1}{}^{\times 1},\, u_{m-1}{}^{\times 5}, v_m{}^{\times 1}\big), \\ \widehat{B}_n &:= \big((u_m = L_m)^{\times 3}, F_m{}^{\times 5}, L_{m-1}{}^{\times 1}, \dots, F_1{}^{\times 5}, L_0{}^{\times 1}\big). \end{aligned}$$

The proof of Proposition 3.3.6 also goes through. In other words, each set of five Cremona moves takes one of the $m - 1$ pairs $\{1, 5\}$ from the left to the right of the central blocks (which now have multiplicities $1, 3$), while introducing no new weights on the right. Since we start with m blocks on the right but only $(m - 1)$ on the left, this means that one pair $\{5, 1\}$ still remains on the right, though all the others become

$\{10, 2\}$. The only other difference is in the interpretation of the first term of $V(n, m)$: when $n = 2m + 1$ and $k = m$ we have $2(n - k + 1) = 2(m + 2)$. Thus we obtain:

Proposition 3.3.13. *For $n = 2m + 1 \geq 5$, the vector $\widehat{V}_n = (g_n g_{n+1}; W(a_n), 1)$ is reduced by $5(m-1)$ Cremona moves to the vector $\widehat{V}_n^1 = \left(f_{2m+4}^2 + f_{2m-1}^2;\, \widehat{A}_n^1, \widehat{B}_n^1\right)$, where*

$$\begin{aligned}\widehat{A}_n^1 &:= \big((u_{m-1} + F_{m-1})^{\times 1},\, {u_{m-1}}^{\times 5},\, {v_m}^{\times 1}\big),\\ \widehat{B}_n^1 &:= \big({L_m}^{\times 3},\, {F_m}^{\times 5},\, {L_{m-1}}^{\times 1},\, {F_{m-1}}^{\times 10}, \dots,\, {F_2}^{\times 10},\, {L_1}^{\times 2},\, {F_1}^{\times 10},\, {L_0}^{\times 2},\, 1^{\times 1}\big).\end{aligned}$$

Notice that the entries in $\widehat{V}_n^1$ still depend explicitly both on k and on $k' = n - k$ since u_{m-1} and v_{m-1} have this structure. This means that there are entries in $\widehat{V}_n^1$ that do not occur anywhere in the reduction of $\widehat{V}_{n+2}$. The next stage takes us to a vector that occurs in the reduction of all $\widehat{V}_{n+2i}$. For even n, this stage consisted of four moves, but now it takes six moves.

Proposition 3.3.14. *When $n = 2m + 1 \geq 5$, the vector $\widehat{V}_n^1$ is reduced by six Cremona moves to*

$$\widehat{V}_n^2 := \big(L_m - F_m; (L_m - 2F_m)^{\times 1}, {F_m}^{\times 7}, {L_{m-1}}^{\times 2}, {F_{m-1}}^{\times 10}, {L_{m-2}}^{\times 2}, \dots\big),$$

where the terms including and after ${F_{m-1}}^{\times 10}$ in $\widehat{V}_n^2$ are the same as those in $\widehat{B}_n^1$.

Proof. Since the proof is much the same as that of Proposition 3.3.9, we simply list the results of each Cremona move. Here are the results of the first four moves:

$$\begin{aligned}&\big(f_{4m+6} - F_{m-1}; {u_{m-1}}^{\times 3}, v_m, f_{4m+3} + f_{4m-1} + F_{m-1}, (f_{4m+3} + f_{4m-1})^{\times 2}, {L_m}^{\times 3}, \dots\big)\\ &\big(f_{4m+5} + L_{m-1}; v_m, f_{4m+3} + f_{4m-1} + F_{m-1}, (f_{4m+3} + f_{4m-1})^{\times 2}, {L_m}^{\times 3},\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad {f_{4m+1}}^{\times 3}, \dots\big)\\ &\big(f_{4m+4} - F_{m-1}; f_{4m+3} + f_{4m-1}, {L_m}^{\times 3}, {f_{4m+1}}^{\times 4}, F_m, L_{m-1} \dots\big)\\ &\big(f_{4m+3} + f_{4m-1}; f_{4m+3} + f_{4m-1} - L_m, L_m, {f_{4m+1}}^{\times 4}, F_m, L_{m-1}, \dots\big).\end{aligned}$$

None of these moves uses up any of the terms of $\widehat{B}_n^1$ after ${L_m}^{\times 3}$, though the multiplicity of F_m, L_{m-1} is increased. Hence after the entry for L_m we have simply listed the extra terms that get added to $\widehat{B}_n^1$. Note also there is just one new number that appears after L_m, namely $f_{4m+1} = L_m - F_m$ which appears with multiplicity 4. Using the same conventions, the next move gives:

$$\big(2f_{4m+1}; {f_{4m+1}}^{\times 3}, f_{4m+1} - F_m, {F_m}^{\times 2}, L_{m-1}, \dots\big).$$

The last move changes the first four terms to the single term $f_{4m+1} = L_m - F_m$. □

Proposition 3.3.15. *For $n = 2m + 1 \geq 5$, the vector $\widehat{V}_n^2$ may be reduced to $(2; 1^{\times 5})$ by Cremona moves.*

Proof. This is just the same as the proof of Proposition 3.3.10. □

4. The interval $[\tau^4, 7]$

This section is devoted to the calculation of $c(a)$ on the interval $[\tau^4, 7]$, thus completing the proof of part (ii) of Theorem 1.1.2. Proposition 1.2.9 gives an easy argument that $c(a) = \frac{a+1}{3}$ when $a \in [6\frac{11}{12}, 7\frac{1}{9}]$. To prove that this holds on the whole interval $[\tau^4, 7]$, we shall adapt the arithmetic approach that works for $a < \tau^4$ rather than using more analytical arguments as in the case $a > 7$. We show in Proposition 4.1.6 that it suffices to restrict attention to some special points with relatively short continued fraction expansions that are related to the convergents of τ^4, and then deal with these special points by largely arithmetic means. The proof that $c(a) = \frac{a+1}{3}$ on $[\tau^4, 7]$ is given at the end of §4.1.

4.1. **Reduction to special points.** As in Section 2.1, given a with weight expansion $\boldsymbol{w}(a)$ and $(d; \boldsymbol{m}) \in \mathcal{E}$, we define ε by $\boldsymbol{m} = \frac{d}{\sqrt{a}}\boldsymbol{w}(a) + \varepsilon$, and denote $E := \sum \varepsilon_i^2$. Also set $\lambda^2 := 1 - E$. Denote

$$y(a) := a + 1 - 3\sqrt{a}.$$

Since $y(\tau^4) = 0$ we have $y(a) > 0$ for $a > \tau^4$. The first result extends Proposition 2.1.1.

Proposition 4.1.1. *Let $a \in (\tau^4, 7)$ and suppose that $(d; \boldsymbol{m}) \in \mathcal{E}$ is such that $\mu(d; \boldsymbol{m})(a) > \frac{a+1}{3}$. Then*

(i) $d < \dfrac{3\sqrt{a}}{\sqrt{a^2 - 7a + 1}}$;

(ii) $\lambda^2 > \dfrac{2d^2 y(a)}{3\sqrt{a}}$.

Proof. By Proposition 2.1.1 (i) we have $\frac{a+1}{3} < \sqrt{a}\sqrt{1 + 1/d^2}$, proving (i).

Since $\frac{a+1}{3} < \boldsymbol{m} \cdot \boldsymbol{w}/d = \sqrt{a} + \varepsilon \cdot \boldsymbol{w}/d$, we have $y(a)d/3 < \varepsilon \cdot \boldsymbol{w}$. Further,

$$d^2 + 1 \;=\; \boldsymbol{m} \cdot \boldsymbol{m} \;=\; \tfrac{d^2}{a}a + \tfrac{2d}{\sqrt{a}}\boldsymbol{w}(a) \cdot \varepsilon + \varepsilon \cdot \varepsilon,$$

i.e., $\varepsilon \cdot \boldsymbol{w}(a) = \frac{\lambda^2\sqrt{a}}{2d}$, proving (ii). □

The continued fraction expansion of $\tau^4 = \frac{7+3\sqrt{5}}{2}$ is $[6; 1, 5, 1, 5, \dots]$. For $k \geq 1$ define its kth convergent c_k by

$$\begin{aligned} c_{2k-1} &:= \big[6; \{1,5\}^{\times(k-1)}, 1\big], = \big[6; \{1,5\}^{\times(k-2)}, 1, 6\big], \\ c_{2k} &:= \big[6; \{1,5\}^{\times k}\big]. \end{aligned}$$

Thus

$$\begin{aligned} &c_1 = 7, \;\; c_2 = 6\tfrac{5}{6} = \tfrac{41}{6}, \;\; c_3 = [6; 1, 5, 1] = [6; 1, 6] = 6\tfrac{6}{7} = \tfrac{48}{7}, \\ &c_4 = [6; 1, 5, 1, 5] = 6\tfrac{35}{41} = \tfrac{281}{41}, \;\; c_5 = [6; 1, 5, 1, 6] = 6\tfrac{41}{48} = \tfrac{329}{48}, \end{aligned} \tag{4.1.1}$$

and more generally

$$c_{2k} < c_{2k+2} < \tau^4 < c_{2k+1} < c_{2k-1} \quad \text{for all } \; k \geq 1.$$

Moreover, for $k \geq 1$ and $j \geq 1$ define the numbers $u_k(j), v_k(j) \in (c_{2k+1}, c_{2k-1})$ by

$$\begin{aligned} u_k(j) &:= \left[6; \{1,5\}^{\times(k-1)}, 1, 6, j\right], \\ v_k(j) &:= \left[6; \{1,5\}^{\times(k-1)}, 1, j\right]. \end{aligned}$$

As in Definition 3.2.4, we define for $k \geq 1$ the kth Lucas number $\ell_k = f_{k-1} + f_{k+1}$, and set $F_k = \frac{1}{3} f_{4k}$ and $L_k = \frac{1}{3} \ell_{4k+2}$. Recall from Lemma 3.2.5 that for all $k \geq 0$,

$$F_{k+1} = L_k + F_k \quad \text{and} \quad L_{k+1} = 5\, F_{k+1} + L_k \tag{4.1.2}$$

as well as

$$F_{k+1}^2 - F_k F_{k+2} \;=\; 1. \tag{4.1.3}$$

Lemma 4.1.2. *For all $k \geq 1$,*

(i) $c_{2k-1} = \dfrac{F_{k+1}}{F_k}; \qquad c_{2k} = \dfrac{L_{k+1}}{L_k};$

(ii) $u_k(j) \;=\; \dfrac{F_{k+1} + jF_{k+2}}{F_k + jF_{k+1}}$ *for all* $j \geq 1$*;*

(iii) $v_k(j) \;=\; \dfrac{L_k + jF_{k+1}}{L_{k-1} + jF_k} \;=\; \dfrac{(j-6)F_{k+1} + F_{k+2}}{(j-6)F_k + F_{k+1}}$ *for all* $j \geq 1$*.*

Proof. Recall that if p_n/q_n is the nth convergent to $[\ell_0; \ell_1, \dots, \ell_N]$ then, for any $n < N$ and any positive $x \in \mathbb{R}$ we have

$$[\ell_0; \ell_1, \dots, \ell_{n-1}, x] \;=\; \frac{p_{n-2} + xp_{n-1}}{q_{n-2} + xq_{n-1}}. \tag{4.1.4}$$

(i) follows from induction on k: The statement is true for $k = 1$. Assume it holds for k. By (4.1.4) with $x = 1$, and by (4.1.2),

$$c_{2k+1} \;=\; \frac{F_{k+1} + L_{k+1}}{F_k + L_k} \;=\; \frac{F_{k+2}}{F_{k+1}}.$$

Then, by (4.1.4) with $x = 5$, and by (4.1.2),

$$c_{2k+2} \;=\; \frac{L_{k+1} + 5F_{k+2}}{L_k + 5F_{k+1}} \;=\; \frac{L_{k+2}}{L_{k+1}}.$$

(ii) follows from (i) by using $c_{2k+1} = [6; \{1,5\}^{\times(k-1)}, 1, 6]$.
(iii) follows from (i) and (4.1.2). □

Corollary 4.1.3. *For all $k \geq 1$, $s \geq 3$ and $t \geq 8$ we have*

$$c_{2k+1} < u_k(s+1) < u_k(s) < \cdots < u_k(2) < u_k(1) = v_k(7) < v_k(t) < v_k(t+1) < c_{2k-1}.$$

Proof. This follows from Lemma 4.1.2 and identity (4.1.3). □

The following corollary will be very useful.

Corollary 4.1.4. *(i) Let $u = u_k(j) =: \frac{p}{q}$, where $j \geq 1$. Then*

$$q^2\left(u^2 - 7u + 1\right) \;=\; j^2 + 7j + 1.$$

(ii) Let $v = v_k(j) =: \frac{p}{q}$, where $j \geq 1$. Then

$$q^2\left(v^2 - 7v + 1\right) \;=\; j^2 - 5j - 5.$$

Proof. The proofs of (i) and (ii) are similar. We prove (ii). In view of Lemma 4.1.2 (iii) we need to show

$$(L_k + jF_{k+1})^2 - 7\,(L_{k-1} + jF_k)\,(L_k + jF_{k+1}) + (L_{k-1} + jF_k)^2 \;=\; j^2 - 5j - 5. \tag{4.1.5}$$

Fix j. Identity (4.1.5) is true for $k = 1, 2, 3$. It therefore holds for all k by Proposition 3.2.3. □

Definition 4.1.5. *We say that a point $a \in [\tau^4, 7]$ is* **regular** *if for all $(d; \boldsymbol{m}) \in \mathcal{E}$ with $\ell(\boldsymbol{m}) = \ell(a)$ we have $\mu(d; \boldsymbol{m})(a) \leq \frac{a+1}{3}$.*

Proposition 4.1.6. *Assume that all the points c_{2k-1} and all the points*

$$u_k(j) \;\text{ with } k \geq 1 \text{ and } j \geq 2 \qquad \text{and} \qquad v_k(j) \;\text{ with } k \geq 1 \text{ and } j \geq 7$$

are regular. Then $c(a) = \frac{a+1}{3}$ on $[\tau^4, 7]$.

A main ingredient in the proof will be the following

Lemma 4.1.7. *Consider the functions $\varphi(a) := \frac{a+1}{3}$ and $\psi(a) := \sqrt{a}$. Fix $k \geq 1$. Then*

(i) *$\varphi\big(u_k(j+1)\big) > \psi\big(u_k(j)\big)$ for all $j \geq 1$;*

(ii) *$\varphi\big(v_k(j)\big) > \psi\big(v_k(j+1)\big)$ for all $j \geq 7$.*

Proof. (i) Abbreviate $u := u_k(j+1)$, $u' = u_k(j)$. We need to show that $\frac{u+1}{3} > \sqrt{u'}$, i.e.,

$$u^2 + 2u + 1 \;>\; 9u'. \tag{4.1.6}$$

Recall from Lemma 4.1.2 (iii) that

$$u = \frac{(j+1)F_{k+2} + F_{k+1}}{(j+1)F_{k+1} + F_k}, \qquad u' = \frac{jF_{k+2} + F_{k+1}}{jF_{k+1} + F_k}. \tag{4.1.7}$$

In particular, the denominator of u is $q := (j+1)F_{k+1} + F_k$. Applying Corollary 4.1.4 (i) to u we therefore find

$$u^2 \;=\; \frac{(j+1)^2 + 7(j+1) + 1}{q^2} + 7u - 1$$

and so (4.1.6) is equivalent to

$$9(u' - u)q^2 \;<\; (j+1)^2 + 7(j+1) + 1. \tag{4.1.8}$$

Using (4.1.7) and $F_{k+1}^2 - F_kF_{k+2} = 1$ we compute

$$u' - u \;=\; \frac{1}{\big((j+1)F_{k+1} + F_k\big)\big(jF_{k+1} + F_k\big)}.$$

Inequality (4.1.8) therefore becomes

$$9\,\frac{(j+1)F_{k+1}+F_k}{jF_{k+1}+F_k} \;<\; (j+1)^2+7(j+1)+1. \tag{4.1.9}$$

For all $j \geq 1$ the second factor on the LHS is < 2 and the RHS is ≥ 19, and so (4.1.9) holds for all $j \geq 1$.

The proof of (ii) is similar (but slightly easier). □

Proof of Proposition 4.1.6: Assume that $c(a) \leq \frac{a+1}{3}$ does not hold on $[\tau^4, 7]$. Since $c(a) \geq \frac{a+1}{3} > \sqrt{a}$ on $(\tau^4, 7]$, Corollary 2.1.4 shows that $c(a)$ is piecewise linear on $(\tau^4, 7]$. Let $S \subset (\tau^4, 7)$ be the set of non-smooth points of c on $(\tau^4, 7)$. This set decomposes as $S = S_+ \cup S_-$, where S_+ (resp. S_-) consists of those $s \in S$ near which c is convex (resp. concave). Note that for $s \in S_+$ we have $c(s) > \frac{s+1}{3}$. By Proposition 1.2.9, $c(a) = \frac{a+1}{3}$ for $a \in [6\frac{11}{12}, 7]$, and so the biggest point of S is in S_-. This and $c(\tau^4) = \frac{\tau^4+1}{3}$ imply that the set S_+ is non-empty. Let $a_0 = \max S_+$. Then $a_0 \in (\tau^4, 7)$. By Corollary 2.1.4 (i) there exists $(d; \boldsymbol{m}) \in \mathcal{E}$ and $\varepsilon > 0$ such that

$$c(z) = \mu(d; \boldsymbol{m})(z) \quad \text{on } [a_0, a_0+\varepsilon]. \tag{4.1.10}$$

Abbreviate $\mu(z) := \mu(d; \boldsymbol{m})(z)$. By (4.1.10), $\mu(a_0) = c(a_0) > \frac{a_0+1}{3} > \sqrt{a_0}$. Let I be the maximal open interval containing a_0 such that $\mu(z) > \sqrt{z}$ for all $z \in I$. By Lemma 2.1.3, there exists a unique $a' \in I$ with $\ell(\boldsymbol{m}) = \ell(a')$, and $\ell(\boldsymbol{m}) < \ell(z)$ for all other $z \in I$. Further, by Proposition 2.3.2, the constraint $\mu(z)$ is given by two linear functions on I:

$$\mu(z) \;=\; \begin{cases} \alpha + \beta z & \text{if } z < a',\ z \in I, \\ \alpha' + \beta' z & \text{if } z > a',\ z \in I, \end{cases}$$

that is, a' is the only non-smooth point of μ on I. By (4.1.10), and since $a_0 \in S_+$ and $\mu \leq c$, the point $a_0 \in I$ is also a non-smooth point of μ, and hence $a' = a_0$. Now (4.1.10) and the fact that c is nondecreasing show that $\beta' \geq 0$.

Let $k \geq 1$ be such that $a_0 \in [c_{2k+1}, c_{2k-1}]$. Since c_{2k+1} and c_{2k-1} are regular by assumption, we have $a_0 \in (c_{2k+1}, c_{2k-1})$. Note that $u_k(j) \to c_{2k+1}$ and $v_k(j) \to c_{2k-1}$ as $j \to \infty$. Let u_-, u_+ be the two neighboring points from the sequence

$$\cdots < u_k(s+1) < u_k(s) < \cdots < u_k(2) < u_k(1) = v_k(7) < v_k(t) < v_k(t+1) < \cdots$$

from Corollary 4.1.3 with $a_0 \in [u_-, u_+]$. Since u_- and u_+ are regular by assumption, we have $a_0 \in (u_-, u_+)$. Then $\mu(a_0) > \frac{a_0+1}{3} > \frac{u_-+1}{3} = \varphi(u_-) > \psi(u_+) = \sqrt{u_+}$ by Lemma 4.1.7. Since $\beta' \geq 0$, it follows that $\mu(u_+) \geq \mu(a_0) > \sqrt{u_+}$, and hence $u_+ \in I$, and hence $\ell(u_+) > \ell(a_0)$. However, $\ell(z) > \ell(u_-)$ and $\ell(z) > \ell(u_+)$ for all $z \in (u_-, u_+)$; in particular, $\ell(a_0) > \ell(u_+)$, a contradiction. □

As we see in the next two lemmas, one can prove that most of the points $u_k(j)$ and $v_k(j)$ are regular by direct arguments.

Lemma 4.1.8. *The points $u_k(j)$ with $k \geq 1$ and $j \geq 2$ are regular.*

Proof. Abbreviate $u := u_k(j)$ and $\frac{p}{q} := u$ in lowest terms. Assume that $(d;\boldsymbol{m}) \in \mathcal{E}$ is such that $\mu(d;\boldsymbol{m})(u) > \frac{u+1}{3}$ and $\ell(u) = \ell(\boldsymbol{m})$. By Proposition 4.1.1 (i) and Corollary 4.1.4 (i) we can estimate

$$\frac{d}{q\sqrt{u}} \;<\; \frac{3}{q\sqrt{u^2-7u+1}} \;=\; \frac{3}{\sqrt{j^2+7j+1}},$$

an estimate independent of k. Note that $\frac{3}{\sqrt{j^2+7j+1}} < 1$ for $j \geq 2$. Since $\ell(u) = \ell(\boldsymbol{m})$, we have $m_i \geq 1$ for all i. Therefore,

$$E \;\geq\; j\left(1-\frac{d}{q\sqrt{u}}\right)^2 \;>\; j\left(1-\frac{3}{\sqrt{j^2+7j+1}}\right) \;=:\; s(j).$$

The function $s(j)$ is increasing in j, and $s(3) > 1$, proving the lemma for $j \geq 3$ and all $k \geq 1$.

Assume now that $j = 2$. In this case, $s(j) < 1$. We therefore need to use the better estimate $E = 1-\lambda^2 < 1 - \frac{2d^2y(u)}{3\sqrt{u}}$ from Proposition 4.1.1 (ii). With this estimate we have

$$\begin{aligned} 0 \;=\; E+\lambda^2-1 \;\;&>\;\; 2\left(1-\frac{d}{q\sqrt{u}}\right)^2 + \frac{2}{3}\frac{d^2y(u)}{\sqrt{u}} - 1 \\ &=\;\; \left(\frac{2}{q^2u}+\frac{2}{3}\frac{y(u)}{\sqrt{u}}\right)d^2 + \left(-\frac{4}{q\sqrt{u}}\right)d + 1 \;=:\; f(d). \end{aligned}$$

We need to show that $f(d) \geq 0$. Since f is a quadratic polynomial in d, this holds if its discriminant is negative,

$$\frac{16}{q^2u} \;<\; 4\left(\frac{2}{q^2u}+\frac{2}{3}\frac{y(u)}{\sqrt{u}}\right).$$

Multiplying by $\frac{u}{8}$ and using $y(u) = u+1-3\sqrt{u}$, this is equivalent to

$$\frac{1}{q^2}+u \;<\; \frac{u+1}{3}\sqrt{u}.$$

Taking squares, replacing u by $\frac{p}{q}$, and multiplying by $9q^4$, this becomes

$$0 \;<\; -9+p^3q-7p^2q^2+pq\left(-18+q^2\right). \tag{4.1.11}$$

Recall now from Lemma 4.1.2 (ii) that $p = 2F_{k+2}+F_{k+1}$ and $q = 2F_{k+1}+F_k$. Since $pq > 9$ for all $k \geq 1$, (4.1.11) follows from the identities

$$1 \;=\; p^2-7pq+\left(-18+q^2\right),$$

which hold true for $k = 1,2,3$ and hence for all k by Proposition 3.2.3. □

Lemma 4.1.9. *The points $v_k(j)$ with $k \geq 1$ and $j \geq 8$ are regular.*

Proof. Abbreviate $v := v_k(j)$ and $\frac{p}{q} := v$ in lowest terms. Assume that $(d;\boldsymbol{m}) \in \mathcal{E}$ is such that $\mu(d;\boldsymbol{m})(v) > \frac{v+1}{3}$ and $\ell(\boldsymbol{m}) = \ell(v)$. Again, by Proposition 4.1.1 (i) and Corollary 4.1.4 (ii),

$$\frac{d}{q\sqrt{v}} < \frac{3}{q\sqrt{v^2-7v+1}} = \frac{3}{\sqrt{j^2-5j-5}},$$

independent of k. Note that $\frac{3}{\sqrt{j^2-5j-5}} < 1$ for $j \geq 8$. Since $\ell(v) = \ell(\boldsymbol{m})$, we have $m_i \geq 1$ for all i. Therefore,

$$E \geq j\left(1 - \frac{d}{q\sqrt{v}}\right)^2 > j\left(1 - \frac{3}{\sqrt{j^2-5j-5}}\right) =: t(j).$$

The function $t(j)$ is increasing on $\{j \geq 8\}$, and $t(8) > 1$, whence the lemma follows. □

For $k \geq 1$ and $i \geq 0$ we set

$$b_k(i) := v_k(1+3i) = \left[6; \{1,5\}^{\times(k-1)}, 1, 1+3i\right]. \tag{4.1.12}$$

Hence $b_k(2) = v_k(7)$, $b_k(3) = v_k(10)$, ...

Lemma 4.1.10. *$c(a) = \frac{a+1}{3}$ for all $a = b_k(i)$, $k \geq 1$, $i \geq 2$. In particular, the points $v_k(7)$ are regular for all $k \geq 1$.*

The proof is postponed to Corollary 4.2.4 in the next subsection. It uses the existence of special (nearly perfect) elements of $\mathcal{E}$, rather than the estimates of Proposition 4.1.1.

Proof of Theorem 1.1.2 part (ii). By Proposition 4.1.6 it suffices to show that for all $k \geq 1$ the points c_{2k-1}, $u_k(j)$, $j \geq 2$, and $v_k(j)$, $j \geq 7$, are regular. Regularity holds for $u_k(j)$, $j \geq 2$, by Lemma 4.1.8 and for $v_k(j)$, $j \geq 8$, by Lemma 4.1.9. Moreover, when a belongs to the subsequence $b_k(i)$, $i \geq 2$, of the $v_k(j)$, then Lemma 4.1.10 makes the stronger statement that $c(a) = \frac{a+1}{3}$. This holds in particular when $a = v_k(7) = b_k(2)$. Hence these points are regular. Further, because the sequence $\big(b_k(i)\big)_{i\geq 2}$ converges to c_{2k-1} as $i \to \infty$, the continuity of c implies that $c(a) = \frac{a+1}{3}$ also at $a = c_{2k-1}$. Hence these points are also regular, which completes the proof. □

4.2. **The classes $E\big(b_k(i)\big)$.** Recall from Section 3.1 that for $n \geq 0$ the points $b_n = \frac{g_{n+2}}{g_n} < \tau^4$ are the break points of the Fibonacci stairs. The next result shows their relation to the numbers $b_k(i)$, $k \geq 1$, $i \geq 0$, defined by (4.1.12).

Lemma 4.2.1. *$b_{2k} = b_k(0)$ and $b_{2k+1} = b_k(1)$ for all $k \geq 1$.*

Proof. Using Proposition 3.2.3 and Definition 3.2.4 we see that $L_k + F_{k+1} = g_{2k+2}$ and $L_k + 4F_{k+1} = g_{2k+3}$ for all $k \geq 0$. Together with Lemma 4.1.2 (iii) we conclude that

$$\begin{aligned} b_k(0) &= v_k(1) = \frac{L_k + F_{k+1}}{L_{k-1} + F_k} = \frac{g_{2k+2}}{g_{2k}} = b_{2k}, \\ b_k(1) &= v_k(4) = \frac{L_k + 4F_{k+1}}{L_{k-1} + 4F_k} = \frac{g_{2k+3}}{g_{2k+1}} = b_{2k+1}, \end{aligned}$$

as required. □

The lemma says that the sequence $\big(b_n\big)$, $n \geq 2$, extends to a double sequence $\big(b_k(i)\big)$, $k \geq 1$, $i \geq 0$, where for each $k \geq 1$ the sequence $\big(b_k(i)\big)$, $i \geq 0$, emanates from the pair $\big(b_{2k}, b_{2k+1}\big) = \big(b_k(0), b_k(1)\big)$.

Recall from Section 3.1 that $E(b_n) := \big(g_{n+1}; W(b_n)\big)$, where $W(b_n) := g_n\, \boldsymbol{w}(b_n)$. In order to prove Lemma 4.1.10, we associate classes $E\big(b_k(i)\big)$ to all $b_k(i)$ as follows. Let $b_k(i) =: \frac{p}{q}$. Let $\boldsymbol{m}_k(i)$ be the tuple obtained from $q\, \boldsymbol{w}\big(b_k(i)\big)$ by replacing its last block $\big(1^{\times(1+3i)}\big)$ by $\big(i, 1^{\times(1+2i)}\big)$, and set $d_k(i) := q\big(1 + b_k(i)\big)/3$. Then define $E\big(b_k(i)\big) := \big(d_k(i); \boldsymbol{m}_k(i)\big)$.

Note that for $i = 0, 1$ we have (with $n = 2k + i$) that $\boldsymbol{m}_k(i) = g_n\, \boldsymbol{w}\big(b_k(i)\big) = W(b_n)$ and, by (3.1.2), $d_k(i) = g_n\big(1 + \frac{g_{n+2}}{g_n}\big)/3 = (g_n + g_{n+2}) = g_{n+1}$. Therefore,

$$E\big(b_k(0)\big) = E\big(b_{2k}\big), \qquad E\big(b_k(1)\big) = E\big(b_{2k+1}\big) \quad \text{for all } k \geq 1. \tag{4.2.1}$$

Proposition 4.2.2. *$E\big(b_k(i)\big) \in \mathcal{E}$ for all $k \geq 1$ and $i \geq 0$.*

Before proving Proposition 4.2.2, we show that it is the key to completing the calculation of c on the interval $[1, 7]$.

Corollary 4.2.3. *Part (i) of Theorem 3.1.1 holds.*

Proof. This is immediate from equation (4.2.1). □

Corollary 4.2.4. *Lemma 4.1.10 holds.*

Proof. We argue as in the proof of Corollary 3.1.2. Let $(d; \boldsymbol{m}) \in \mathcal{E}$. Write $b' := b_k(i)$, $d' := d_k(i)$, $\boldsymbol{m}' := \boldsymbol{m}_k(i)$, so that $(d'; \boldsymbol{m}') = E(b_k(i))$. If $(d; \boldsymbol{m}) = \big(d'; \boldsymbol{m}'\big)$, then

$$\mu(d; \boldsymbol{m})\big(b'\big) = \frac{\boldsymbol{m}' \cdot \boldsymbol{w}(b')}{d'} = \frac{3b'}{b' + 1} < \sqrt{b'}$$

because $b' > \tau^4$. If $(d; \boldsymbol{m}) \neq (d'; \boldsymbol{m}')$, then $d\, d' \geq \boldsymbol{m} \cdot \boldsymbol{m}'$ by positivity of intersections. By the definition of d' and $\boldsymbol{m}'$, and since $\boldsymbol{m}$ is ordered, this spells out to

$$q\, d\, \tfrac{1+b'}{3} \geq \boldsymbol{m} \cdot \boldsymbol{m}' \geq q\, \boldsymbol{m} \cdot \boldsymbol{w}(b'),$$

i.e., $\frac{1+b'}{3} \geq \frac{1}{d} \boldsymbol{m} \cdot \boldsymbol{w}(b') =: \mu(d; \boldsymbol{m})(b')$. □

Remark 4.2.5. (i) Note that the classes $E\big(b_k(i)\big) = \big(d_k(i); \boldsymbol{m}_k(i)\big)$ are perfect for $i = 0, 1$, but are not perfect for $i \geq 2$. Since, nevertheless, at $b_k(i) =: \frac{p}{q}$ we have

$$\boldsymbol{m}_k(i) \cdot \boldsymbol{w}\big(b_k(i)\big) \;=\; q\, \boldsymbol{w}\big(b_k(i)\big) \cdot \boldsymbol{w}\big(b_k(i)\big) \;=\; q\, b_k(i) \quad \text{for all } i, \tag{4.2.2}$$

these classes are useful also for $i \geq 2$.

(ii) For $i \geq 3$ there are other choices for $\boldsymbol{m}_k(i)$ that can be used to prove Lemma 4.1.10. All one needs is that (4.2.2) holds. For this, one just needs to alter the last block $1^{\times j}$ so that the sum of entries stays intact, while the sum of the squares goes up by $i^2 - i$.

If $i = 2$, one has no choice: In order to make the sum of squares go up by 2 one must replace $1^{\times 7}$ by $2, 1^{\times 5}$. If $i = 3$, however, instead of replacing $1^{\times 10}$ by $3, 1^{\times 7}$, one can replace it by $2^3, 1^{\times 4}$. The resulting class $\big(d_k(3); \boldsymbol{m}'_k(3)\big)$ lies in $\mathcal{E}$ for all k. If $i = 4$, instead of replacing $1^{\times 13}$ by $4, 1^{\times 9}$, one can make the sum of squares go up by 12 also by replacing $\boldsymbol{m}_k(4) = \big(\dots, 1^{\times 13}\big)$ by one of

$$\boldsymbol{m}_k^{(1)}(4) := \big(\dots, 3^{\times 2}, 1^{\times 7}\big), \quad \boldsymbol{m}_k^{(2)}(4) := \big(\dots, 3, 2^{\times 3}, 1^{\times 4}\big), \quad \boldsymbol{m}_k^{(3)}(4) := \big(\dots, 2^{\times 6}, 1\big).$$

However, while the classes $\big(d_k(4); \boldsymbol{m}_k^{(2)}(4)\big)$ and $\big(d_k(4); \boldsymbol{m}_k^{(3)}(4)\big)$ reduce to $(1;1,1)$, and thus lie in $\mathcal{E}$, the classes $\big(d_k(4); \boldsymbol{m}_k^{(1)}(4)\big)$ do not reduce correctly. $\diamond$

We now turn to the proof of Proposition 4.2.2. It follows from Lemma 4.2.6 and Proposition 4.2.7 below.

Lemma 4.2.6. *The class $\big(d_k(i); \boldsymbol{m}_k(i)\big)$ satisfies the Diophantine conditions* (1.2.4) *for the elements of $\mathcal{E}$.*

Proof. Write $b := b_k(i)$ and $(d; \boldsymbol{m}) = \big(d_k(i); \boldsymbol{m}_k(i)\big)$. By Lemma 1.2.6, $\sum m_s = \sum q\, w_s(b) = q(b+1) - 1 = 3d - 1$. Further, $\sum m_s^2 = q^2 b + i^2 - i$, and by Corollary 4.1.4 (ii),

$$q^2\left(b^2 - 7b + 1\right) \;=\; (1+3i)^2 - 5\,(1+3i) - 5 \;=\; 9i^2 - 9i - 9.$$

Therefore,

$$\begin{aligned} d^2 + 1 \;=\; \tfrac{1}{9} q^2\,(1+b)^2 + 1 \;&=\; \tfrac{1}{9} q^2\left(b^2 - 7b + 1\right) + q^2 b + 1 \\ &=\; i^2 - i + q^2 b \;=\; \sum m_s^2, \end{aligned}$$

as required. $\square$

Proposition 4.2.7. *The classes $E\big(b_k(i)\big)$ reduce to $(1;1,1)$ for all $k \geq 1$ and $i \geq 0$.*

Proof. We fix $i \geq 0$, and argue by induction on $k \geq 1$.

Step 1. Assume that $k = 1$. Set $j = 1 + 3i$. The weight expansion of $b_1(i) = [\,6; 1, j\,] =: \frac{p}{q}$ is

$$\tfrac{1}{q}\big((j+1)^{\times 6}, j, 1^{\times j}\big). \tag{4.2.3}$$

Together with (1.2.3) we find $d_1(i) = \frac{q}{3}\big(1+b_1(i)\big) = \frac{1}{3}\big(6(j+1)+j+j+1\big) = \frac{1}{3}\,(8j+7) = 8i+5$, and so

$$\big(d_1(i); \boldsymbol{m}_1(i)\big) \;=\; \big(8i+5; (3i+2)^{\times 6}, 3i+1, i, 1^{\times(2i+1)}\big).$$

Applying five standard Cremona moves yields, successively,

$$\begin{aligned} \big(7i+4; (3i+2)^{\times 3}, 3i+1, (2i+1)^{\times 3}, i, 1^{\times(2i+1)}\big);\\ \big(5i+2; 3i+1, (2i+1)^{\times 3}, i^{\times 4}, 1^{\times(2i+1)}\big);\\ \big(3i+1; 2i+1, i^{\times 5}, 1^{\times(2i+1)}\big);\\ \big(2i+1; i+1, i^{\times 3}, 1^{\times(2i+1)}\big);\\ \big(i+1; i, 1^{\times(2i+2)}\big). \end{aligned}$$

The standard Cremona move maps $(s+1; s, 1^{\times t})$ to $(s; s-1, 1^{\times(t-2)})$ for any $s \geq 1$ and $t \geq 2$. Applying i more standard Cremona moves therefore moves $\big(i+1; i, 1^{\times(2i+2)}\big)$ to $(1; 1, 1)$.

Step 2. Assume by induction that $\big(d_k(i); \boldsymbol{m}_k(i)\big)$ reduces to $(1;1,1)$. We shall show that $\big(d_{k+1}(i); \boldsymbol{m}_{k+1}(i)\big)$ reduces to $\big(d_k(i); \boldsymbol{m}_k(i)\big)$ by five standard Cremona moves.

The end of the weight expansion $q\boldsymbol{w}\big(b_k(i)\big) = q\boldsymbol{w}\big(v_k(j)\big)$ is

$$\big(\ldots,\ (48j+41)^{\times 5},\ 41j+35,\ (7j+6)^{\times 5},\ 6j+5,\ (j+1)^{\times 5},\ j,\ 1^{\times j}\big).$$

Using $F_{k+1} = L_k + F_k$ and $L_{k+1} = 5F_{k+1} + L_k$ from (4.1.2), and

$$L_{k+1} - L_k \;=\; 5L_k + (L_k - L_{k-1}),$$

which follows from these two formulae, we see that in general,

$$\begin{aligned} \text{(4.2.4)} \quad \boldsymbol{m}_k(i) \;=\; \big(&(jF_k + L_{k-1})^{\times 6},\ (j+1)L_{k-1} - L_{k-2},\\ &(jF_{k-1} + L_{k-2})^{\times 5},\ (j+1)L_{k-2} - L_{k-3},\ \ldots,\\ &(jF_2 + L_1)^{\times 5},\ (j+1)L_1 - L_0,\ (j+1)^{\times 5},\ j,\ i,\ 1^{\times(j-i)}\big). \end{aligned}$$

It will be convenient to express the numbers $d_k(i)$ in terms of the *even* Fibonacci numbers $h_k := f_{2k}$. Thus

$$\text{(4.2.5)} \qquad h_1 = 1,\ h_2 = 3,\ h_3 = 8,\ h_4 = 21,\ h_5 = 55,\ h_6 = 144,\ \ldots$$

Lemma 4.2.8. $d_k(i) = h_{2k+2} + (i-2)h_{2k+1}$

Proof. We fix i and $j = 1+3i$, and write $d_k = d_k(i)$, $\boldsymbol{m}_k = \boldsymbol{m}_k(i)$, $b_k = b_k(i)$, etc. Set $b_k = \frac{p_k}{q_k}$. By Lemma 4.1.2 (iii),

$$p_k = F_{k+2} + (j-6)F_{k+1}, \qquad q_k = F_{k+1} + (j-6)F_k.$$

Therefore, $3d_k = q_k(b_k+1) = p_k + q_k = F_{k+2} + (j-5)F_{k+1} + (j-6)F_k$. We thus need to show that

$$F_{k+2} + (j-5)F_{k+1} + (j-6)F_k \;=\; 3h_{2k+2} + 3(i-2)h_{2k+1}.$$

This holds true for $k = 1$ and $k = 2$, and so it holds true for all $k \geq 1$ by Proposition 3.2.3. □

Lemma 4.2.9. *The class* $\big(d_{k+1}(i); \boldsymbol{m}_{k+1}(i)\big)$ *reduces to* $\big(d_k(i); \boldsymbol{m}_k(i)\big)$ *by five Cremona transforms.*

Proof. Fix i. By (4.2.4) and Lemma 4.2.8, the entries of $\big(d_{k+1}(i); \boldsymbol{m}_{k+1}(i)\big)$ are given by linear formulas in Fibonacci numbers, that depend only on k. Using (4.2.4) and Lemma 4.2.8 one checks for $k = 1$ and $k = 2$ that $\big(d_{k+1}(i); \boldsymbol{m}_{k+1}(i)\big)$ reduces to $\big(d_k(i); \boldsymbol{m}_k(i)\big)$ by five Cremona transforms with equal reordering at each stage. The lemma thus follows from Proposition 3.2.3. □

The proof of Proposition 4.2.7 is complete.

4.3. **The ghost stairs.** In this section we compute the contribution of the classes $E\big(b_k(i)\big) = \big(d_k(i); \boldsymbol{m}_k(i)\big)$, $i \geq 2$, to the graph of $c(a) = \frac{a+1}{3}$ on $\big[\tau^4, 7\big]$. The lemma and the proposition below are not needed for the results of this paper, but they illuminate the role of these classes.

Lemma 4.3.1. *Assume that* $i \geq 3$. *Then* $\mu\big(d_k(i); \boldsymbol{m}_k(i)\big)(a) \leq \sqrt{a}$ *for all* $a > 1$ *and all* $k \geq 1$.

Proof. Write $d = d_k(i)$, $\boldsymbol{m} = \boldsymbol{m}_k(i)$. Assume that $\mu(d; \boldsymbol{m})(a) > \sqrt{a}$ for some $a \geq 1$. Let I be the open interval such that $\mu(d; \boldsymbol{m})(z) > \sqrt{z}$ and $a \in I$. Let a_0 be the unique point in I with $\ell(a_0) = \ell(\boldsymbol{m})$. Recall that

$$\boldsymbol{m} \,=\, \big(\dots,\, i,\, 1^{\times(2i+1)}\big).$$

Since $i \geq 3$, the last block of $\boldsymbol{w}(a_0)$ must have length $2i+1$ according to Lemma 2.1.7 (i). But

$$|i - (2i+1)| \,=\, i + 1 \,\geq\, \sqrt{2i+1}\,,$$

in contradiction to Lemma 2.1.8 (i). □

Thus, surprisingly, for $i \geq 3$ the classes $E\big(b_k(i)\big)$ give no embedding constraints, but nevertheless are most useful to find $c(a)$ on $[\tau^4, 7]$.

We now look at the case $i = 2$. For $k \geq 1$ write $e_k := b_k(2) = v_k(7)$ and $E(e_k) = (d_k; \boldsymbol{m}_k) := \big(d_k(2); \boldsymbol{m}_k(2)\big)$. Recall from Lemma 4.2.8 that $d_k = h_{2k+2}$, where $h_{2k+2} = f_{4k+4}$ is an even Fibonacci number. We now show that the corresponding constraint functions $\mu(d_k; \boldsymbol{m}_k)$ form a staircase whose properties echo that of the Fibonacci stairs on the other side of τ^4. However this staircase does not add anything new to the graph of $c(a)$ because it never rises above the line $y = \frac{a+1}{3}$. Thus we call it the *ghost stairs.* Note also that, although $E(e_k)$ is made from $\boldsymbol{w}(e_k)$ and so influences $c(a)$ at $a = e_k$, it gives a constraint that is centered at the convergent $c_{2k+1} < e_k$.

Proposition 4.3.2 (The ghost stairs)**.**

$$\mu(d_k;\boldsymbol{m}_k)(z) \;=\; \begin{cases} \frac{z+1}{3} & \textit{for } z\in[c_{2k},c_{2k+1}]\,, \\ \frac{h_{2k+3}}{h_{2k+2}} & \textit{for } z\in[c_{2k+1},e_k]\,. \end{cases}$$

Since for each k we have $c_{2k}<\tau^4$, the proposition shows that $\mu(d_k;\boldsymbol{m}_k)(z)=c(z)=\frac{z+1}{3}$ on $[\tau^4,c_{2k+1}]$.

Proof. Fix k, and recall from Lemma 4.1.2 and Corollary 4.1.3 that

$$c_{2k} \;=\; \big[6;\{1,5\}^{(k-1)},1,5\big], \tag{4.3.1}$$

$$c_{2k+1} \;=\; \big[6;\{1,5\}^{(k-1)},1,6\big], \tag{4.3.2}$$

$$e_k \;=\; \big[6;\{1,5\}^{(k-1)},1,7\big] \;=\; \tfrac{L_k+7F_{k+1}}{L_{k-1}+7F_k}, \tag{4.3.3}$$

and that $c_{2k}<c_{2k+1}<e_k$. If $z\in(c_{2k},c_{2k+1})$ then $z=\big[6;\{1,5\}^{(k-1)},1,g,h,\dots\big]$ with $g=5$ and $h\ge 1$, while if $z\in(c_{2k+1},e_k)$ then $z=\big[6;\{1,5\}^{(k-1)},1,g,h,\dots\big]$ with $g=6$ and $h\ge1$. In both cases, the weight expansion has the form

$$\begin{aligned} \boldsymbol{w}(z) \;=\; &\big(1^{\times 6}, z-6, (7-z)^{\times 5}, 6z-41,\dots,x_{2k-1}(z)=zL_{k-1}-L_k, \\ &\big(x_{2k}(z)=F_{k+1}-zF_k\big)^{\times g},\big(x_{2k+1}(z)\big)^{\times h},\dots\big). \end{aligned} \tag{4.3.4}$$

Further $\boldsymbol{w}(e_k)$ begins the same way, but ends at the block of terms

$$\big(x_{2k}(z)\;=\;F_{k+1}-zF_k\big)^{\times 7}\;=\;\big(\alpha^e_{2k}+z\beta^e_{2k}\big)^{\times 7},$$

where the last expression uses the elements α^e_j and β^e_j of equation (2.2.2) with $a=e:=e_k$.

Since e_k has $N+1$ blocks where N is even, Corollary 2.2.7 implies that

$$\sum_j \ell_j\,x_j(e_k)\big(\alpha^e_j+z\beta^e_j\big)=e_k=\tfrac{L_k+7F_{k+1}}{L_{k-1}+7F_k}.$$

Further, $\boldsymbol{m}_k=q\,\boldsymbol{w}(e_k)$ where $q:=L_{k-1}+7F_k$, except for the last block where we have $2,1^{\times5}$ instead of $1^{\times7}$. When z has $g=6$ it follows that

$$\boldsymbol{m}_k\cdot\boldsymbol{w}(z)=\boldsymbol{m}_k\cdot\boldsymbol{w}(e_k)=q\,\boldsymbol{w}(e_k)\cdot\boldsymbol{w}(e_k)=q\,e_k=L_k+7F_{k+1}.$$

Thus $\mu(d_k;\boldsymbol{m}_k)(z)$ is constant on this interval. On the other hand, if $g=5$ then $x_{2k+1}(z)=zL_k-L_{k+1}$ and we find

$$\begin{aligned} \boldsymbol{m}_k\cdot\boldsymbol{w}(z) &= \sum_j q\,\ell_j\,x_j(e_k)\big(\alpha^e_j+z\beta^e_j\big)-x_{2k}(z)+x_{2k+1}(z) \\ &= q\,e_k-F_{k+1}+zF_k+zL_k-L_{k+1} \\ &= 6F_{k+1}+L_k-L_{k+1}+z(L_k+F_k) \\ &= (1+z)F_{k+1}. \end{aligned}$$

But

$$3d_k=q(1+e_k)=L_{k-1}+7F_k+L_k+7F_{k+1}=9F_{k+1}.$$

Therefore $\mu(d_k; \boldsymbol{m}_k)(z) = (z+1)/3$ for $z \in [c_{2k}, c_{2k+1}]$ and it remains to check that $c_{2k+1} + 1 = 3h_{2k+3}/h_{2k+2}$. Since $c_{2k+1} = F_{k+2}/F_{k+1}$ and $F_{k+1} = 3h_{2k+2}$ this reduces to the identity

$$f_{4k+8} + f_{4k+4} = 3f_{4k+6},$$

which is readily checked using Proposition 3.2.3. □

Remark 4.3.3. (i) At the points $v_k(j)$ with $j \geq 7$, Theorem 1.1.2 (ii) implies that $c\big(v_k(j)\big) = \frac{v_k(j)+1}{3}$, which by Lemma 4.1.2 (iii) can be written as

$$c\left(\frac{\ell_{4k+2} + jf_{4k+4}}{\ell_{4k-2} + jf_{4k}}\right) = \frac{\ell_{4k} + jf_{4k+2}}{\ell_{4k-2} + jf_{4k}},$$

where $\ell_{4k+2} = 3L_k$ as in Definition 3.2.4. In particular, at $b_k(2) = v_k(7)$ and $b_k(3) = v_k(10)$,

$$c\left(\frac{h_{2k+3}}{h_{2k+1}}\right) = \frac{h_{2k+2}}{h_{2k+1}} \qquad \text{and} \qquad c\left(\frac{\ell_{4k+5}}{\ell_{4k+1}}\right) = \frac{\ell_{4k+3}}{\ell_{4k+1}},$$

where the h_k are the even Fibonacci numbers of (4.2.5). On the other hand, on the left of τ^4, where $\frac{a+1}{3} < \sqrt{a}$, we have by Theorem 1.1.2 that

$$c(b_n) = c\left(\frac{g_{n+2}}{g_n}\right) = \frac{g_{n+2}}{g_{n+1}} = \frac{b_{n+1}+1}{3}.$$

In other words, the function c attains the value $\frac{b_n+1}{3}$ already at b_{n-1}.

(ii) Recall from Section 3.1 that $a_n = \big(\frac{g_{n+1}}{g_n}\big)^2$, and that on the left of τ^4, the classes $W'(a_n)$ obtained from $W(a_n) = g_n^2\, \boldsymbol{w}(a_n)$ by adding one 1 were very useful to establish the Fibonacci stairs. One may try to define similar classes at $a_n' := \big(\frac{h_{n+1}}{h_n}\big)^2$. Denote by $W''(a_n')$ the sequence obtained from $h_n^2\, \boldsymbol{w}(a_n')$ by removing three of the 1s at the end, and adding one 2. Thus when $n = 3$ we get

$$a_3' = \big(\tfrac{21}{8}\big)^2, \qquad W''(a_3') = \big(64^{\times 6}, 57, 7^{\times 8}, 2, 1^{\times 4}\big).$$

It is easy to check that the tuple $\big(h_n h_{n+1}; W''(a_n')\big)$ satisfies the Diophantine equations (1.2.4). However, when $n \geq 3$ this is *not* an element of $\mathcal{E}$ because it has negative intersection with the class $(3; 2, 1^{\times 6}) \in \mathcal{E}$. On the other hand, when $n = 2$ this gives $(24; 9^{\times 7}, 2, 1^{\times 6}) \in \mathcal{E}$ which as we will see in Theorem 5.2.3 does give an obstruction near $a = 7\frac{1}{7}$, and when $n = 1$ we get $(3; 2, 1^{\times 6})$ itself. ◇

5. The interval [7, 9]

This section calculates c on the interval $[7, 9]$. The main arguments are contained in §5.2 and §5.3. We begin in §5.1 by establishing some estimates that are most useful on $[8, 9]$ but are also needed for some of the arguments concerning $[7, 8]$ such as Lemma 5.2.7.

5.1. **Preliminaries.** We begin with a simple result about continued fractions. Let $q_n(a)$ be the denominator of the nth convergent $[\ell_0; \ell_1, \dots, \ell_n]$ to the continued fraction

$$a := [\ell_0; \ell_1, \dots, \ell_N] = \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \dots}}.$$

Thus $q_1(a) = \ell_1, q_2(a) = 1 + \ell_1\ell_2$ and, in general, $q_n(a) = \ell_n q_{n-1}(a) + q_{n-2}(a)$. Then an easy induction argument shows that:

Sublemma 5.1.1. *Let $L := \sum_{j=1}^N \ell_j$. Then $q_N(a) \geq L$.*

In the sequel, we abbreviate $\sigma := \sum_{i>\ell_0} \varepsilon_i^2 < 1$ and $\sigma' := \sum_{\ell_0 < i \leq M - \ell_N} \varepsilon_i^2 \leq \sigma'$.

Lemma 5.1.2. *Assume that $(d; \boldsymbol{m}) \in \mathcal{E}$ is such that $\mu(d; \boldsymbol{m})(a) > \sqrt{a}$ for some $a \in (\tau^4, 9)$ with $\ell(a) = \ell(\boldsymbol{m})$. Assume further that $y(a) > \frac{1}{q}$ where $q := q_N(a)$, and denote $v_M := \frac{d}{q\sqrt{a}}$. Then*

(i) $|\sum_{i\geq 1} \varepsilon_i| \leq \sqrt{\sigma L}$

(ii) *If $v_M < 1$ then $|\sum_{i\geq 1} \varepsilon_i| \leq \sqrt{\sigma' L}$.*

(iii) *If $v_M \leq \frac{1}{2}$, then $v_M > \frac{1}{3}$ and $\sigma' \leq \frac{1}{2}$. If $v_M \leq \frac{3}{4}$, then $\sigma' \leq \frac{7}{8}$.*

(iv) *Define $\delta := y(a) - \frac{1}{q} > 0$. Then*

$$d \leq \tfrac{\sqrt{a}}{\delta}\left(\sqrt{\sigma L} - 1\right) \leq \tfrac{\sqrt{a}}{\delta}\left(\sqrt{\sigma q} - 1\right) < \tfrac{\sqrt{a}}{\delta}\left(\tfrac{\sigma}{\delta v_M} - 1\right).$$

Further, if $v_M < 1$, then σ can be replaced by σ'. In particular, always

$$d < \tfrac{\sqrt{a}}{\delta}\left(\tfrac{2}{\delta} - 1\right) < \tfrac{2\sqrt{a}}{\delta^2}.$$

Proof. **Step 1:** $\sum_{i\geq 1} \varepsilon_i < 0$.

Proposition 2.1.1 (iv) states that

$$-\sum \varepsilon_i = 1 + \tfrac{d}{\sqrt{a}}\big(y(a) - \tfrac{1}{q}\big). \tag{5.1.1}$$

Since we assume that $y(a) > \frac{1}{q}$, Step 1 is immediate.

Step 2: $\sum_{i>\ell_0} |\varepsilon_i| \geq |\sum_{i\geq 1} \varepsilon_i|$.

If $\sum_{i\leq \ell_0} \varepsilon_i \geq 0$, then by Step 1 we have

$$|\sum_{i\geq 1} \varepsilon_i| \leq |\sum_{i>\ell_0} \varepsilon_i| \leq \sum_{i>\ell_0} |\varepsilon_i|,$$

as required. Therefore, suppose that $\sum_{i\leq \ell_0} \varepsilon_i < 0$. Let $P = \{i > \ell_0 \mid \varepsilon_i > 0\}$ and $Q = \{i > \ell_0 \mid \varepsilon_i \leq 0\}$. Because $w_i = 1$ for $i \leq \ell_0$, we have

$$\begin{aligned} 0 < \varepsilon \cdot \boldsymbol{w} &= \sum_{i\leq \ell_0} \varepsilon_i + \sum_{i\in P} \varepsilon_i w_i - \sum_{i\in Q} |\varepsilon_i| w_i \\ &< \sum_{i\leq \ell_0} \varepsilon_i + \sum_{i\in P} \varepsilon_i. \end{aligned}$$

Therefore

$$0 > \sum_{i\geq 1}\varepsilon_i \geq \sum_{i\in Q}\varepsilon_i = -\sum_{i\in Q}|\varepsilon_i| \geq -\sum_{i>\ell_0}|\varepsilon_i|.$$

Step 3: *Proof of (i).* Let $a = [\ell_0; \ell_1, \dots, \ell_N]$ as above, and write ε as $N+1$ blocks each of length ℓ_j. Assume first that ε_i is constant on each block with absolute value δ_j. Let $\nu_j = \ell_j\delta_j^2$ so that $|\delta_j| = \sqrt{\frac{\nu_j}{\ell_j}}$. Then

$$\sum_{i>\ell_0}\varepsilon_i^2 = \sum_{j\geq 1}\ell_j\delta_j^2 = \sum \nu_j = \sigma.$$

Hence, by Step 2,

$$\begin{aligned} \Big|\sum_{i\geq 1}\varepsilon_i\Big| &\leq \sum_{i>\ell_0}|\varepsilon_i| = \sum \ell_j\,|\delta_j| \\ &= \sum \ell_j\sqrt{\frac{\nu_j}{\ell_j}} \\ &= \sum \sqrt{\nu_j\ell_j} \\ &\leq \sqrt{\sum \ell_j}\sqrt{\sum \nu_j} \leq \sqrt{\sigma L}. \end{aligned}$$

This proves (i) in the case when the ε_i are constant on the jth block for all $j \geq 1$. But by Lemma 2.1.7 the only other possibility is that there is precisely one block, say the Jth, on which ε_i is not constant. In that case we subdivide this block into two subblocks of lengths $\ell_J - 1$ and 1. Since the upper bound $\sqrt{\sigma L}$ depends only on the sum of the ℓ_j, the argument goes through as before.

Step 4: *Proof of (ii).* We abbreviate $M' = M - \ell_N$, and write $v_i := \frac{d}{\sqrt{a}}w_i$. If the v_i are constant on the last block and if $v_M := \frac{d}{q\sqrt{a}} < 1$, then $m_{M'+1} = \cdots = m_M = 1$, and so $\varepsilon_{M'+1} = \cdots = \varepsilon_M = 1 - v_M > 0$. Since also $\sum_i \varepsilon_i < 0$, we have $|\sum_i \varepsilon_i| \leq \Big|\sum_{i=1}^{M'}\varepsilon_i\Big|$. Hence, the argument in Step 3 adapts to show that

$$\Big|\sum_{i\geq 1}\varepsilon_i\Big| \leq \left|\sum_{i=1}^{M'}\varepsilon_i\right| \leq \sum_{i>\ell_0}^{M'}|\varepsilon_i| \leq \sqrt{\sigma' L}.$$

Proof of (iii). Assume that $v_M \leq \frac{1}{3}$. If $\ell_N \geq 3$, then

$$1 > \varepsilon_{M'+1}^2 + \cdots + \varepsilon_M^2 \geq 3\left(\tfrac{2}{3}\right)^2 > 1,$$

a contradiction. If $\ell_N = 2$, then $v_{M-2} = v_{M-1} + v_M = 2v_M \leq \frac{2}{3}$, and so

$$1 > \varepsilon_{M-2}^2 + 2\varepsilon_M^2 \geq \left(\tfrac{1}{3}\right)^2 + 2\left(\tfrac{2}{3}\right)^2 = 1,$$

a contradiction. Further, $\varepsilon_M \geq \frac{1}{2}$ implies that

$$\sigma' \leq \sum_{\ell_0 < i \leq M-2}\varepsilon_i^2 \leq \sigma - \tfrac{1}{2} \leq \tfrac{1}{2}.$$

The second claim in (iii) is proved similarly.

Proof of (iv). We use Sublemma 5.1.1 and equation (5.1.1) to estimate

$$\sqrt{\sigma q} \ \geq\ \sqrt{\sigma L} \ \geq\ 1 + \tfrac{d}{\sqrt{a}}\left(y(a) - \tfrac{1}{q}\right) \ =\ 1 + \tfrac{d}{\sqrt{a}}\delta \ =\ 1 + \delta q v_M \ >\ \delta q v_M. \tag{5.1.2}$$

Therefore, $\sqrt{q} < \frac{\sqrt{\sigma}}{\delta v_M}$, and so, using again (5.1.2),

$$d \ \leq\ \tfrac{\sqrt{a}}{\delta}\left(\sqrt{\sigma L} - 1\right) \ \leq\ \tfrac{\sqrt{a}}{\delta}\left(\sqrt{\sigma q} - 1\right) \ <\ \tfrac{\sqrt{a}}{\delta}\left(\tfrac{\sigma}{\delta v_M} - 1\right).$$

If $v_M < 1$, we repeat this argument with σ replaced by σ'. This completes the proof. □

5.2. **The interval** $[7,8]$**.** In this section we calculate $c(a)$ on the interval $[7,8]$. At some places, we will use the computer. We will therefore first prove a weaker result that does not use the computer.

Proposition 5.2.1. *There are only finitely many* $(d;\boldsymbol{m}) \in \mathcal{E}$ *for which there is* $a \geq 7$ *with* $c(a) = \mu(d;\boldsymbol{m})(a) > \sqrt{a}$.

Proof. Suppose that $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ for some $a \geq 7$. Let I be the maximal open interval containing a on which $\mu(d;\boldsymbol{m})(z) > \sqrt{z}$, and let $a_0 \in I$ be the unique element with $\ell(a_0) = \ell(\boldsymbol{m})$. (This exists by Lemma 2.1.3.) If $7 \in I$, then clearly $a_0 = 7$ so that $(d;\boldsymbol{m})$ belongs to the finite set $\mathcal{E}_7$. Otherwise, $a_0 > 7$. In particular, $y(a_0) > y(7) = 8 - 3\sqrt{7} > \frac{1}{20}$. Moreover $a_0 < 9$ by Corollary 1.2.4.

Now write $a_0 = p/q$. There are only finitely many $a = \frac{p}{q} \in [7,9]$ with $q \leq 40$, and for each of them Corollary 2.1.4 shows that there are only finitely many obstructive $(d;\boldsymbol{m})$. We can therefore assume that $q := q(a_0) \geq 40$ so that $y(a_0) - \frac{1}{q} \geq \frac{1}{40} > 0$. Since $\ell(a_0) = \ell(\boldsymbol{m})$ we can apply the last statement of Lemma 5.1.2 to conclude that

$$d \ \leq\ 2(40)^2\sqrt{a_0} \ <\ 6(40)^2.$$

Since for each D there are only finitely many $(d;\boldsymbol{m}) \in \mathcal{E}$ with $d \leq D$, this completes the proof. □

Remark 5.2.2. The result in Proposition 5.2.1 clearly extends to any interval of the form $[a,b]$ provided that $a > \tau^4$. ◇

We already know that $c(a) = \frac{8}{3}$ on $[7, 7\frac{1}{9}]$ by Proposition 1.2.9. We can therefore assume that $a \in [7\frac{1}{9}, 8]$.

In order to explain our notation in Theorem 5.2.3 below, we work out the constraint given by the class

$$(d;\boldsymbol{m}) \ =\ \left(48; 18^{\times 7}, 3, 2^{\times 7}\right) \ \in\ \mathcal{E}.$$

Note that $\ell(\boldsymbol{m}) = 7 + 8 = \ell(7\frac{1}{8})$. It gives the constraint $\mu(d;\boldsymbol{m})(7\frac{1}{8}) = \frac{1025}{384} > \sqrt{7\frac{1}{8}}$ at $7\frac{1}{8}$. For $a = 7 + x$ with $x \in [\frac{1}{9}, \frac{1}{8}]$ we have $\boldsymbol{w}(a) = \left(1^{\times 7}, x^{\times 8}, \dots\right)$. Therefore,

$$\boldsymbol{m} \cdot \boldsymbol{w}(a) = 7 \cdot 18 + 3x + 14x = 126 + 17x = 7 + 17a,$$

and so $\mu(d;\boldsymbol{m})(a) = \frac{1}{48}(7+17a)$. Note that $\frac{1}{48}(7+17a) = \sqrt{a}$ at $u_{\frac{1}{8}} := 7.12499$ (where the last decimal is rounded). Similarly, for $a = 7+x$ with $x \in [\frac{1}{8}, \frac{1}{7}]$ we have $\boldsymbol{w}(a) = \left(1^{\times 7}, x^{\times 7}, 1-7x, \dots\right)$. Therefore,

$$\boldsymbol{m} \cdot \boldsymbol{w}(a) = 7 \cdot 18 + 3x + 12x + 2 - 14x = 128 + x = 121 + a,$$

and so $\mu(d;\boldsymbol{m})(a) = \frac{1}{48}(121+1a)$. Note that $\frac{1}{48}(121+a) = \sqrt{a}$ at $v_{\frac{1}{8}} := 7.12501$ (where the last decimal is rounded). The interval containing $a = 7\frac{1}{8}$ on which this class gives a constraint is therefore $I_{\frac{1}{8}} := [u_{\frac{1}{8}}, v_{\frac{1}{8}}]$, and

$$\mu(d;\boldsymbol{m})(z) = \begin{cases} \frac{1}{48}(7+17z) & \text{if } z \in \left[u_{\frac{1}{8}}, 7\frac{1}{8}\right] \\ \frac{1}{48}(121+z) & \text{if } z \in \left[7\frac{1}{8}, v_{\frac{1}{8}}\right]. \end{cases}$$

All this is expressed in the first row of the table below. In the same way we compute (A,B), (A',B'), u_x, v_x and $\mu(a) := \mu(d;\boldsymbol{m})(a)$ at $a = 7+x$ for the other seven classes in the table below, where we write $\mu(z) = \frac{1}{d}(A+Bz)$ for z just less than a and $\mu(z) = \frac{1}{d}(A'+B'z)$ for z just greater than a. Note that the eight intervals $[u_x, v_x]$ are all disjoint.

Theorem 5.2.3. *For $a \in [7\frac{1}{9}, 8]$ we have $c(a) = \sqrt{a}$ except for the eight intervals $[u_x, v_x]$ where $c(a)$ is as described in the following table.*

(5.2.1)

a	$(d;\boldsymbol{m})$	(A,B)	(A',B')	u_x	v_x	$\mu(a)$	$\mu(a)-\sqrt{a}$
$7\frac{1}{8}$	$(48; 18^{\times 7}, 3, 2^{\times 7})$	$(7,17)$	$(121,1)$	7.12499	7.12501	$\frac{1025}{384}$	$1.27\,10^{-6}$
$7\frac{2}{15}$	$(64; 24^{\times 7}, 3^{\times 7}, 1^{\times 2})$	$(14,22)$	$(121,7)$	7.1333	7.1334	$\frac{641}{240}$	$3.25\,10^{-6}$
$7\frac{1}{7}$	$(24; 9^{\times 7}, 2, 1^{\times 6})$	$(7,8)$	$(57,1)$	7.1428	7.1429	$\frac{449}{168}$	$6.63\,10^{-6}$
$7\frac{2}{13}$	$(40; 15^{\times 7}, 2^{\times 6}, 1^{\times 2})$	$(14,13)$	$(107,0)$	7.151	7.156	$\frac{107}{40}$	$332.5\,10^{-6}$
$7\frac{1}{5}$	$(16; 6^{\times 7}, 1^{\times 5})$	$(7,5)$	$(43,0)$	7.1665	7.22	$\frac{43}{16}$	$4218.4\,10^{-6}$
$7\frac{1}{4}$	$(35; 13^{\times 7}, 4, 3^{\times 3})$	$(0,13)$	$(87,1)$	7.2485	7.252	$\frac{377}{140}$	$274.7\,10^{-6}$
$7\frac{1}{2}$	$(8; 3^{\times 7}, 1^{\times 2})$	$(7,2)$	$(22,0)$	7.328	7.56	$\frac{11}{4}$	$11387.2\,10^{-6}$
8	$(6; 3, 2^{\times 7})$	$(1,2)$	$(17,0)$	7.97	8.03	$\frac{17}{6}$	$4906.2\,10^{-6}$

Remark 5.2.4. (i) The above table gives just enough decimal places of the (irrational) numbers u_x, v_x to describe their important features. For example $u_{\frac{1}{2}} = \frac{1}{2}\left(9+4\sqrt{2}\right) \approx 7.328 < 7\frac{1}{3}$.

(ii) In the above table there is one constraint centered at each point of the form $7\frac{1}{k}$ for $2 \le k \le 8$, except for $k=3$ and $k=6$. In fact, there are classes $(d;\boldsymbol{m})$ giving constraints centered at $7\frac{1}{6}$ and $7\frac{1}{3}$, namely

$$\big(96; 36^{\times 6}, 35, 6^{\times 6}\big) \;\text{ at }\; 7\tfrac{1}{6} \qquad \text{and} \qquad \big(24; 9^{\times 6}, 8, 3^{\times 3}\big) \;\text{ at }\; 7\tfrac{1}{3}.$$

But these $(d;\boldsymbol{m})$ have the property that $\mu(d;\boldsymbol{m})(a) = c(a)$ only at their center points. (See the proof of Theorem 5.2.3 at the end of this section for details).

(iii) The four steps at the points $7\frac{1}{8}$, $7\frac{2}{15}$, $7\frac{1}{7}$, $7\frac{1}{4}$ are the only ones in the graph of $c(a)$ that are not flat to the right. ◇

To prove Theorem 5.2.3 we will proceed as follows. Assume that $(d;\boldsymbol{m}) \in \mathcal{E}$ is a class with $\ell(a) = \ell(\boldsymbol{m})$ and $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ for some $a \in [7\frac{1}{9}, 8]$. We first assume that $a = 7\frac{1}{k}$ for some $k \in \{1, \dots, 8\}$, and find all such classes $(d;\boldsymbol{m})$. We then assume that $a \in \,]7\frac{1}{k+1}, 7\frac{1}{k}[$, and prove an upper bound $D(z_k)$ for d if $a = z_k := 7\frac{2}{2k+1}$ and an upper bound D_k if $a \neq z_k$. In both cases, we also show that $m_1 = \cdots = m_7$. We then use a simple computer program to find all classes $(d;\boldsymbol{m})$ as above at z_k with $d \le D(z_k)$. Finally, we use another computer program to find all classes $(d;\boldsymbol{m})$ as above at some $a \neq z_k$ with $d \le D_k$.

We start by looking at the boundary points $7\frac{1}{k}$ of our subintervals $[7\frac{1}{k+1}, 7\frac{1}{k}]$.

Lemma 5.2.5. *The classes $(d;\boldsymbol{m}) \in \mathcal{E}$ such that $\ell(7\frac{1}{k}) = \ell(\boldsymbol{m})$ and $\mu(d;\boldsymbol{m})(7\frac{1}{k}) > \sqrt{7\frac{1}{k}}$ are*

(5.2.2)

k	$(d;\boldsymbol{m})$	k	$(d;\boldsymbol{m})$
8	$\big(48; 18^{\times 7}, 3, 2^{\times 7}\big)$	8	$\big(384; 144^{\times 6}, 143, 18^{\times 8}\big)$
7	$\big(24; 9^{\times 7}, 2, 1^{\times 6}\big)$	7	$\big(168; 63^{\times 6}, 62, 9^{\times 7}\big)$
6	$\big(96; 36^{\times 6}, 35, 6^{\times 6}\big)$	5	$\big(16; 6^{\times 7}, 1^{\times 5}\big)$
4	$\big(35; 13^{\times 7}, 4, 3^{\times 3}\big)$	3	$\big(24; 9^{\times 6}, 8, 3^{\times 3}\big)$
2	$\big(8; 3^{\times 7}, 1^{\times 2}\big)$	1	$\big(6; 3, 2^{\times 7}\big)$

Proof. We first look at the case $a = 7\frac{1}{1} = 8$. Then $\ell(\boldsymbol{m}) = \ell(8) = 8$. By Lemma 2.1.7 we need to consider 3 cases, namely $\boldsymbol{m} = (M^{\times 8})$, $\boldsymbol{m} = (M+1, M^{\times 7})$, $\boldsymbol{m} = (M^{\times 7}, M-1)$. Consider the case $\boldsymbol{m} = (M^{\times 8})$. From the Diophantine equations

$$\begin{cases} 3d &= 8M + 1 \\ d^2 &= 8M^2 - 1 \end{cases}$$

we obtain $(8M+1)^2 = 9\big(8M^2 - 1\big)$, i.e. $4M^2 - 8M - 5 = 0$. This equation has no solution in $\mathbb{N}$. In the case $\boldsymbol{m} = (M+1, M^{\times 7})$, the Diophantine equations give

$$(8M + 1 + 1)^2 \;=\; 9\big(8M^2 + 2M + 1 - 1\big)$$

whose only solution in $\mathbb{N}$ is $M = 2$, giving the solution $(d;\boldsymbol{m}) = (6; 3, 2^{\times 7})$. In the case $(M^{\times 7}, M-1)$, the Diophantine equations give $(8M-1+1)^2 = 9\big(8M^2 - 2M + 1 - 1\big)$, which has no solution in $\mathbb{N}$.

Assume now that $k \in \{2, \dots, 8\}$. In view of Lemma 2.1.7, there are five possibilities for $\boldsymbol{m}$, namely

$$(M^{\times 7}, m^{\times k}),\ \ (M+1, M^{\times 6}, m^{\times k}),\ \ (M^{\times 6}, M-1, m^{\times k}),$$
$$(M^{\times 7}, m+1, m^{\times (k-1)}),\ \ (M^{\times 7}, m^{\times (k-1)}, m-1).$$

Since $\ell(\boldsymbol{m}) = \ell(7\frac{1}{k}) = 7+k$, in the first four cases we can assume that $m \geq 1$ and in the last case we can assume that $m-1 \geq 1$. We define ε_M and ε_m by

$$M = \frac{d}{\sqrt{7\frac{1}{k}}} + \varepsilon_M, \quad m = \frac{d}{k\sqrt{7\frac{1}{k}}} + \varepsilon_m.$$

Case 1. $\boldsymbol{m} = (M^{\times 7}, m^{\times k})$**.** Then $|M - km| = |\varepsilon_M - k\varepsilon_m| \leq |\varepsilon_M| + k|\varepsilon_m|$. Since $|\varepsilon_M|^2 + k|\varepsilon_m|^2 < 1$, we find $|\varepsilon_M| + k|\varepsilon_m| < \sqrt{k+1}$, and so $|M - km| \leq \lceil \sqrt{k+1} - 1 \rceil \in \{0,1,2\}$. Set

$$s \,=\, M - km \,\in\, \begin{cases} \{0, \pm 1\} & \text{if } \ k \in \{2,3\}, \\ \{0, \pm 1, \pm 2\} & \text{if } \ k \in \{4, \dots, 8\}. \end{cases}$$

From the Diophantine equations

$$\begin{cases} 3d &= \ 7M + km + 1 \\ d^2 &= \ 7M^2 + km^2 - 1 \end{cases}$$

we obtain $(7M+km+1)^2 = 9\left(7M^2 + km^2 - 1\right)$. Since $M = km + s$, this becomes

$$10 + km\,(16 - 9m + km) + 14s\,(1 - km - s) \,=\, 0.$$

If $s = 1$, this is

$$10 + km\,(2 - 9m + km) \,=\, 0,$$

which has solutions in $\mathbb{N}$ only if $k = 5$ or 2, namely $m = 1$, giving

$$\left(16; 6^{\times 7}, 1^{\times 5}\right) \text{ at } 7\tfrac{1}{5}, \qquad \left(8; 3^{\times 7}, 1^{\times 2}\right) \text{ at } 7\tfrac{1}{2}.$$

No other allowed values for s and k yield integer solutions m.

Case 2. $\boldsymbol{m} = (M+1, M^{\times 6}, m^{\times k})$**.** Then $\sigma = k|\varepsilon_m|^2 \leq \frac{1}{7}$. Therefore, $|M - km| \leq |\varepsilon_M| + k|\varepsilon_m| \leq \frac{1}{\sqrt{6}} + \sqrt{\frac{k}{7}}$, and so

$$s := M - km \,\in\, \begin{cases} \{0\} & \text{if } \ k = 2, \\ \{0, \pm 1\} & \text{if } \ k \in \{3, \dots, 8\}. \end{cases} \tag{5.2.3}$$

In this case, the Diophantine equations translate to

$$(7M + km + 2)^2 \,=\, 9\left(7M^2 + 2M + 1 + km^2 - 1\right).$$

With $M = km + s$ this becomes

$$-4 - km\,(14 - 9m + km) + 2s\,(-5 + 7km + 7s) \,=\, 0. \tag{5.2.4}$$

If $s=0$, this becomes

$$-4-km\,(14-9m+km)\;=\;0$$

which has no solution in $\mathbb{N}$ for $k\in\{2,\dots,8\}$. For $s=\pm1$ and $k\in\{3,\dots,8\}$ equation (5.2.4) has no solution in $\mathbb{N}$.

Case 3. $\boldsymbol{m}=(M^{\times 6},M-1,m^{\times k})$**.** As in Case 2 we have (5.2.3). In this case, the Diophantine equations translate to

$$(7M+km)^2\;=\;9\left(7M^2-2M+1+km^2-1\right).$$

With $M=km+s$ this becomes

$$-km\,(18-9m+km)+2s\,(-9+7km+7s)\;=\;0. \tag{5.2.5}$$

If $s=0$, this becomes

$$18-9m+km\;=\;0$$

which has a solution in $\mathbb{N}$ for four k, namely $k=8,7,6$, and 3. This gives the first four of the five entries in the table with $m_1\neq m_7$.

If $s=1$, equation (5.2.5) becomes

$$-4-km\,(4-9m+km)\;=\;0,$$

which has a solution in $\mathbb{N}$ only for $k=4$. We get the solution $(13;5^{\times 6},4,1^{\times 4})$, which is, however, not obstructive, since it gives $\mu(d;\boldsymbol{m})(7\frac14)=\frac{35}{13}<\sqrt{7\frac14}$.

If $s=-1$, equation (5.2.5) becomes

$$32-km\,(32-9m+km)\;=\;0.$$

It has a solution in $\mathbb{N}$ only for $k=2$, and gives $(19;7^{\times 6},6,4^{\times 2})$. But again this class is not obstructive, since $\mu(d;\boldsymbol{m})(7\frac12)=\frac{52}{19}<\sqrt{7\frac12}$.

Case 4. $\boldsymbol{m}=(M^{\times 7},m+1,m^{\times (k-1)})$**.** Note that for $\varepsilon\in\mathbb{R}$ and $k\in\mathbb{N}$ with $(k-1)\varepsilon^2+(\varepsilon+1)^2\le 1$ we have $\varepsilon\in\left[-\frac{2}{k},0\right]$ and hence

$$|(k-1)\varepsilon+(\varepsilon+1)|\;=\;|k\varepsilon+1|\;\le\;1.$$

Using this and $\sigma\ge\frac{k-1}{k}$ we estimate

$$\begin{aligned}|M-km-1|\;=\;|M-(m+1)-(k-1)m| &\;=\; |\varepsilon_M-(\varepsilon_m+1)-(k-1)\varepsilon_m|\\ &\;\le\; |\varepsilon_M|+|(k-1)\varepsilon_m+\varepsilon_m+1|\\ &\;\le\; \sqrt{\tfrac{1}{7k}}+1\;<\;2.\end{aligned}$$

Therefore,

$$M-1\;=\;km+s\quad\text{ with }\;s\in\{0,\pm1\}.$$

In this case, the Diophantine equations translate to

$$(7M+km+1+1)^2\;=\;9\left(7M^2+km^2+2m+1-1\right).$$

With $M=km+1+s$ this becomes

$$-18+18m-km\,(18-9m+km)+14s\,(km+s)\;=\;0. \tag{5.2.6}$$

If $s=1$, this is

$$-4+18m-km\,(4-9m+km)\;=\;0,$$

which has a solution in $\mathbb{N}$ only when $k=8, m=2$ and $k=7, m=1$, giving us two more entries in our table. If $s=0$, equation (5.2.6) becomes

$$-18+18m-km\,(18-9m+km)\;=\;0,$$

which has a solution in $\mathbb{N}$ only for $k=4, m=3$. This gives the entry in the table at $k=4$. If $s=-1$, equation (5.2.6) has no solution in $\mathbb{N}$ for $k\in\{2,\dots,8\}$.

Case 5. $\boldsymbol{m}=(M^{\times 7}, m^{\times(k-1)}, m-1)$**.** As in Case 4 we find

$$M+1\;=\;km+s \quad \text{ with } s\in\{0,\pm 1\}.$$

In this case, the Diophantine equations translate to

$$(7M+km-1+1)^2\;=\;9\left(7M^2+km^2-2m+1-1\right).$$

With $M=km-1+s$ this becomes

$$-14+18m+km\,(14-9m+km)+14s\,(2-km-s)\;=\;0. \tag{5.2.7}$$

If $s=1$, this becomes

$$18-9km+k^2m\;=\;0.$$

It has a solution in $\mathbb{N}$ only for $k=6$ and $k=3$, namely $m=1$. Since we assumed that $m-1\geq 1$, the corresponding classes $(d;\boldsymbol{m})$ are not relevant. If $s=0$ or if $s=-1$, equation (5.2.7) has no solution in $\mathbb{N}$ for $k\in\{2,\dots,8\}$.

The above calculations show that the elements listed in Table 5.2.2 are the only obstructive solutions to the Diophantine equations. One readily checks that these elements all reduce to $(0;-1)$ under standard Cremona moves, and therefore belong to $\mathcal{E}$. □

Remark 5.2.6. Proceeding as in the proof of Lemma 5.2.5, one can find all classes $(d;\boldsymbol{m})\in\mathcal{E}$ with $\mu(d;\boldsymbol{m})(z_k)>\sqrt{z_k}$ and $\ell(z_k)=\ell(\boldsymbol{m})$ at the points $z_k:=7\frac{2}{2k+1}$, $k\in\{1,\dots,8\}$, namely

$$\left(64;24^{\times 7},3^{\times 7},1^{\times 2}\right) \text{ at } 7\tfrac{2}{15} \qquad \text{and} \qquad \left(40;15^{\times 7},2^{\times 6},1^{\times 2}\right) \text{ at } 7\tfrac{2}{13}.$$

For convenience, we will find these classes by a different method, that involves the first of the two computer programs of Appendix B. ◇

We next derive upper bounds for d if $a\in\,]7\frac{1}{k+1},7\frac{1}{k}[$. There are various ways to do this. We will give arguments that give rather low upper bounds, so that our method of finding $c(a)$ depends as little as possible on computer computations (compare Remark 5.2.9 below). Note that $a\in\,]7\frac{1}{k+1},7\frac{1}{k}[$ has $N+1$ blocks with $N\geq 2$, and that $L:=\sum_{i\geq 1}\ell_i\geq 2+k$ with equality exactly if $a=[7;k,2]=7\frac{2}{2k+1}=z_k$.

Lemma 5.2.7. *Suppose that* $(d;\boldsymbol{m})\in\mathcal{E}$ *is such that* $\mu(d;\boldsymbol{m})(a)=c(a)>\sqrt{a}$ *for some* a *with* $\ell(a)=\ell(\boldsymbol{m})$*. Suppose also that* a *has* $N+1$ *blocks for some* $N\geq 2$ *and that* $a\in\,]7\frac{1}{k+1},7\frac{1}{k}[$ *where* $8\geq k\geq 1$*. Then* $m_1=m_7$*. Further*

(i) *When $L = 2 + k$, the following table gives the maximum possible values $D(z_k)$ of d for the different k.*

(5.2.8)

k	8	7	6	5	4	3	2	1
$D(z_k)$	104	98	92	86	79	73	69	75

(ii) *When $L > 2 + k$, the following table gives the maximum possible values D_k of d for the different k.*

(5.2.9)

k	8	7	6	5	4	3	2	1
D_k	88	81	74	67	61	56	64	66

Remark 5.2.8. By Lemma 5.2.5 one cannot conclude $m_1 = m_7$ without the assumption $a \neq 7\frac{1}{k}$. ◇

Proof. The proof of this lemma is based on an analysis of the equation (5.1.1) using the estimates for $|\sum \varepsilon_i|$ obtained in Lemma 5.1.2. Recall from (iv) of that lemma that for $a = 7\frac{p}{q}$ and with $v_M := \frac{d}{q\sqrt{a}}$ we have the estimates

$$d \leq \frac{\sqrt{a}}{\delta}\left(\sqrt{\sigma L} - 1\right) \leq \frac{\sqrt{a}}{\delta}\left(\sqrt{\sigma q} - 1\right) < \frac{\sqrt{a}}{\delta}\left(\frac{\sigma}{v_M \delta} - 1\right) \tag{5.2.10}$$

whenever $\delta := y(a) - \frac{1}{q} > 0$.

(i) The only number in $]7\frac{1}{k+1}, 7\frac{1}{k}[$ with $L = k + 2$ is $z_k := [7; k, 2] = 7\frac{2}{2k+1}$. Note that

$$y(z_k) - \tfrac{1}{q} \;=\; 8 + \tfrac{1}{2k+1} - 3\sqrt{7 + \tfrac{2}{2k+1}} \;>\; 0$$

for all k. By (5.2.10) we therefore have

$$d \leq \frac{\left(\sqrt{\sigma(k+2)} - 1\right)\sqrt{z_k}}{y(z_k) - \frac{1}{2k+1}} = \frac{\left(\sqrt{\sigma(k+2)} - 1\right)\sqrt{7 + \frac{2}{2k+1}}}{8 + \frac{1}{2k+1} - 3\sqrt{7 + \frac{2}{2k+1}}}$$

With $\sigma \leq 1$ this yields Table 5.2.8. If $m_1 \neq m_7$ we may take $\sigma \leq \frac{1}{7}$. The largest value of d is then ≤ 6 when $k \leq 7$ and ≤ 9 when $k = 8$. But there are clearly no suitable $(d; \boldsymbol{m})$ with such small d. Therefore this case does not occur. This proves $m_1 = m_7$ for $L = 2 + k$.

We will prove (ii) and the claim that $m_1 = m_7$ together. We will give separate arguments for the three cases $k = 1$, $k = 2$ and $k \in \{3, \dots, 9\}$. Denote $a_k = 7\frac{1}{k+1}$ for some $1 \leq k \leq 8$. We have the table (rounded down to 3 decimal places)

$k =$	8	7	6	5	4	3	2	1
$y(a_k) \geq$	$\frac{1}{9} = 0.111$	0.117	0.125	0.135	0.150	0.172	0.209	0.284

The case $k = 1$. Assume that $(d; \boldsymbol{m}) \in \mathcal{E}$ is a class with $\mu(d; \boldsymbol{m})(a) > \sqrt{a}$ and $\ell(a) = \ell(\boldsymbol{m})$ for some $a \in]7\frac{1}{2}, 8[$ other than $7\frac{2}{3}$. We first prove that $m_1 = m_7$. If not, then $\sigma \le \frac{1}{7}$. Therefore, $v_M = \frac{d}{q\sqrt{a}} \ge 1 - \frac{1}{\sqrt{14}} > 0.73$, since otherwise $\sigma \ge \varepsilon_M^2 + \varepsilon_{M-1}^2 > 2\frac{1}{14} = \frac{1}{7}$. Also, $q \ge L \ge 3 + k = 4$, and so $y(a) - \frac{1}{q} \ge y(7\frac{1}{2}) - \frac{1}{4} \ge 0.28 - \frac{1}{4} > 0$. We can therefore apply (5.2.10):

$$\sqrt{\tfrac{q}{7}} \;\ge\; \sqrt{\sigma L} \;\ge\; 1 + \tfrac{d}{\sqrt{a}}\left(y(a) - \tfrac{1}{q}\right) \;>\; 1,$$

showing that $q \ge 8$. Therefore, $y(a) - \frac{1}{q} > y(7\frac{1}{2}) - \frac{1}{8} > 0.28 - \frac{1}{8} > \frac{1}{7}$. Using again (5.2.10) we finally find

$$5\sqrt{a} \;<\; 8 \cdot 0.73\sqrt{a} \;\le\; 0.73q\sqrt{a} \;<\; d \;<\; \frac{\sqrt{a}}{\frac{1}{7}}\left(\frac{\sigma}{v_M\frac{1}{7}} - 1\right) \;<\; 7\sqrt{a}\left(\frac{1}{0.73} - 1\right) \;<\; 3\sqrt{a},$$

a contradiction.

We now prove that $d \le 66$. For a as above, both numbers

$$\begin{aligned} f(a,q) \;&:=\; \frac{\sqrt{a}}{a + 1 - 3\sqrt{a} - \frac{1}{q}}\left(\sqrt{q} - 1\right), \\ g(a,q) \;&:=\; \frac{\sqrt{a}}{a + 1 - 3\sqrt{a} - \frac{1}{q}}\left(\frac{2}{a + 1 - 3\sqrt{a} - \frac{1}{q}} - 1\right), \end{aligned}$$

are positive. Moreover, by (5.2.10) we have $d \le f(a,q)$, and using also Lemma 5.1.2 (iv) we see that $d < g(a,q)$. We saw above that $q \ge 4$. We first use the function f to see that for $q \in \{4,5,6,7,8\}$ we have $d \le 26$. Assume now that $q \ge 9$. We then view a and q as independent variables of the functions f and g. Both $f(a,q)$ and $g(a,q)$ are decreasing functions of a. With $a_1 = 7\frac{1}{2}$ we therefore have

$$d \;\le\; \max_{q \ge 9} \min\left\{f(a_1,q),\, g(a_1,q)\right\}.$$

One readily checks that $f(a_1,q)$ is increasing on $\{q \ge 9\}$ and that $g(a_1,q)$ is decreasing in q. Since $d \le f(a_1,56) < 67$ if $q \le 56$ and $d \le g(a_1,57) < 67$ if $q \ge 57$ we conclude that $d \le 66$, as claimed. ◇

Remark 5.2.9. This method for estimating d can be used for all $k \le 8$. However, the estimates get worse, e.g. for $k = 8$ (with the factor $\sqrt{q} - 1$ of f replaced by $\sqrt{8 + \frac{q}{8}} - 1$, see (5.2.11) below) one finds $d \le 410$. One could also omit checking that obstructive classes have $m_1 = m_7$, and use a variant of our computer code `SolLess` from Appendix B.1 that does not use $m_1 = m_7$. ◇

The case $k = 2$**.** The class $(8; 3^{\times 7}, 1^{\times 2})$ gives the constraint $c(a) \ge \mu_0(a) = \frac{7+2a}{8} > \sqrt{a}$ on $[7\frac{1}{3}, 7\frac{1}{2}]$. Assume that $(d;\boldsymbol{m}) \in \mathcal{E}$ is a class with $\mu(d;\boldsymbol{m})(a) = c(a) \ge \frac{7+2a}{8}$ for some $a \in [7\frac{1}{3}, 7\frac{1}{2}]$. Proposition 2.1.1 (i) implies that

$$\frac{7 + 2a}{8} \le \mu(d;\boldsymbol{m})(a) \le \sqrt{a}\sqrt{1 + 1/d^2}.$$

When $a = 7\frac{1}{3}$ this gives the estimate $d \le 64$. Since $\frac{7+2a}{\sqrt{a}}$ decreases on $[7\frac{1}{3}, 7\frac{1}{2}]$, we find $d \le 64$ everywhere. We will check $m_1 = m_7$ for $k = 2$ and $L \ge 3 + k = 5$ at the same time as for $k \ge 3$. $\diamond$

The case $k \in \{3, \dots, 9\}$. Suppose that $a \in \,]7\frac{1}{k+1}, 7\frac{1}{k}[$ for some $k \ge 2$, and that $L \ge 3 + k$. Then we may write

$$a = [7; k, \ell_2, \dots, \ell_N] = 7 + \frac{1}{k + \frac{p'}{q'}}$$

where $\frac{p'}{q'} := a' := [0; \ell_2, \dots, \ell_N]$. Thus $q' = q_{N-1}(a') \ge \sum_{j \ge 2} \ell_j =: L'$ by Sublemma 5.1.1, and so

$$L := \sum_{j \ge 1} \ell_j \;=\; k + L' \;\le\; k + q'.$$

Since $q = kq' + p'$ we find $L \le k + \frac{q}{k}$. Moreover, $q' \ge L' = L - k \ge 3$, and so $q \ge 3k + 1$. Therefore, for $a \in \,]a_k, a_{k-1}[$ we have $y(a) \ge y(a_k) > \frac{1}{3k+1} > \frac{1}{q}$. Thus the inequality (5.2.10) implies that

$$\left(1 - \tfrac{d}{q\sqrt{a}}\right) + \tfrac{d}{\sqrt{a}} y(a) \;=\; 1 + \tfrac{d}{\sqrt{a}} \left(y(a) - \tfrac{1}{q}\right) \;\le\; \sqrt{\sigma(k + \tfrac{q}{k})}. \tag{5.2.11}$$

Case 1: $\frac{1}{2} \le v_M := \frac{d}{q\sqrt{a}} \le \frac{3}{4}$.

Because $y(a) \ge y(a_k)$ for all $a \in [a_k, a_{k-1}]$ and $y(a_8) = \frac{1}{9}$, we must have

$$\tfrac{q}{18} \;\le\; \tfrac{1}{4} + \tfrac{q}{2} y(a_k) \;\le\; \tfrac{1}{4} + \tfrac{d}{\sqrt{a}} y(a_k) \;\le\; \left(1 - \tfrac{d}{q\sqrt{a}}\right) + \tfrac{d}{\sqrt{a}} y(a) \;\le\; \sqrt{\sigma'(k + \tfrac{q}{k})}\,,$$

where $\sigma' \le \frac{7}{8}$ is as in Lemma 5.1.2 (iii).

Note that the squared error of the last two ε_i is at least $2\left(\frac{1}{4}\right)^2 = \frac{1}{8}$. Therefore, if also $m_1 \ne m_7$, we have $\sigma' < \frac{1}{7} - \frac{1}{8} = \frac{1}{56}$. But, for each $k \in [2, 8]$, the inequality

$$\tfrac{q}{18} \le \sqrt{\tfrac{1}{56}(k + \tfrac{q}{k})}$$

holds only if $q^2 \le 6(k + \frac{q}{k})$. Since this quadratic inequality holds for $q = 0$ and does not hold when $q = 3k + 1$, it does not hold for any $q \ge 3k + 1$. Therefore, for each k we have $m_1 = m_7$, and $\sigma' \le \frac{7}{8}$.

Now suppose that $k = 8$, and consider the inequality

$$\tfrac{1}{4} + \tfrac{q}{2} y(a_8) \;=\; \tfrac{1}{4} + \tfrac{q}{18} \le \sqrt{\tfrac{7}{8}(8 + \tfrac{q}{8})}.$$

This holds when $q = 0$ but does not hold for $q \ge 63$. Thus $q \le 62$ so that $\frac{d}{\sqrt{a}} \frac{1}{9} \le \sqrt{\frac{7}{8}\left(8 + \frac{62}{8}\right)} - \frac{1}{4}$. Since $a \le 7\frac{1}{8}$ we get $d \le 83$. The same argument works for the other k, and we obtain the following upper bounds for q and then for d.

$k =$	8	7	6	5	4	3
$q \le$	62	58	53	49	45	41
$d \le$	83	77	71	66	61	56

Case 2: $v_M \le \frac{1}{2}$.

Since the squared error $\ell_N \delta_N^2$ on the last block is now at least $\frac{1}{2}$, we must have $m_1 = m_7$ and $\sigma' < \frac{1}{2}$. Further, by Lemma 5.1.2 (iii), $\frac{d}{q\sqrt{a}} \ge \frac{1}{3}$. Therefore (5.2.11) gives

$$\tfrac{1}{2} + \tfrac{q}{3} y(a_k) \;\le\; \tfrac{1}{2} + \tfrac{d}{\sqrt{a}} y(a_k) \;\le\; \sqrt{\tfrac{1}{2}(k + \tfrac{q}{k})}\,.$$

This gives the following upper bounds for q and d.

$k =$	8	7	6	5	4	3
$q \le$	62	57	53	49	45	42
$d \le$	55	51	47	43	40	37

Case 3: $\frac{3}{4} \le v_M \le 1$.

Now (5.2.11) gives

$$\tfrac{3}{4}\, q\, y(a_k) \;\le\; \tfrac{d}{\sqrt{a}}\, y(a_k) \;\le\; \sqrt{\sigma(k + \tfrac{q}{k})}\,.$$

If $\sigma \le \frac{1}{7}$, this is not satisfied when $q \ge 3k+1$ for any $k \in \{2, \dots, 8\}$. Thus $m_1 = m_7$.

Further, taking $\sigma = 1$ we obtain the following upper bounds for q and d.

$k =$	8	7	6	5	4	3
$q \le$	44	40	37	33	30	26
$d \le$	88	81	74	67	60	53

Case 4: $1 \le v_M$.

In this case, (5.2.11) gives

$$q\, y(a_k) \;\le\; 1 + q\left(y(a_k) - \tfrac{1}{q}\right) \;\le\; 1 + \tfrac{d}{\sqrt{a}}\big(y(a_k) - \tfrac{1}{q}\big) \;\le\; 1 + \tfrac{d}{\sqrt{a}}\big(y(a) - \tfrac{1}{q}\big) \;\le\; \sqrt{\sigma(k + \tfrac{q}{k})}\,.$$

We have already seen in Case 3 that $q\, y(a_k) \le \sqrt{\sigma(k + \frac{q}{k})}$ is impossible for $\sigma \le \frac{1}{7}$. Thus $m_1 = m_7$.

Further, taking $\sigma = 1$ we obtain the following upper bounds for q and d.

$k =$	8	7	6	5	4	3
$q \le$	31	28	25	22	19	17
$d \le$	82	75	68	61	54	46

Taking for each k the worst upper bound for d in the different cases, we obtain Table 5.2.9. This completes the proof of Lemma 5.2.7. $\square$

Corollary 5.2.10. (i) *The only classes $(d;\boldsymbol{m}) \in \mathcal{E}$ such that $\ell(z_k) = \ell(\boldsymbol{m})$ and such that $\mu(d;\boldsymbol{m})(z_k) > \sqrt{z_k}$ are*

$$7\tfrac{2}{15}\colon\ \big(64; 24^{\times 7}, 3^{\times 7}, 1^{\times 2}\big) \qquad \textit{and} \qquad 7\tfrac{2}{13}\colon\ \big(40; 15^{\times 7}, 2^{\times 6}, 1^{\times 2}\big)$$

(ii) *There are no classes $(d;\boldsymbol{m}) \in \mathcal{E}$ such that $\ell(a) = \ell(\boldsymbol{m})$ and $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ for some $a \in \,]7\frac{1}{9}, 8[$ not of the form $7\frac{1}{k}$ or $7\frac{2}{2k+1}$.*

Proof. (i) The computer code `SolLess[a,D]` given in Appendix B.1 finds for a rational number a and a natural number D all classes $(d;\boldsymbol{m}) \in \mathcal{E}$ with $\ell(\boldsymbol{m}) = \ell(a)$ and $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ and $d \leq D$. For $k \in \{1,\dots,8\}$ we choose $D = D(z_k)$ as given by Table (5.2.8). The code `SolLess[a,D]` with $D = D(z_k)$ and $a = z_k$ tells us that for $k = 7$ and $k = 6$, the only such classes are the ones given in the corollary, while for the other k there are no such classes. Finally, one checks that the two classes in (i) reduce to $(0;-1)$ under standard Cremona moves, and hence belong to $\mathcal{E}$.

(ii) The computer code `InterSolLess[k,D]` given in Appendix B.2 provides for a natural number D a finite list of candidate classes $(d;\boldsymbol{m}) \in \mathcal{E}$ with $\ell(\boldsymbol{m}) = \ell(a)$ and $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ and $d \leq D$ for some $a \in \,]7\frac{1}{k+1}, 7\frac{1}{k}[$. For $k \in \{1,\dots,8\}$ we choose $D = D_k$ as given by Table (5.2.9). The code `InterSolLess[k,D]` with $D = D_k$ tells us that for $k \neq 4$ there are no candidate classes, while for $k = 4$ the only candidate class is $(d;\boldsymbol{m}) = \big(59; 22^{\times 7}, 5^{\times 3}, 4, 1^{\times 3}\big)\big)$. Since the length of the second block is 4 and the length of last block is ≥ 2, the a in question must be $[7;4,3]$ or $[7;4,1,2]$. The second possibility is excluded by Lemma 2.1.8 (i) applied to the third block. Moreover, at $a = [7;4,3] = 7\frac{3}{13}$ we have $\mu(d;\boldsymbol{m})(a) = \frac{2062}{767} < \sqrt{a}$, which excludes also the first possibility. $\square$

Proof of Theorem 5.2.3. Recall from Proposition 1.2.9 that $c(7\frac{1}{9}) = \sqrt{7\frac{1}{9}}$. Moreover, by Lemma 2.1.3 any class $(d;\boldsymbol{m}) \in \mathcal{E}$ with $\mu(d;\boldsymbol{m})(8) > \sqrt{8}$ must lie in $\mathcal{E}_8$. By looking at the list of elements in $\mathcal{E}_8$ given in Lemma 1.2.7 one checks that the only such class is $\big(6;3,2^{\times 7}\big)$. By using Lemma 2.1.3 once more, we conclude that all constraints on $[7\frac{1}{9}, 8]$ come from the ten classes of Lemma 5.2.5 and the two classes from Corollary 5.2.10.

In the paragraph just before Theorem 5.2.3 we worked out the constraint $\mu(d;\boldsymbol{m})$ given by the class centered at $7\frac{1}{8}$. Similar computations show that all the eight classes in Table 5.2.1 behave as described there. In order to prove Theorem 5.2.3, it therefore remains to check that the four classes from Lemma 5.2.5 that do not appear in Theorem 5.2.3 give no further constraints. However, one can calculate the corresponding functions $\mu(d;\boldsymbol{m})$ just as before, obtaining the following data.[7]

(5.2.12)

a	$(d;\boldsymbol{m})$	(A,B)	(A',B')	$\mu(a)$	$N(A,B)$	s
$7\frac{1}{8}$	$(384;144^{\times 6},143,18^{\times 8})$	$(-1,144)$	$(1025,0)$	$\frac{1025}{384}$	74322	18
$7\frac{1}{7}$	$(168;63^{\times 6},62,9^{\times 7})$	$(-1,63)$	$(449,0)$	$\frac{449}{168}$	14373	9
$7\frac{1}{6}$	$(96;36^{\times 6},35,6^{\times 6})$	$(-1,36)$	$(257,0)$	$\frac{257}{96}$	4758	6
$7\frac{1}{3}$	$(24;9^{\times 6},8,3^{\times 3})$	$(-1,9)$	$(65,0)$	$\frac{65}{24}$	327	3

In all cases the new constraint takes the same value at its center point as the old one but the slope to the left is steeper (because $A=-1$) and it is flat (i.e. with $B'=0$) rather than increasing to the right. This completes the proof. □

5.3. **The interval** $[8,9]$**.** In this section we compute $c(a)$ on the interval $[8,9]$. We first prove that $c(a)=\sqrt{a}$ for $a\ge 8\frac{1}{36}$.

Lemma 5.3.1. *Suppose that* $\mu(d;\boldsymbol{m})(a)>\sqrt{a}$ *for some* $a\in\left[8\frac{1}{36},9\right)$ *with* $\ell(a)=\ell(\boldsymbol{m})$*. Then* $d\le 16$ *and* $m_1=\cdots=m_8$*.*

Proof. Note that $y(a)\ge y(8\frac{1}{36})=\frac{19}{36}>\frac{1}{q}$ for all $q\ge 2$. Assume first that $q\ge 12$. Then $\delta:=y(a)-\frac{1}{q}\ge\frac{19}{36}-\frac{1}{12}=\frac{4}{9}$. Suppose that $m_1\neq m_8$. Then $\sigma\le\frac{1}{8}$, and hence $v_M\ge\frac{3}{4}$. This and $\delta>\frac{1}{6}$ shows that $\frac{\sigma}{v_M\delta}<1$, which is impossible by Lemma 5.1.2 (iv). In order to prove that $d\le 16$, note that

if $v_M\in[\frac{1}{3},\frac{1}{2}]$, then $\frac{\sigma'}{v_M}\le\frac{1/2}{1/3}=\frac{3}{2}$;
if $v_M\in[\frac{1}{2},\frac{2}{3}]$, then $\frac{\sigma'}{v_M}\le\frac{7/9}{1/2}=\frac{14}{9}$;
if $v_M\ge\frac{2}{3}$, then $\frac{\sigma}{v_M}\le\frac{3}{2}$.

Lemma 5.1.2 (iv) with $\sqrt{a}\le 3$ therefore shows that

$$d\;\le\;\frac{3}{4/9}\left(\frac{14}{9}\frac{1}{4/9}-1\right)\;=\;\frac{27}{4}\frac{5}{2}\;<\;17$$

and hence $d\le 16$.

Assume now that $q\le 11$. Note that $a\le 8\frac{q-1}{q}$ and $\delta=y(a)-\frac{1}{q}\ge y(8\frac{1}{q})-\frac{1}{q}$. Lemma (5.1.2) (iv) therefore shows that

$$d\;\le\;\frac{\sqrt{8\frac{q-1}{q}}}{y(8\frac{1}{q})-\frac{1}{q}}\left(\sqrt{q}-1\right).$$

The RHS is <17 for all $q\in\{2,\dots,11\}$, and so $d\le 16$. Suppose that $m_1\neq m_9$. Then $\sigma\le\frac{1}{8}$. If $q\le 8$, then $\sqrt{\sigma q}-1\le 0$, contradicting (iv) of Lemma 5.1.2. If $q\in\{9,10,11\}$, then

$$v_M\;=\;\frac{d}{q\sqrt{a}}\;\le\;\frac{16}{9\sqrt{8}}\;<\;\frac{2}{3},$$

[7]We also calculated the number $N(A,B)$ of integer points in the triangle $T^a_{A,B}$ and the number s of integer points on its slant edge because of their relevance to Remark 2.4.5.

and hence $\varepsilon \cdot \varepsilon \geq \frac{7}{8} + 2 \cdot \frac{1}{9} > 1$, a contradiction. □

Proposition 5.3.2. *$c(a) = \sqrt{a}$ for $a \in \big[8\frac{1}{36}, 9\big)$.*

Proof. Suppose to the contrary that $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ for some $a \geq 8\frac{1}{36}$. By Lemma 2.1.3 we may choose a_0 with $\ell(a_0) = \ell(\boldsymbol{m})$ in the interval I containing a on which this inequality holds.

We first claim that $a_0 > 8$. By Lemma 2.1.3 it suffices to see that $\ell(\boldsymbol{m}) > 8$. One can prove this by explicit calculation since $\mathcal{E}_8$ is finite. In fact, the last obstruction given by the elements of $\mathcal{E}_8$ is that centered on $a = 8$ which is discussed in Remark 1.2.10. As we saw there, this is not effective when $a > 8\frac{1}{36}$.

It then follows that $a_0 \geq 8\frac{1}{36}$. For if not, because I contains $a \geq 8\frac{1}{36}$, it must also contain $8\frac{1}{36}$. But clearly $\ell(z) > \ell(8\frac{1}{36}) = 8 + 36 = 44$ for $z \in (8, 8\frac{1}{36})$. Therefore the minimum of $\ell(z)$ on I cannot occur in this interval.

We may therefore apply Lemma 5.3.1 to a_0. Hence $d \leq 16$ and $m := m_1 = m_8$. Since $\sum m_i = 3d - 1 \leq 47$, we must have $m \leq 5$. It remains to check that there are no solutions to the Diophantine equations (1.2.4) for any choice of $m \leq 5$.

Suppose first that $m = 5$. We then look for solutions of

$$3d - 1 = 40 + \sum_{i>8} m_i, \quad d^2 + 1 = 200 + \sum_{i>8} m_i^2. \tag{5.3.1}$$

The second equation shows that $d \in \{15, 16\}$. For $d = 15$, (5.3.1) becomes $4 = \sum_{i>8} m_i$, $26 = \sum_{i>8} m_i^2$, which has no solution. For $d = 26$, (5.3.1) becomes $7 = \sum_{i>8} m_i$, $57 = \sum_{i>8} m_i^2$, which has no solution either. For $m \leq 5$ there are no solution either. □

Corollary 5.3.3. *$c(a) = \frac{17}{6}$ for $a \in [8, 8\frac{1}{36}]$.*

Proof. The class $(d;\boldsymbol{m}) = (6; 3, 2^{\times 7})$ gives $c(a) \geq \mu(d;\boldsymbol{m})(a) = \frac{17}{6} = \sqrt{8\frac{1}{36}}$ for $a \geq 8$. Therefore c must be constant on this interval because it cannot decrease. □

Appendix A. Weight expansions and Farey diagrams

In this section, we show that the weight expansion $\boldsymbol{w}(a)$ described above agrees with the expansion considered in [16]. For clarity, we call the latter the Farey weight expansion; see Definition A.7. It arose from a procedure of constructing an outer approximation to an ellipsoid by repeated blowing up. After explaining this, we establish the equivalence of the two definitions in Corollary A.10. No doubt, versions of this result are already known. However, since it is not hard, we give a direct proof in our context.

Definition A.4. *Let $(\rho_i = p_i/q_i)$, $i = 0, \dots, N$, be a sequence of rational numbers in lowest terms, with $\rho_0 = 0/1, \rho_1 = 1/1$ and $\rho_i > 0$ for $i > 0$. We say that the two elements ρ_j, ρ_k are* **adjacent** *in (ρ_i) if they are neighbors when the numbers $\rho_0, \dots, \rho_N$ are arranged in increasing order. Further (ρ_i) is called a* **Farey expansion** *of the rational number a if the following conditions hold:*

(i) *$\rho_N = a$;*

(ii) *$q_i < q_{i+1}$ for all $i \geq 1$;*

(iii) *adjacent pairs $p/q, p'/q'$ of elements of (ρ_i) have the property that*

$$|pq' - p'q| = 1; \tag{A.0.2}$$

(iv) *condition* (iii) *does not hold if any term is removed from this expansion.*

Example A.5. The Farey expansion of $\frac{4}{7}$ is $1, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{4}{7}$ which may be arranged as

$$\tfrac{1}{2} < \tfrac{4}{7} < \tfrac{3}{5} < \tfrac{2}{3} < \tfrac{1}{1}.$$

Lemma A.6. (i) *Every positive rational number a has a Farey expansion.*
(ii) *This expansion is unique. Moreover, if $\rho_1, \dots, \rho_N$ is the Farey expansion of $a = \rho_N$, then for all $n < N$, $\rho_1, \dots, \rho_n$ is the Farey expansion of ρ_n.*

Sketch of proof. Given positive fractions $\rho_i := \frac{p_i}{q_i}$ and $\rho_j := \frac{p_j}{q_j}$, we define their Farey sum to be

$$\rho_i \oplus \rho_j := \frac{p_i + p_j}{q_i + q_j}.$$

If $0 < a < 1$, the expansion is constructed inductively starting with $\rho_0 = 0$ and $\rho_1 = 1$, in such a way that $\rho_{i+1} := \rho_i \oplus \rho_j$ where j is the largest number $< i$ such that a lies between ρ_i and ρ_j. Thus $\rho_2 = \frac{1}{2}$, and ρ_3 is either $\frac{1}{3}$ (if $a < \rho_2$) or $\frac{2}{3}$ (if $a > \rho_2$). The construction stops when $a = \rho_N$.

If a lies between k and $k+1$, then the expansion begins with the terms $\rho_i := \frac{i}{1}$, $i = 1, \dots, k+1$. Then $\rho_{k+2} = \rho_k \oplus \rho_{k+1} = \frac{2k+1}{2}$ and the expansion proceeds as in the previous case. Further details may be found in Hardy and Wright, [8, Ch. III]. □

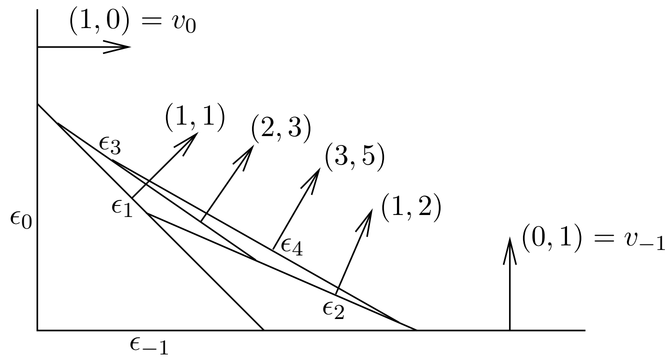

Figure A.1. The Farey diagram for $p/q = 5/3$. Here $v_1 = (1,1)$, $v_2 = (1,2)$, $v_3 = (2,3)$, $v_4 = (3,5)$. The edge ϵ_2 meets ϵ_1 and ϵ_0, while ϵ_3 meets ϵ_1 and ϵ_2, and ϵ_4 meets ϵ_2 and ϵ_3.

One can build a diagram in $\mathbb{R}^2$ corresponding to a given Farey expansion by associating to each fraction p_i/q_i a line segment ϵ_i (called an *edge*) with normal vector $v_i := (q_i, p_i)$ of slope p_i/q_i. See Figure A.1. One starts with the first quadrant whose edges $\epsilon_{-1}, \epsilon_0$ are the positive coordinate axes with (inward) normals $(q_{-1}, p_{-1}) = (1, 0)$ and $(q_0, p_0) = (0, 1)$, and builds up a sequence of edges by cutting along certain directions. The first cut is along an edge ϵ_1 going from ϵ_{-1} to ϵ_0 with normal $v_1 = (1, 1)$. In general, if ρ_i is the Farey sum of ρ_j with ρ_{i-1} for some $j < i - 1$ then the ith cut is along an edge ϵ_i with normal v_i that meets the edges ϵ_j and ϵ_{i-1} (but none of the others). The collection of edges $\epsilon_1, \dots, \epsilon_N$ is called the **Farey diagram**; the **extended Farey diagram** also includes the edges $\epsilon_{-1}, \epsilon_0$.

As described in [16, §3], adding a new edge whose normal is the sum of the two adjacent normals corresponds to a (smooth) blow up, since in the toric model, each blow up corresponds to cutting off a corner of the moment polytope. Therefore we can think of the process of constructing the Farey expansion for a as the process of blowing up the first quadrant repeatedly and in as efficient a way as possible, in order to obtain a (smooth) polytope with one edge whose normal has slope a. In the language of [16], this is an outer approximation; see Figure 3.1 and Lemma 3.8 ff. in [16]. For further discussion of the relation between weight sequences and the resolution of singularities by blow up, see the end of [17]. This contains a description of the Riemenschneider staircase that links the weight expansion for a to the Hirzebruch–Jung continued fraction expansions for the two singular points at the vertices of the toric model of the ellipsoid $E(1, a)$.

Given such a sequence of edges $\epsilon_1, \dots, \epsilon_N$ one can define an associated sequence of **Farey labels** $\lambda_1, \dots, \lambda_N$ as follows, starting with the last ϵ_N that is labeled by $\lambda_N := 1$.

(∗) *If $\epsilon_j, j > n$, is labeled by λ_j, label ϵ_n with the sum of the labels of the edges $\epsilon_j, j > n$, that intersect ϵ_n.*

Definition A.7. *If $a = p/q$ has Farey diagram with labels $\lambda_i, 1 \le i \le N$, the* **Farey weights** *of a are the numbers $u_i := \lambda_i/\lambda_1$, for $i = 1, \dots, N$.*

Note that reflection in the line $p = q$ converts the Farey diagram for p/q into that for q/p. Therefore the Farey weights for p/q and q/p are equal.

These Farey weights are the weights considered in [16]. Our aim in this section is to show that these agree with the weights $\boldsymbol{w}(a)$ of Definition 1.2.5.

Example A.8. One can see from Figure A.1 that when $p/q = 5/3$ the labels λ_i (in decreasing order) are $\lambda_4 = 1$, $\lambda_3 = 1$, $\lambda_2 = 2$, $\lambda_1 = 3$ which gives the Farey weights $1, \frac{2}{3}, \frac{1}{3}, \frac{1}{3}$. These agree with the weight expansion constructed in Definition 1.2.5.

In the following we will denote the distinct Farey labels for $a = p/q$ by $h_1 > h_2 > \cdots > h_S > 0$ and will suppose that they occur with multiplicities $n_1, \dots, n_S$. Thus we write

$$(\lambda_1, \lambda_2, \dots, \lambda_N) = \big(h_1^{\times n_1}, \dots, h_S^{\times n_S}\big).$$

Proposition A.9. *Let $h_1 > \cdots > h_S > 0$ be the distinct Farey labels for $a = p/q$ and suppose that they occur with multiplicities $n_1, \dots, n_S$.*

(i) *If $a \in (k, k+1]$ then $n_1 = k$ and $h_1 = q, h_2 = p - kq$;*
(ii) *If $a \in [1/(k+1), 1/k)$ then $n_1 = k$ and $h_1 = p, h_2 = q - kp$;*

(iii) *In both cases the* h_i *for* $1 \le i < S$ *satisfy the recursion relation*

$$h_i = n_{i+1}h_{i+1} + h_{i+2},$$

where $h_{S+1} := 0$.

Corollary A.10. *For all* $a > 1$ *the weights* $\boldsymbol{w}(a)$ *of Definition 1.2.5 are the Farey weights of* a.

Proof. If we write $\boldsymbol{w}(a)$ as $\boldsymbol{w}(a) = \left(1^{\times \ell_1}, x_2^{\times \ell_2}, \dots, x_K^{\times \ell_K}\right)$ then Definition 1.2.5 implies that the x_i are characterized by the properties that $x_1 = 1$, $x_i > x_{i+1} \ge 0$ and the recursive relation

$$x_i = \ell_{i+1}x_{i+1} + x_{i+2}.$$

Since the λ_i are positive and nonincreasing, Proposition A.9 shows that Farey weights λ_i/λ_1 have precisely the same characterization. □

Proof of Proposition A.9. Reflection in the line $p = q$ converts the Farey diagram for p/q into that for q/p. Therefore statements (i) and (ii) are equivalent. We will prove all three statements together by an inductive argument.

We use the extended diagram obtained by adding to the edges $\epsilon_1, \dots, \epsilon_N$ the edge ϵ_0 with normal $v_0 = (0,1)$ and the edge ϵ_{-1} with normal $v_{-1} = (1,0)$. When $a > 1$ we order them as $\epsilon_{-1}, \epsilon_0, \epsilon_1, \dots,$ and then label these as in $(*)$ above; when $a < 1$ we order them as $\epsilon_0, \epsilon_{-1}, \epsilon_1, \dots,$ and then label them using $(*)$.

If $k < a < k+1$, the Farey expansion starts with $1, 2, \dots, k, k+1$ and then contains further elements between k and $k+1$. It follows that $n_1 = k$. Further, because the only edges meeting ϵ_0 are $\epsilon_1, \dots, \epsilon_{k+1}$, we have

$$\lambda_0 = k\lambda_1 + \lambda_{k+1} = n_1h_1 + h_2. \tag{A.0.3}$$

Similarly, because the only edges meeting ϵ_{-1} are ϵ_0 and ϵ_1, we have $\lambda_{-1} = \lambda_0 + \lambda_1$. Therefore (i) is equivalent to

(iv) $n_1 = k$ *and* $\lambda_0 = p,\ \ \lambda_{-1} = p + q$ *when* $p/q > 1$.

Similarly, if $a \in (1/(k+1), 1/k)$ we find that

$$n_1 = k, \quad \lambda_{-1} = k\lambda_1 + \lambda_{k+1} = n_1h_1 + h_2, \quad \lambda_0 = \lambda_{-1} + \lambda_1.$$

Hence (ii) is equivalent to

(v) $n_1 = k$ *and* $\lambda_{-1} = q,\ \ \lambda_0 = p + q$ *when* $p/q < 1$.

We argue by induction on N, the length of the Farey expansion of a. By symmetry, it suffices to consider the case when $a \in (k, k+1]$. The result is clear when $a = k+1$ (and also for the trivial case $a = 1$ which has a single label $\lambda_1 = 1$). Point (i) is easily checked when $N = k+2$ since then $p/q = (2k+1)/2$. Similarly, one can check it for the two numbers $(3k+1)/3, (3k+2)/3$ with $N = k+3$. (Note that $v_{k+2} = (2, 2k+1)$ and v_{k+3} is either $(3, 3k+1) = (1,k) \oplus (2, 2k+1)$ or $(3, 3k+2) = (2, 2k+1) \oplus (1, k+1)$.) Now, consider the matrix

$$A_k = \begin{pmatrix} k+1 & -1 \\ -k & 1 \end{pmatrix}$$

that takes the vectors $(1,k),(1,k+1)$ to $(1,0)=v_{-1},(0,1)=v_0$. Then

$$A_k\begin{pmatrix} q\\ p\end{pmatrix} = \begin{pmatrix} (k+1)q-p\\ -kq+p\end{pmatrix} =: \begin{pmatrix} q'\\ p'\end{pmatrix}.$$

Therefore if $p/q =: k+x$, we find $p'/q' = x/(1-x)$. In particular, $p'/q'>1$ if and only if $x>1-x$, that is, exactly if $n_2=1$.

Because $\det A_k = 1$, A_k preserves the Farey addition relation between adjacent normals. Hence if $v_1=(1,1),v_2,\dots,v_N$ are the normals in the diagram for p/q, the normals for the diagram for p'/q' are

$$A_kv_{k+2}=(1,1),\ A_kv_{k+3},\ \dots,\ A_kv_N.$$

In fact one could construct the diagrams for p/q and p'/q' so that there is an affine transformation obtained by following A_k by a suitable translation that takes the standard diagram for p/q to the extended diagram for p'/q'. Therefore, if $p'/q'>1$, the labels $\lambda_k,\lambda_{k+1},\dots,\lambda_N$ for p/q equal the labels $\lambda'_{-1},\lambda'_0,\lambda'_1,\dots,\lambda'_{N-k-1}$ of the extended diagram for p'/q'. Hence:

> *if $p'/q'>1$, then $n_2=1$ and the multiplicities for p'/q' are $n_3,\dots,n_S$ with corresponding labels $h_3,\dots,h_S$.*

Therefore because the recursive relation (iii) holds for p'/q' it holds for p/q and $i\ge 3$. Further, by equation (A.0.3)) and (iv) applied to p'/q', $\lambda'_0 = n_3h_3+h_4 = p'$ and $\lambda'_{-1}=\lambda'_0+\lambda'_1 = p'+h_3$. Therefore, since $n_2=1$,

$$h_2=\lambda_{k+1}=\lambda'_0=n_3h_3+h_4=p',\quad h_1=\lambda_1=\lambda'_{-1}=n_2h_2+h_3.$$

This shows that (iii) holds for p/q. Moreover, $h_2=p'=p-kq$ and, by (i) for p'/q', we find $h_1=p'+h'_1=p'+q'=q$.

This completes the proof when $p'/q'>1$. When $p'/q'<1$, the proof is similar. By (v), the labels $\lambda'_0,\lambda'_{-1},\lambda'_1,\dots$ (note the reordering) for the extended diagram for p'/q' are $\lambda_k+\lambda_{k+1},\lambda_k,\lambda_{k+2},\dots$, with first multiplicity $n'_1=n_2-1$. Further details will be left to the reader. □

Appendix B. Computer programs

B.1. **Computing $c(a)$ at a point a.** In this section we describe a `Mathematica` program `SolLess[a,D]` that finds for a rational number a and a natural number D all classes $(d;\boldsymbol{m})\in\mathcal{E}$ with $\ell(\boldsymbol{m})=\ell(a)$ and $\mu(d;\boldsymbol{m})(a)>\sqrt{a}$ and $d\le D$. We have applied this program in the proof of Theorem 5.2.3 to eight numbers $z_k = 7\frac{2}{2k+1}$ in $[7\frac{1}{9},8]$. The present program can be used for all a. By removing one line, one obtains a program finding *all* obstructive solutions at a with $d\le D$ (not just those with $\ell(\boldsymbol{m})=\ell(a)$).

Recall from Remark 5.2.6 that instead of using the code `SolLess`, one can use the algebraic method from the proof of Lemma 5.2.5 to find all obstructive classes $(d;\boldsymbol{m})$ at z_k with $\ell(\boldsymbol{m})=\ell(z_k)$. We have chosen to use this code for convenience, and because it might be helpful for understanding the more involved code of § B.2.

We start with computing the weight expansion $\boldsymbol{w}(a)$ of a rational number a. For convenience, we use that the multiplicities of $\boldsymbol{w}(a)$ are given by the continued fraction expansion of a.

```
W[a_] := Module[{aa=a,M,i=2,L,u,v},
                 M = ContinuedFraction[aa];
                 L = Table[1, {j,M[[1]]}];
                 {u,v} = {1,aa-Floor[aa]};
                 While[i <= Length[M],
                       L = Join[L, Table[ v, {j,M[[i]]}] ];
                       {u,v} = {v,u - M[[i]] v};
                       i++];
                 Return[L] ]
```

For instance, `W[3+2/3]` yields {`1,1,1,2/3,1/3,1/3`}.

We next give for each natural number k a list of 4 vectors, from which we will construct candidates for the vectors $\boldsymbol{m}$.

```
P[k_] := Module[{kk=k,PP,T0,i},
                 T0  = Table[0,{u,1,k}];
                 T0p = ReplacePart[T0,1,1];
                 T1  = Table[1,{u,1,k}];
                 T1m = ReplacePart[T1,0,-1];
                 PP = {T0,T0p,T1,T1m};
                 Return[PP] ]
```

For instance, `P[3]` yields {`0, 0, 0`}, {`1, 0, 0`}, {`1, 1, 1`}, {`1, 1, 0`} .

Our next task is to construct for given a all candidate vectors $\boldsymbol{m}$. To this end we first take a given multiplicity vector `M`, say (k_1, k_2, k_3), and associate to it all vectors of length $k_1+k_2+k_3$ such that that the j th block is a vector from `P[`k_j`]`. In the example we thereby obtain 4^3 vectors . We use the sets `P[k]` and a recursion:

```
Difference[M_] := Module[{V=M,vN,V1,l,L={},D,PP,i,j,N},
                          l = Length[V];
                          If[ l == 1, L = P[ V[[1]] ]];
                          If[ l >  1,
                              vN = V[[-1]];
                              V1 = Delete[V,-1];
                              D  = Difference[V1];
                              PP = P[vN];
                              i  = 1;
                              While[ i <= Length[D],
                                     j=1;
                                     While[j <= Length[PP],
```

```
                                        N = Join[ D[[i]], PP[[j]] ];
                                        L = Append[L,N];
                                        j++];
                                i++]
                    ];
                 Return[L] ]
```

We now take a positive rational number a and $d \in \mathbb{N}$ and compute all solutions of the Diophantine equation with d given that are obstructive at a: We first take the multiplicity vector `M` of $\boldsymbol{w}(a)$, and then round down each of its entries, getting F. In view of Lemma 2.1.7, an obstructive multiplicity vector $\boldsymbol{m}$ must be of the form `F+D[[i]]`, where `D[[i]]` is the `i`th vector from the list `Difference[W[a]]`. We therefore run through this list, and each time check whether `V=F+D[[i]]` is a solution of the Diophantine system, has last entry positive, and is obstructive: $\mu(d;V)(a) > \sqrt{a}$. If all three conditions are fulfilled, we add `V` to our list, and also retain d.

```
Sol[a_,d_] := Module[{aa=a,dd=d,M,F,D,i,V,L={}},
                M = ContinuedFraction[aa];
                F = Floor[ dd/Sqrt[aa] W[aa] ];
                D = Difference[M];
                i=1;
                While[i <= Length[D],
                      V = Sort[F+D[[i]], Greater];
                      SV = Sum[ V[[j]], {j,1,Length[V]} ];
                      If[ {SV, V.V} == {3dd-1, dd^2+1}
                           && V[[-1]] > 0
                           && W[aa].V / dd >= Sqrt[aa],
                           L = Append[L, V]
                        ];
                      i++];
                Return[{dd,Union[L]}] ]
```

For instance, `Sol[7 + 1/8, 48]` yields

$$\{48, \{\{18, 18, 18, 18, 18, 18, 18, 3, 2, 2, 2, 2, 2, 2, 2\}\}\}.$$

Remark B.1. (i) We were not at all economical when constructing the list `Difference[M]`: In view of Lemma 2.1.7, for an obstructive vector `F+D[[i]]` there is at most one k_j such that the vector `P[`k_j`]` appearing in `D[[i]]` can have both 0 and 1 as entries. We have chosen this form of the program to make it more readable.

(ii) In the main body of the paper, we applied this program only to the eight numbers $z_k = 7\frac{2}{2k+1}$, and for these numbers we know that $m_1 = m_7$ by Lemma 5.2.7. We did not use this information so as to make the program applicable also at other points, e.g. to $7\frac{1}{k}$ in order to check Lemma 5.2.5 (at least for all $d \leq 2000$ or so).

(iii) Recall from Lemma 2.1.3 that for every $(d;\boldsymbol{m})$ that gives an obstruction at a we have $\ell(a) \geq \ell(\boldsymbol{m})$. The condition `V[[-1]] > 0` asked in `Sol[a,d]` is therefore equivalent to $\ell(a) = \ell(\boldsymbol{m})$. By removing this condition, we obtain a program finding *all* obstructive solutions $(d;\boldsymbol{m})$ at a. $\diamond$

We finally collect, for given a and $D \in \mathbb{N}$, all solutions that are obstructive at a and have $d \leq D$:

```
SolLess[a_,D_] := Module[{aa=a,DD=D,d=1,Ld,L={}},
                          While[d <= D,
                                Ld = Sol[aa,d];
                                If[ Length[ Ld[[2]] ] > 0,
                                    L = Append[L,Sol[aa,d]]
                                  ];
                                d++];
                          Return[L] ]
```

B.2. **Computing $c(a)$ on an interval.** In this section we describe a `Mathematica` program `InterSolLess[k,D]` that provides for a natural number D a finite list of candidate classes $(d;\boldsymbol{m}) \in \mathcal{E}$ with $\ell(\boldsymbol{m}) = \ell(a)$ and $\mu(d;\boldsymbol{m})(a) > \sqrt{a}$ and $d \leq D$ for some $a \in]7\frac{1}{k+1}, 7\frac{1}{k}[$, $a \neq z_k$, where $z_k = [7;k,2]$. We have applied this program in the proof of Theorem 5.2.3 to the eight intervals $]7\frac{1}{k+1}, 7\frac{1}{k}[$, $k \in \{1,\dots,8\}$. Throughout we assume that a, b and the m_i are positive integers.

Our first goal is to list for a given pair a, b all solutions of the Diophantine system

$$\text{(B.2.1)} \qquad \begin{cases} a &= \sum_i m_i \\ b &= \sum_i m_i^2 \end{cases}$$

To illustrate our method, let us find in an algorithmic way the solutions $\boldsymbol{m}$ of (B.2.1) for $(a,b) = (4,6)$. It suffices to list solutions $\boldsymbol{m} = (m_1, m_2, \dots, m_M)$ with $m_1 \geq m_2 \geq \dots \geq m_M$. We must have $m_1 \leq \lfloor\sqrt{6}\rfloor = 2$. We therefore try with $m_1 = 1$ and $m_1 = 2$. For a solution $\boldsymbol{m}$, the next numbers $(m_2, m_3, \dots)$ must fulfill (B.2.1) with $(a,b) = (4-1, 6-1) = (3,5)$ and $(a,b) = (4-2, 6-4) = (2,2)$. In the first case (when $m_1 = 1$), we only need to try with $m_2 = 1$. The next numbers $(m_3, \dots)$ of a solution must then fulfill (B.2.1) with $(a,b) = (3-1, 5-1) = (2,4)$. Then $a^2 = b$, so that the only solution is $m_3 = 2$. But $m_3 = 2 > 1 = m_2$, whence we discard the solution $(1,1,2)$. In the second case (when $m_1 = 2$), we try to find numbers $(m_2, m_3, \dots)$ solving (B.2.1) with $(a,b) = (4-2, 6-4) = (2,2)$. We only need to try with $m_2 = 1$, and then want to solve (B.2.1) with $(a,b) = (2-1, 2-1) = (1,1)$ for m_3. Since $a^2 = b$, the only solution is $m_3 = 1$. We therefore find the solution $\boldsymbol{m} = (2,1,1)$.

The code `Solutions[a,b]` below does the same thing by a recursion. Note that if $a^2 < b$, then (B.2.1) has no solution.

```
Solutions[a_,b_] := Solutions[a,b,Min[a,Floor[Sqrt[b]]]]
```

```
Solutions[a_,b_,c_] := Module[{A=a,B=b,C=c,i,m,K,j,V,L={}},
                              If[ A^2 < B, L={}];
                              If[ A^2== B,
                                  If[ A > C, L={}, L={{A}} ] ];
                              If[ A^2 > B,
                                  i=1;
                                  m = Min[Floor[Sqrt[B]],C];
                                  While[i <= m,
                                        K = Solutions[A-i,B-i^2,i];
                                        j=1;
                                        While[j <= Length[K],
                                              V = Prepend[ K[[j]], i];
                                              L = Append[L,V];
                                              j++
                                             ];
                                        i++]
                                ];
                              Return[Union[L]] ]
```

Notice that applied to $(a, b) = (3d-1, d^2+1)$, the above algorithm lists all solutions of our principal Diophantine equation. For large d, however, there are many solutions. We shall therefore directly choose the first $7+k+1$ numbers m_i, using that for obstructive solutions the vectors $\boldsymbol{m}$ and $\boldsymbol{w}(a)$ must be essentially parallel, and shall then use the code `Solutions` only to choose the remaining $m_{7+k+2}, \dots$.

It will be useful to have a short expression for the sum of the entries of a vector `L`:

```
sum[L_] := Sum[ L[[j]], {j,1,Length[L]} ]
```

For given $k \geq 1$ the following code gives three vectors of length $7+k$ that have all entries equal to 0 except that the last entry of the first vector is -1 and the eighth entry of the third vector is 1.

```
P[k_] := Module[{kk=k,PP,T0,i},
                T0 = Table[0,{i,7+kk}];
                Tm = ReplacePart[T0,-1,-1];
                Tp = ReplacePart[T0,1,8];
                PP = {Tm,T0,Tp};
                Return[PP] ]
```

For $k=4$, this gives $\{0^7, 0, 0, 0, -1\}$, $\{0^7, 0, 0, 0, 0\}$, $\{0^7, 1, 0, 0, 0\}$. We shall use these vectors to take into account that the m_i may not be constant on the second block.

Fix $d \in \mathbb{N}$ and $k \in \{1, \dots, 8\}$. Assume that $\boldsymbol{m}$ is such that $(d; \boldsymbol{m}) \in \mathcal{E}$ and such that $\ell(\boldsymbol{m}) = \ell(a)$ and $\mu(d; \boldsymbol{m})(a) > \sqrt{a}$ for some $a \in \,]7\frac{1}{k+1}, 7\frac{1}{k}[$.

In view of Lemma 2.1.7 we have $m_1 = m_7$. Moreover, since $|\varepsilon_1| \le \frac{1}{\sqrt{7}} < \frac{1}{2}$ and $a < 7\frac{1}{k}$,

$$m_1 = \tfrac{d}{\sqrt{a}} + \varepsilon_1 > \tfrac{d}{\sqrt{7\frac{1}{k}}} - \tfrac{1}{2}$$

and hence $m_1 \ge \texttt{m1} := \texttt{Round}\left(\frac{d}{\sqrt{7+\frac{1}{k}}}\right)$. In the same way we see that $m_1 \le \texttt{M1} := \texttt{Round}\left(\frac{d}{\sqrt{7+\frac{1}{k+1}}}\right)$. The number $m_1 = m_7$ must therefore be in the interval $[\texttt{m1}, \texttt{M1}]$.

Next, consider m_j for $j \in \{8, \dots, 7+k\}$. For $a = 7 + x$ we have $\frac{1}{k+1} < x < \frac{1}{k}$. Therefore,

$$m_j = \tfrac{d}{\sqrt{a}}x + \varepsilon_j > \tfrac{d}{\sqrt{7\frac{1}{k}}}\tfrac{1}{k+1} - 1$$

and hence $m_j \ge \texttt{mx} := \left\lceil \frac{d}{\sqrt{7\frac{1}{k}}}\frac{1}{k+1} \right\rceil - 1 = \texttt{Ceiling}\left(\frac{d}{\sqrt{7\frac{1}{k}}}\frac{1}{k+1}\right) - 1$. In the same way we see that $m_j \le \texttt{Mx} := \left\lfloor \frac{d}{\sqrt{7\frac{1}{k+1}}}\frac{1}{k} \right\rfloor + 1 = \texttt{Floor}\left(\frac{d}{\sqrt{7\frac{1}{k+1}}}\frac{1}{k}\right) + 1$. The numbers $m_8, \dots, m_{7+k}$ must therefore be in the interval $[\texttt{mx}, \texttt{Mx}]$.

Since $a \in]7\frac{1}{k+1}, 7\frac{1}{k}[$, we have $\ell(\boldsymbol{m}) = \ell(a) \ge 7+k+1$. Lemma 2.1.8 (i) applied to the second block shows that $|m_7 - (m_8 + \cdots + m_{7+k}) - m_{7+k+1})| \le \texttt{Ceiling}\left(\sqrt{k+2}\right) - 1$. Note that we can assume that $m_1 \ge m_2 \ge \dots m_{7+k+1}$ and that $m_{7+k+1} \ge 1$.

Using this information about $m_1, \dots, m_{7+k+1}$, we build a preliminary list of vectors $\boldsymbol{m}$ as follows: We first take all possibilities for $m_1, \dots, m_{7+k+1}$ into account. Since the full vector $\boldsymbol{m}$ solves the Diophantine equation for d, the remaining numbers $m_{7+k+2}, \dots, m_M$ must solve the Diophantine equation (B.2.1) with

$$a = 3d - 1 - \sum_{i=1}^{7+k+1} m_i, \qquad b = d^2 + 1 - \sum_{i=1}^{7+k+1} m_i^2.$$

Note that we can assume that $a \ge 0$ and $b \ge 0$. We then take the list `Solutions[a,b,M[[-1]]` of all solutions to (B.2.1) for which all m_j are at most m_{7+k+1}, and append each such solution to $(m_1, \dots, m_{7+k+1})$. It could be that the only solution is 0 (namely if $a = b = 0$); in this case we remove the entry 0.

```
Prelist[k_,d_] := Module[{kk=k,dd=d,u,v,m1,M1,mx,Mx,f,t,
                               PP,M,MM,i=0,j=0,s=1,S,T,K,l,L={}},
                    u  = 1/(kk+1);
                    v  = 1/kk;
                    m1 = Round[dd/Sqrt[7+v]];
                    M1 = Round[dd/Sqrt[7+u]];
                    mx = Floor[dd/Sqrt[7+v] u]-1;
                    Mx = Ceiling[dd/Sqrt[7+u] v]+1;
                    f  = Ceiling[Sqrt[kk+2]-1];
                    t  = -f;
```

```
                  PP = P[kk];
                  While[i <= M1-m1,
                    While[j <= Mx-mx,
                      While[s <= 3,
                        While[t <= f,
                              M = Join[ Table[m1+i, {u,7}], Table[mx+j, {u,kk}] ];
                              M = M + PP[[s]];
                              S = Sum[ M[[u]], {u,8,8+kk-1}];
                              M = Append[M, M[[7]]-S+t];
                              T=1;
                              If[ M == Sort[M,Greater] && M[[-1]] > 0, T=1, T=0];
                              S = sum[M];
                              A = 3dd-1-S;
                              B = dd^2+1-M.M;
                              If[ Min[A,B] < 0, T=0];
                              If[ T==1,
                                  K = Solutions[A,B,M[[-1]]];
                                  l=1;
                                  While[l <= Length[K],
                                        MM = Join[ M,K[[l]] ];
                                        While[ MM[[-1]] == 0, MM=Drop[MM,-1] ];
                                        L = Append[L,MM];
                                        l++
                                       ]
                                ];
                        t++];
                        t=-f;
                      s++];
                      s=1;
                   j++];
                   j=0;
                  i++];
                  Return[{dd,Union[L]}] ]
```

Many of the solutions $(d;\boldsymbol{m})$ in `Prelist[k,d]` are not obstructive. The next code removes most of these solutions:

```
InterSol[k_,d_] := Module[{kk=k,dd=d,L,M,T,K={},i=1,l,rest},
                   L = Prelist[kk,dd][[2]];
                   While[i <= Length[L],
                         M = L[[i]];
                         l = Length[M];
                         T = 1;
                         If[ l <= 7 + kk + 2, T=0];
```

```
            If[ M[[-2]]-M[[-1]] > 1, T=0 ];
            If[ M[[-3]] > M[[-2]] + 1
                && Abs[ M[[-3]]-M[[-2]]-M[[-1]] ] > 1, T=0 ];
            If[ kk==1 && l >= 10,
                If[ M[[9]] - M[[10]] > 1 &&
                    Abs[ M[[8]] - (M[[9]] + M[[10]]) ] > 1,
                    T=0 ]];
            rest = Sum[ M[[j]], {j,8+kk,l} ];
            If[ M[[7+kk]] - rest >= Sqrt[l-kk-6], T=0 ];
            If[ T==1, K = Append[K, M] ];
            i++];
    Return[{dd,K}] ]
```

Recall that $a \in \,]7\frac{1}{k+1}, 7\frac{1}{k}[$, $a \neq z_k$, where $z_k = [7; k, 2]$. In particular, $\ell(\boldsymbol{m}) = \ell(a) > 7 + k + 2$. We then test the very end of $\boldsymbol{m}$: Since the last block has length at least 2, we must have $m_{M-1} = m_M + 1$ or $m_{M-1} = m_M$ in view of Lemma 2.1.7.

We next exploit Lemma 2.1.8: If $m_{M-2} > m_{M-1} + 1$, then we know that the length of the last block is 2, and m_{M-2} belongs to the before the last block. Lemma 2.1.8 (ii) then shows that $|m_{M-2} - (m_{M-1} + m_M)| \leq 1$. In the next test we apply the same lemma to the special situation where $k = 1$ and where we know that the third block has length 1.

In the last test we apply Lemma 2.1.8 (ii) with $j = 7 + k$.

We finally take the union over $d \leq D$ of the solutions in InterSol[k,D]:

```
InterSolLess[k_,D_] := Module[{kk=k,DD=D,LL={},Q,d=1},
                            While[d <= DD,
                                  Q = InterSol[kk,d];
                                  If[Length[Q[[2]]] > 0,
                                     LL = Append[LL,Q]];
                                  d++];
                            Return[LL] ]
```

References

[1] P. Biran, Symplectic packing in dimension 4, *Geom. Funct. Anal.* **7** (1997) 420–437.

[2] P. Biran, Constructing new ample divisors out of old ones, *Duke Math. J.* **98** (1999) 113–135.

[3] P. Biran, From Symplectic Packing to Algebraic Geometry and back, European Congress of Mathematics, Vol. II (Barcelona, 2000), 507–524, *Progr. Math.* **202**, Birkhäuser, Basel, 2001.

[4] O. Buse and R. Hind, Symplectic embeddings of ellipsoids in dimension greater than four, in preparation.

[5] K. Cieliebak, H. Hofer, J. Latschev and F. Schlenk, Quantitative symplectic geometry, Dynamics, ergodic theory, and geometry, 1–44, *Math. Sci. Res. Inst. Publ.* **54**, Cambridge Univ. Press, Cambridge, 2007.

[6] I. Ekeland and H. Hofer, Symplectic topology and Hamiltonian dynamics. II *Math. Z.* **203** (1990) 553–567.

[7] L. Guth, Symplectic embeddings of polydisks, *Invent. Math.* **172** (2008) 477–489.

[8] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, OUP, Oxford (1938).
[9] R. Hind and E. Kerman, New obstructions to symplectic embeddings, arXiv:0906.4296.
[10] M. Hutchings, Quantitative embedded contact homology, arXiv:1005.2260, to appear in *J. Differential Geom.*
[11] M. Hutchings, Recent progress on symplectic embedding problems in four dimensions, arXiv:1101.1069, to appear in *PNAS.*
[12] M. Hutchings and C. H. Taubes, Gluing pseudoholomorphic curves along branched covered cylinders, I, *J. Symplectic Geom.* **5** (2007) 43–137.
[13] Bang-He Li and T.-J. Li, Symplectic genus, minimal genus and diffeomorphisms, *Asian J. Math.* **6** (2002) 123-144.
[14] T.-J. Li and A. K. Liu, Uniqueness of symplectic canonical class, surface cone and symplectic cone of 4-manifolds with $b^+ = 1$, *J. Differential. Geom.* **58** (2001) 331–370.
[15] D. McDuff, From symplectic deformation to isotopy, Topics in symplectic 4-manifolds (Irvine, CA, 1996) 85–99, *First Int. Press Lect. Ser., I*, Int. Press, Cambridge, MA, 1998.
[16] D. McDuff, Symplectic embeddings of 4-dimensional ellipsoids, *J. Topol.* **2** (2009) 1–22.
[17] D. McDuff, Symplectic embeddings and continued fractions: a survey, *Jpn. J. Math.* **4** (2009) 121–139.
[18] D. McDuff, The Hofer conjecture on embedding symplectic ellipsoids, arXiv:1008.1885, to appear in *J. Differential Geom.*
[19] D. McDuff and L. Polterovich, Symplectic packings and algebraic geometry, *Invent. Math.* **115** (1994) 405–29.
[20] D. Müller, Symplectic embeddings of ellipsoids into polydiscs, PhD thesis, Université de Neuchâtel, in preparation.
[21] E. Opshtein, Maximal symplectic packings in $\mathbb{P}^2$, *Compos. Math.* **143** (2007) 1558–1575.
[22] E. Opshtein, Singular polarizations and symplectic embeddings, arXiv:1011.6358.
[23] P. Seidel, Lectures on four-dimensional Dehn twists, Symplectic 4-manifolds and algebraic surfaces, 231–267, *Lecture Notes in Math.* **1938**, Springer, Berlin, 2008.
[24] J. J. Sylvester, On a remarkable modification of Sturm's theorem, *Philosophical Magazine* **V** (1853) 446-456.
[25] C. H. Taubes, Embedded contact homology and Seiberg-Witten Floer cohomology I. *Geom. Topol.* **14** (2010) 2497–2581.

(D. McDuff) Department of Mathematics, Barnard College, Columbia University, New York, NY 10027-6598, USA.
E-mail address: dmcduff@barnard.edu

(F. Schlenk) Institut de Mathématiques, Université de Neuchâtel, Rue Émile Argand 11, CP 158, 2009 Neuchâtel, Switzerland
E-mail address: schlenk@unine.ch