The Hurwitz Complex Continued Fraction

Doug Hensley

January 9, 2006

Abstract

The Hurwitz complex continued fraction algorithm generates Gaussian rational approximations to an arbitrary complex number α by way of a sequence $(a_0, a_1, \ldots)$ of Gaussian integers determined by $a_0 = [\alpha]$, $z_0 = \alpha - a_0$, (where $[u]$ denotes the Gaussian integer nearest u) and for $j \geq 1$, $a_j = [1/z_{j-1}]$, $z_j = 1/z_{j-1} - a_j$. The rational approximations are the finite continued fractions $[a_0; a_1, \ldots, a_r]$. We establish a result for the Hurwitz algorithm analogous to the Gauss-Kuz'min theorem, and we investigate a class of algebraic α of degree 4 for which the behavior of the resulting sequences $\langle a_j \rangle$ and $\langle z_j \rangle$ is quite different from that which prevails almost everywhere.

1 Complex Continued Fractions

A *complex continued fraction algorithm* is, broadly speaking, an algorithm that provides approximations by ratios of Gaussian integers, or integers of another complex number field, to a given complex number. We are faced with an immediate choice: does speed matter, or does approximation quality trump everything else?

If quality of approximation is paramount, then the algorithm of choice is due to Asmus Schmidt[7]. The basic idea is to partition (we ignore issues concerning breaking ties at the boundaries) the complex plane into pieces bounded by line segments and arcs of circles, and to successively refine this partition. The partition component to which the target ξ belongs shrinks down to ξ as this process is iterated.

Each component of the partition is associated with a linear fractional map $z \to (az + b)/(cz + d)$, where the coefficients a, b, c, d are Gaussian integers, as well as with the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

The Schmidt algorithm resembles the Lagarias multidimensional continued fraction algorithm [5] in that both generate, for each ξ, a sequence $\langle M_n(\xi) \rangle$ of matrices with 'integer' entries, whose inverses also have 'integer' entries. The required Diophantine approximation to the target ξ can then be read off from $M_n(\xi)$.

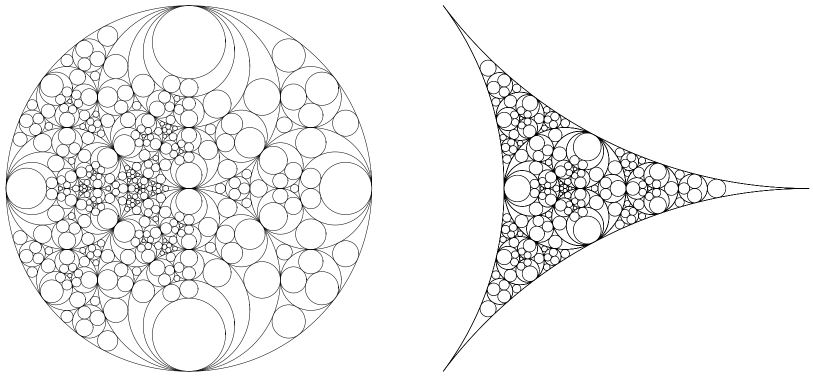

Figure 1: Fourth generation and (zoom to left-center) fifth.

It produces a tesselation of its space of possible targets, with the property that each set of ξ associated to a given M_n splits into some number of shards, each associated with a specific one of the possible successors M_nT. We show the fourth-generation partition of that portion of $\mathbb{C}$ bounded by the circle with diameter $(1, 1+i)$:

On average, the Schmidt algorithm would appear to require only polynomial time in the binary length of the numerator and denominator of Gaussian-rational inputs. The worst-case is a different matter. Given a real input, or one equivalent at an early stage to a real number, the algorithm can be quite slow.

2 The Hurwitz Complex Continued Fraction

If we relax the priority we set on quality of approximation, what do we get in exchange? Any complex continued fraction algorithm should take as input a complex number $z = z_0$, and return as output a sequence of approximations (p_n/q_n) to z, where $p_n, q_n \in G = \{a + ib\colon\ a \in \mathbb{Z}, b \in \mathbb{Z}\}$ are relatively prime Gaussian integers. The following properties are desirable:

1. The sequence $(|q_n|)$ increases at least exponentially.

2. The sequence (p_n/q_n) terminates in p/q if the input z is a Gaussian rational.

3. All approximations are of quality at least comparable to that which is known to exist by virtue of the Dirichlet pigeonhole principle:

$$\left|\frac{p_n}{q_n} - z\right| \ll |q_n|^{-2}$$

uniformly over n and over all inputs z.

It would be nice if the algorithm were also easily executed, provided that basic arithmetic operations and truncations present no particular difficulty. (If the input is given only as the limit of a sequence of decimal approximations, for instance, it may be impossible to resolve a truncation operation.) We might hope that the p_n/q_n take the form

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}} \quad .$$

We might also hope that every reasonably good Gaussian rational approximation p/q to z, of quality comparable to or better than that vouchsafed by the Dirichlet pigeonhole principle, was either a continued fraction convergent, or simply related to the two convergents whose denominators bracketed q in absolute value.

Finally, we might hope to understand both what typically happens for a 'random' initial z, and what happens in selected special cases such as inputs that satisfy a quadratic polynomial with Gaussian integer coefficients.

The classical algorithm due to A. Hurwitz [2] meets most of these desiderata. It takes as input a complex number z, and outputs a sequence $\langle a_n \rangle$ of Gaussian integers, from which further sequences $\langle p_n \rangle$ and $\langle q_n \rangle$ of Gaussian integers can be computed exactly as in the case of the classical continued fraction. The Gaussian rationals p_n/q_n then come to us in reduced form, and they furnish decent approximations to z. The arithmetic needed to decide on the next step is decidedly simpler than with the Schmidt algorithm, while the approximations are comparable if not always quite as good.

We denote by $[z]$ the Gaussian integer *nearest* z, rounding down, in both the real and imaginary components, to break ties. The Hurwitz complex continued fraction algorithm, like the near-classical *centered* continued fraction algorithm for real numbers, is simplest when restricted to inputs inside the fundamental domain of the truncation function. Here, that domain is $B = \{x + iy \mid -1/2 \leq x < 1/2, \ -1/2 \leq y < 1/2\}$. The algorithm proceeds by steps of the form $z_{n+1} = 1/z_n - [1/z_n]$. If $z = z_0 \in \mathbb{Q}(i)$, the algorithm terminates when, as must eventually occur, $z_n = 0$, and the final finite-depth continued fraction gives a reduced fraction p_n/q_n equal to z. The next-to-last convergent pair gives a solution to $uq_n - vp_n = 1$.

If initially $z \notin \mathbb{Q}(i)$ then the algorithm continues indefinitely, or in practice, until some other terminating condition is met.

Once $(a_0, a_1, \ldots a_n)$ have been computed, we compute the *convergents* p_k/q_k using the usual formula. (Details are provided in the next section.) As in the case of the classical algorithm,

$$\begin{vmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{vmatrix} = (-1)^n.$$

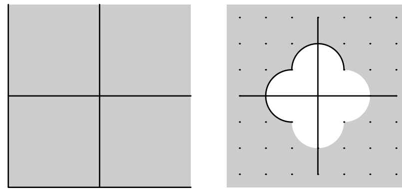

Figure 2: B and $1/B$.

Here, we show that this algorithm has excellent convergence properties, and we study some surprising cases in which algebraic inputs z lead to expansions that exhibit behavior that is neither periodic (along the lines of the classical continued fraction expansion of $\sqrt{n}$), nor 'typical'. We close by working out the details of what that 'typical' behavior is, and give an analog, for the case of the Hurwitz algorithm, to the classical Gauss-Kuz'min theorem, itself discussed in Chapter 9.

3 Notation

It will be convenient to introduce some additional notation. We assume that $z = z_0$ is given, with $z_0 \in B$, and that $p_{-1} = q_0 = 1$ while $p_0 = q_{-1} = 0$. For $n \geq 1$ let $a_n = [1/z_{n-1}]$, $p_n = a_n p_{n-1} + p_{n-2}$, and $q_n = a_n q_{n-1} + q_{n-2}$. Let $x_n = 1/z_{n-1}$, and let $w_n = q_{n-1}/q_n$. Then

$$z_{n+1} = \frac{1}{z_n} - a_{n+1}$$
$$w_{n+1} = \frac{1}{a_{n+1} + w_n}.$$

Let $1/B$ denote the set of reciprocals of the nonzero elements of B; $1/B$ is bounded by arcs of circles of radius 1 about ± 1 and $\pm i$, and these arcs pass through $\pm 1 \pm i$ and through ± 2 and $\pm 2i$. Let G' denote $G \backslash \{0, \pm 1, \pm i\}$. From the figure, it is evident that only elements of G' can occur as *partial quotients* a_n. On the other hand, it is not so simple to say which sequences can actually occur. There is a variant on this algorithm, due to Julius Hurwitz[3] in which the a_n are restricted further to be multiples of $(1 + i)$, and in this variant, it was apparently possible to ferret out the details of which sequences of (a_n) can occur.

For the algorithm now under investigation it is possible, at any rate, to work out that certain finite combinations of consecutive a_k cannot occur. The following table gives the set of all possible pairs (a, b) of consecutive a_j. The reason for these constraints is a matter of geometry, and we explain two of the cases in some detail. If $a_j = 1+i$, then x_j belongs to the intersection of $1/B$ and the square $1/2 \leq \Re x < 3/2$, $1/2 \leq \Re y < 3/2$, so that $z_j = x_j - (1 + i)$ belongs

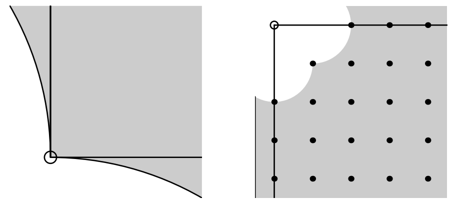

Figure 3: $a_j = 1 + i$: Possible z_j, Possible a_{j+1}.

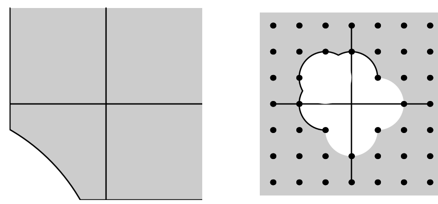

Figure 4: $a_j = 2 + i$: Possible z_j, Possible a_{j+1}.

to the intersection of $1/B - (1 + i)$ and B. The set of reciprocals of complex numbers in this set, in turn, is the intersection of the regions $\Re[z] \geq -1/2$, $\Im[z] \leq 1/2$, $|z + 1| > 1$, and $|z + i| > 1$.

If $a_j = 2 + i$, then z_j belongs to that part of B lying on or outside the circle of radius 1 about $-1 - i$, so $1/z_j$ lies in that part of $1/B$ lying on or outside the circle of radius 1 about $-1 + i$, with the result that all of G' apart from $-1 + i$ is a possibility for a_{j+1}.

The dots indicate Gaussian integers that are available for a_{j+1}. For other values of a_j, we have the following little table; negating or taking the conjugate of a_j negates or conjugates the set of possible a_{j+1}, while if $|a_j| \geq 2\sqrt{2}$, any element of G' is a possibility for a_{j+1}.

a_j	possible a_{j+1}
$1 + i$	$x - iy, x + y \geq 2$
$1 - i$	$x + iy, x + y \geq 2$
2	$x + iy, x \geq 0, \|x\| + \|y\| \geq 2$
$2 + i$	$G' \backslash \{-1 + i\}$
$1 + 2i$	$G' \backslash \{-1 + i\}$

4 Growth of $|q_n|$ and the Quality of the Hurwitz Approximations

Hurwitz continued fraction convergents to a target z_0 are good approximations. With the notation established above, we have

$$z_0 - \frac{p_n}{q_n} = \frac{(-1)^n z_n}{q_n^2(1+z_n w_n)}.$$

Since $|z_n| < 1/\sqrt{2}$, and $|w_n| < 1$, it follows immediately that $|z_0 - p_n/q_n| < 2\sqrt{2}|z_n|/|q_n|^2$. When z_0 is a Gaussian quadratic irrational, such as $\sqrt{2+i}-1$ or, for that matter, $\sqrt{2}-1$, the Hurwitz continued fraction expansion is (eventually) periodic, and $|z_n|$ is bounded below by a positive constant. On the other hand, a simple computation with (algebraic) conjugates shows that when the target z_0 is a Gaussian quadratic irrational, there exists $\epsilon > 0$, depending on z_0, such that for all Gaussian integers p, q with $q \neq 0$, $|z_0 - p/q| > \epsilon |q|^{-2}$. Thus for such targets, the Hurwitz continued fraction gives approximations that are essentially best possible. Whenever z_n is small, we get particularly good approximations, but this alone does not exclude the possibility that there might be particularly good approximations that the Hurwitz algorithm misses. The algorithm does, however, achieve approximations that are essentially best possible, in the following sense:

Theorem 1. *Suppose $\alpha \in \mathbb{C}$ has a Hurwitz algorithm sequence of convergents to depth at least n, and suppose (p_{n-1}, q_{n-1}) and (p_n, q_n) are the numerators and denominators of the $n-1$th and nth convergents. Suppose $q \in G$ with $|q_{n-1}| < |q| \leq |q_n|$, $p \in G$, and $p/q \neq p_n/q_n$. Then*

$$\left|\frac{p}{q} - \alpha\right| \geq \frac{1}{5}\left|\frac{p_n}{q_n} - \alpha\right| \cdot \left|\frac{q_n}{q}\right|.$$

We defer the proof as we shall need some of the machinery developed below.

We now discuss the growth of the continuants, that is, the denominators q_n. The first result along these lines is due to Hurwitz, who showed that for all initial $z_0 \in B$, and all n such that the expansion has not yet terminated, $|q_n| > |q_{n-1}|$. His proof is based on the observations given above.

To this, we now add

Theorem 2. *If $z \in B$ has a Hurwitz continued fraction expansion to depth $n+2$, then*

$$\left|\frac{q_{n+2}}{q_n}\right| \geq 3/2.$$

Proof. Recall that

$$\begin{aligned} q_{n+1} &= a_{n+1}q_n + q_{n-1} \\ z_{n+1} &= \frac{1}{z_n} - a_{n+1} \\ w_{n+1} &= \frac{1}{a_{n+1} + w_n} \\ a_{n+1} &= \left[\frac{1}{z_n}\right] \\ \frac{q_n}{q_{n+1}} &= w_{n+1}. \end{aligned}$$

Let G_a denote the set of possible values in G' for the successor a_{j+1} to a_j if $a_j = a$. We shall need some estimates for w_n.

Lemma 1. *Suppose $z \in B$, $n \geq 1$, and z has a Hurwitz continued fraction to depth $n+2$. If $|w_n| \geq 2/3$, then $|w_{n+1}| < 2/3$. Furthermore, either $|w_n| < 2/3$, or $\frac{2}{3} \leq |w_n| < 1$ and one of the following, or its negative or complex-conjugate counterpart, holds:*

$$\left|w_n - \frac{9}{14}(1-i)\right| < \frac{3}{7} \text{ and } a_n = 1+i$$

$$\left|w_n - \frac{9}{16}\right| < \frac{3}{16} \text{ and } a_n = 2.$$

The proof of the lemma is inductive and begins with the observation that the claim is true at the outset because $w_0 = 0$. We use a fact about reciprocals of disks: If D is a disk in $\mathbb{C}$ with center z and radius r, and if $|z| > r$, then with the notation $D(s,r) := \{z \in \mathbb{C} : |z - s| < r\}$,

$$\frac{1}{D(s,r)} = D\left(\frac{\overline{s}}{|s|^2 - r^2}, \frac{r}{|s|^2 - r^2}\right).$$

Now let D_0 be the disk about 0 with radius 2/3, and let $D_1 = 1/(1 + i + D_0) = D((9/14)(1-i), 3/7)$, $D_2 = 1/(-1 + i + D_0)$, $D_3 = -1/(1 + i + D_0)$, $D_4 = 1/(1-i+D_0)$, $D_5 = 1/(2+D_0)$, $D_6 = 1/(2i+D_0)$, $D_7 = -1/(2+D_0)$, and $D_8 = 1/(-2i + D_0)$. Assuming the claim to be true for $k \leq n$, we have either that $|w_n| \in D_0$, or that w_n lies in the intersection of one of 8 other disks with the open unit disk, and a_n takes a corresponding specific value, either $\pm 1 \pm i$ or ± 2 or $\pm 2i$.

Now if $|a| \geq \sqrt{5}$, then $1/(a + D_0) \subseteq D_0$, while if a is one of the 8 Gaussian integers in G' nearest the origin, $1/(a + D_0)$ is one of $D_1, \ldots, D_8$. Thus the inductive step cannot fail in the case that $w_n \in D_0$. On the other hand, for $1 \leq k \leq 8$, one readily checks that for any eligible successor a to a_n, $1/(D_k + a) \subseteq D_0$. Thus, for instance, when $w_n \in D_1 \backslash D_0$, we have $a_n = 1 + i$ so that $a_{n+1} \in G_{1+i} = \{u - iv : u \geq 0, v \geq 0, u + v \geq 2\}$. For $a \in G_{1+i}$, though, if $|a| \geq 2\sqrt{2}$ and $|w| < 1$ then $|1/(a + w)| < 2/3$, while if $a \in \{2, 2-i, 1-i, 1-2i, -2i\}$ then, from case by case calculation, $1/(a+D_1) \subset D_0$. For instance, $1/(1-i+D_1) = D((23/73)(1+i), 6/73)$, while $1/(2-i+D_1) = D(37/133+23i/133, 6/133)$. Thus the inductive step cannot fail in any of these other cases either, as $|w_{n+1}| < 2/3$. We now note that we have shown that whenever $|w_n| \geq 2/3$, $|w_{n+1}| < 2/3$.

At this point the proof of the theorem falls right out: $|w_n||w_{n+1}| < 2/3$ because both factors are less than 1 and one of them is less than 2/3, and so $|q_{n+2}/q_n| = 1/|w_n w_{n+1}| > 3/2$. □

We now give the proof of theorem 1. We first write $(p,q) = s(p_{n-1}, q_{n-1}) + t(p_n, q_n)$. If $s = 0$ the estimate is immediate. We now break the question into two main cases: $|s| = 1$ and $|s| > 1$.

If $|s| = 1$, then we may multiply s and t by the same unit and so take $s = 1$. Now $|\alpha - p_n/q_n| = |z_n|/(|q_n(q_n + z_n q_{n-1})|)$. Thus we must show that if $q \neq q_n$ then

$$\left| \frac{p_n + z_n p_{n-1}}{q_n + z_n q_{n-1}} - \frac{tp_n + p_{n-1}}{tq_n + q_{n-1}} \right| > \frac{1}{5} \frac{|z_n|}{|q(q_n + z_n q_{n-1})|},$$

or equivalently, $|t - 1/z_n| > 1/5$. Now $q = q_{n+1}$ if and only if $t = [1/z_n]$. But $|q| \leq |q_n| < |q_{n+1}|$, so $t \neq [1/z_n]$, so $|t - 1/z_n| \geq 1/2$. This completes the proof for the case $|s| = 1$.

We break the case $|s| > 1$ down into subcases depending on the value of a_n. There is sufficient symmetry that rotations and reflections are of no consequence, so these cases reduce to the following: $|a_n| \geq 3$, $a_n = 2+2i$, $a_n = 2+i$, $a_n = 2$, and $a_n = 1+i$. We now assume that p/q gives a counterexample at depth n. The premise $|q_{n-1}| < |q| \leq |q_n|$ gives $|w_n| < |sw_n + t| \leq 1$, while our assumption that p/q is a good approximation to α gives $|s/z_n - t| \leq (1/5)$. Equivalently,

$$\left| \frac{1}{z_n} - \frac{t}{s} \right| \leq \frac{1}{5|s|}, \quad \left| w_n + \frac{t}{s} \right| \leq \frac{1}{|s|}.$$

From these it follows that $|w_n + 1/z_n| < 6/5|s|)$.

Now the value of a_n constrains both w_n and z_n. It constrains w_n because $w_n = 1/(a_n + w_{n-1})$ and $|w_{n-1}| < 1$. Thus, w_n belongs to the disk $D(\overline{a_n}/(|a_n|^2 - 1), 1/(|a_n|^2 - 1))$ where $D(c,r)$ denotes the disk $\{z \in \mathbb{C} : |z - c| < r\}$. The effect of a_n upon the possible values of z_n comes from the fact that $z_n \in D(a_n, 1) \cap B$, which becomes important only when $|a_n| \leq \sqrt{5}$.

For the case $|a_n| \geq 3$, we have $|w_n| < 1/2$ while $|1/z_n| \geq \sqrt{2}$, so $|w_n + 1/z_n| > \sqrt{2} - 1/2 > 6/5\sqrt{2}$. If $a_n = 2+2i$ then $w_n \in D((2-2i)/7, 1/7)$. Thus it will suffice to complete this case, that $D((-2+2i)/7, (1/7 + 6/(5\sqrt{2}))$ sits inside the union of the disks about ± 1 and $\pm i$ of radius 1. It does, since $|-(1+i) + (2-2i)/7| = 5\sqrt{2}/7 > 6/(5\sqrt{2}) + 1/7$.

If $a_n = 2+i$ then $z_n \subseteq B \backslash D(-1-i, 1)$, so $1/z_n$ lies outside the union of the disks of radius 1 about ± 1, $\pm i$, and $-1+i$. On the other hand, $w_n \in D((2-i)/4, 1/4))$. The disk about $-(2-i)/4$ of radius $(1/4 + 6/(5\sqrt{2}))$ fits comfortably within the region from which $1/z_n$ is excluded, which completes the proof for the case $a_n = 2+i$.

If $a_n = 2$, then $w_n \in D(2/3, 1/3)$ while $z_n \in B \backslash D(-1, 1)$ so that $1/z_n$ belongs outside the union of our disks and the half-space $\Re z < -1/2$. The disk

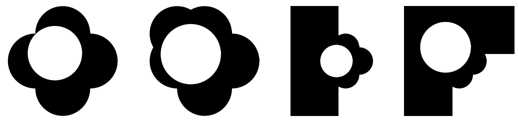

Figure 5: Excluded values of $1/z_n$ and disks for $-w_n$.

$D(-2/3, 1/3 + 6/(5\sqrt{2}))$ fits comfortably inside the region from which $1/z_n$ is excluded, which completes the proof for the case $a_n = 2$.

Finally, if $a_n = 1 + i$, then $w_n \in D(1 - i, 1)$ while $z_n \in B \backslash (D(-1, 1) \cup D(-i, 1))$. Thus $1/z_n$ is excluded from the union of our four unit disks about ± 1 and $\pm i$, and the half-planes $\Re z < -1/2$ and $\Im z > 1/2$. But $D(-1 + i, 1 + 6/(5\sqrt{2}))$ fits comfortably within this excluded zone.

5 Distribution of the Remainders

The sequence $\langle z_n \rangle$ of remainders arising out of execution of the Hurwitz continued fraction algorithm on input z_0 depends, of course, on the input z_0. If z_0 is rational, then it terminates. If z_0 is a Gaussian quadratic irrational, that is, if z_0 satisfies a quadratic polynomial over $\mathbb{Q}(i)$, then it is ultimately periodic. [2]. These things are no surprise in view of what is known about the classical continued fraction algorithm.

In the classical theory of continued fractions, we have the famous Gauss-Kuz'min theorem to the effect that if X is taken at random with uniform distribution in $[0, 1]$, then the probability density function of $T^k X$ converges exponentially fast to the Gauss density $\frac{1}{\log 2} \frac{1}{1+x}$ on $[0, 1]$. This density is characterized by the fact that it is absolutely continuous with respect to Lebesgue measure and is invariant under the map $x \to 1/x - \lfloor 1/x \rfloor$. Numerical experiments suggest that something in the same vein might be true for the Hurwitz complex continued fraction. We do obtain such a theorem, but for the moment, we just show a picture of the density that plays the rôle for the Hurwitz continued fraction that the Gauss density plays for the classical real continued fraction.

The invariant density ρ for the Hurwitz algorithm is, in its own way, rather pretty: we show here a false-color image of a somewhat crudely computed approximation to it. A simple expression for this invariant density is not known, but the topic is new and there may well be one. The figure suggests, and we later prove, that this density is real-analytic on the interiors of the 12 regions into which arcs of circles of radius 1 centered at ± 1, $\pm i$, and $\pm 1 \pm i$ cut B, and that it has the symmetry group of the rotations and reflections that map $\overline{B}$ to itself. A key fact is that T maps the union of these arcs, and the boundary of B, into that same union.

Knowing this density, we can determine the relative frequency of the partial quotients a_j. For instance, the limiting frequency with which $a_j = (1 + i)$, as

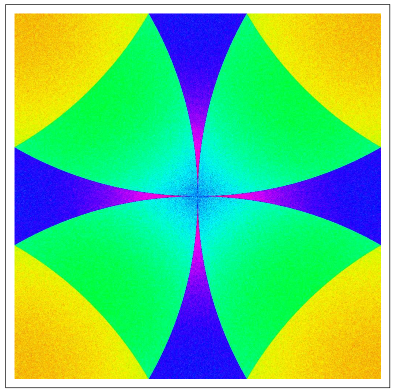

Figure 6: Invariant density for the Hurwitz algorithm.

$j \to \infty$, is $\int_{E[1+i]} \rho$, where $E[a] = \{z \in B : [1/z] = a\}$, so that $E[1+i]$ is the region bounded by the line segments $1/2+it$, $-1/2 \leq t \leq -\sqrt{3}/6$, $t-i/2$, $\sqrt{3}/6 \leq t \leq 1/2$, and by arcs of the circles $|z-1/3| = 1/3$ and $|z+i/3| = 1/3$ joining $(1-i)/3$ and $(1-i)/2$. We can also make a well-founded conjecture as to the average speed with which the algorithm brings a random Gaussian-rational input to zero.

As in the case of the rational integers and the classical gcd algorithm, a continued fraction algorithm executed on rational inputs is, en passant, computing the gcd of the input pair, as well as auxiliary information. Consider an input $z_0 = s_0/t_0$. As execution of the Hurwitz algorithm proceeds, successive $z_n = s_n/t_n$ are computed according to the rule $z_{n+1} = 1/z_n - [1/z_n] = 1/z_n - a_{n+1}$. From this, we have $s_{n+1}t_{n+1} = s_n t_n(1 - a_{n+1}z_n) = s_n t_n z_{n+1} z_n$. We conjecture that for typical random rational inputs drawn from $|s_0| < R$, $|t_0| < R$ say, the distribution of (z_n) more or less conforms to ρ, except at the beginning, where it will be more uniform, or near the end when it will of necessity be concentrated on a few simple Gaussian rationals. From this it would follow that the average number of steps needed to execute the Gaussian gcd algorithm should be proportional to $\log |t_n|$, with the constant of proportionality being $1/\int_B \log |z| \rho(z)\, dm(z)$. This integral was estimated by Monte Carlo methods to be 1.092766. Numerical tests of the conjecture on 100 pseudorandom pairs with s, t chosen at random among Gaussian integers with real and imaginary parts of absolute value less than 10^{1000} gave an aggregate of 210863 steps to finish all 100 inputs, as against a predicted value of 210679.

It would be nice to know the asymptotic behavior of the growth of q_n. For this, it would seem at first that we would need the distribution of w_n; if we had a function ω along the lines of ρ for the distribution of the successive w_n, $n\int_D(-\log |z| \omega(z))$ would be the expected value of $\log |q_n|$. This hypothetical ω can be guessed at by amassing some statistics; here is a pictures of the result. The range of the picture is the square $|\Re(z)| \leq 1$, $|\Im(z)| \leq 1$. On grounds of

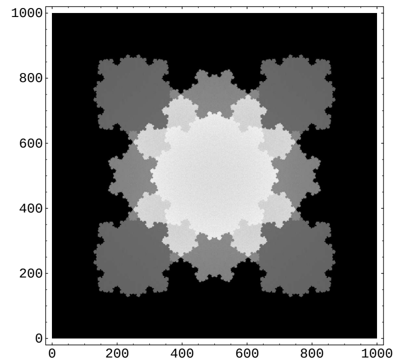

Figure 7: grayscale representation of ω.

Hurwitz's result, we know in advance that the picture sits within the unit circle tangent to the four edges of this frame. What is not clear in advance is that the set in which ω is positive is a fractal, and that within this domain, there are interior fractal boundaries delimiting regions of greater and lesser intensity.

On the other hand, we can turn around the speculations about the rate at which the Hurwitz algorithm deals with random rational inputs. Suppose the algorithm is given s_0/t_0, and requires n steps to reduce this to zero. As the algorithm is executing it is at the same time building a fraction p_n/q_n equal to the input. In general, $\log|q_n|$ will be not much less than $|\log t_n|$ because it is rare that s_0 and t_0 will have a large common factor. Thus, we should expect the rate at which t_0 is eroded to zero to be equal to the rate at which q_k builds up, and if this rate is the same for typical random rational inputs as it is for the average complex input (taken at random from B with a uniform distribution), then we should expect that almost always, $\log|q_n| \approx 1.092766n$. A simple numerical experiment is not quite trivial, as it is not feasible to carry a computation to an appropriate depth with a pseudorandom number. What we can do is to take advantage of the rough and ready randomizing effect of rounding to a fixed number of digits as we go, which has the side benefit of restoring 'precision' to floating-point numbers that lose precision during extracting their Hurwitz expansion. The exact value one arrives at will depend on the details of the truncation and precision rules employed; the author got a rate of 1.09302 during 10^6 steps, on one run, and 1.09299 on another.

We defer the proof of the existence of this invariant density ρ for the moment, and take up a rather different topic.

6 A Class of Algebraic Approximants with Atypical Hurwitz Continued Fraction Expansions

Algebraic numbers may be expected to perhaps have atypical diophantine approximation properties. When we are dealing with real numbers, the situation is not all that well understood but certain things are well-known. Rational numbers have terminating continued fraction expansions. Real roots α of an irreducible quadratic polynomial over $\mathbb{Q}$ have ultimately periodic continued fraction expansions. As a result of this, the sequence of remainders is ultimately periodic, while the sequence of scaled errors $q_n^2(\alpha - p_n/q_n)$ converges to a periodic loop. For real α of degree greater than 2, there are results concerning simultaneous diophantine approximation of the powers of α, but concerning α alone, all we know is that the continued fraction expansion cannot involve extraordinarily large integers, such as would be typical of Liouville numbers. Thus, the theoretical frequency of $1, 2, \ldots, 10$ according to the Gauss density is $(0.415037, 0.169925,$ 0.0931094, 0.0588937, 0.040642, 0.0297473, 0.0227201, 0.0179219, 0.0144996, 0.0119726), while the actual frequency of these same partial quotients in the expansion of $2^{1/3}$ to a depth of 10000 is (0.4173, 0.1675, 0.0946, 0.0636, 0.0421, 0.0295, 0.024, 0.0163, 0.0122, 0.0118).

In the case of the Hurwitz algorithm, at first it seems that we are in for more of the same. When the approximation target is in $\mathbb{Q}(i)$, the algorithm terminates, having found a reduced Gaussian rational equal to the original target. When the approximation target satisfies an irreducible quadratic polynomial over $\mathbb{Q}(i)$, the Hurwitz expansion is ultimately periodic. [2]. But if we look a bit further, we come upon a new behavior. Somewhat typical here is what happens with $z_0 = \sqrt{2} - 1 + i(\sqrt{5} - 2)$. This z_0 satisfies an irreducible quartic polynomial over $\mathbb{Q}(i)$, to wit, $x^4 + (4+8i)x^3 - (12-24i)x^2 - (32-16i)x + 24$. But the Hurwitz continued fraction expansion of z_0 is neither ultimately periodic, nor anything like the typical expansion. Instead, what we see is that

1. The sequence of remainders is confined to a certain finite set of arcs of a circle. None of these arcs goes through the origin, and so the sequence of remainders is bounded away from the origin.

2. The scaled errors (which we now define to be $|q_n|^2(z_0 - p_n/q_n)$), all lie very near one or another of a set of six parallel line segments cutting through B, equally spaced except that the one that would have gone through the origin is missing. The lines have equations of the form $\sqrt{2}x + \sqrt{5}y = n/2$ where n is an integer, $-3 \le n \le 3$.

3. The ratios of successive q_n are confined as well. They are not distributed like the fractal we saw previously. Instead, they appear to lie very near one or another of a set of circular arcs, as with the remainders.

For most of these observations, we can provide reasons.

Theorem 3. *For arbitrary positive integers u, v such that u, v, and uv are not*

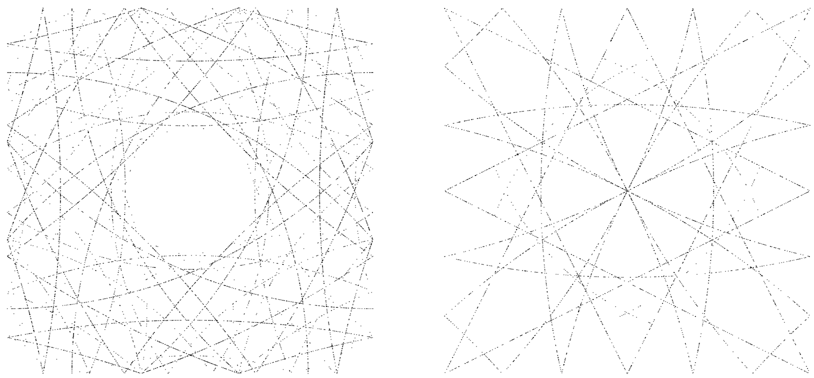

Figure 8: Remainders for $\sqrt{2}+i\sqrt{5}$ and $\sqrt{2}+i\sqrt{3}$.

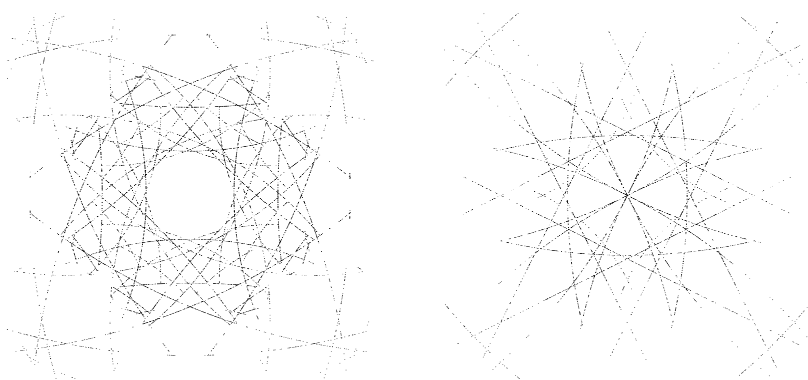

Figure 9: Ratios of successive q_j for $\sqrt{2}+i\sqrt{5}$ and $\sqrt{2}+i\sqrt{3}$.

square, there is a finite set $\mathcal{A}$ of circles and lines in $\mathbb{C}$ such that for all $j \geq 1$, $T^j z_0 \in \cup_{A\in\mathcal{A}} A \cap B$.

Proof. Let $C(\gamma, r)$ denote the circle in $\mathbb{C}$ about γ of radius r. Fix positive integers u and v, neither square, and assume that uv is also not square. Let $a_0 = [\sqrt{u} + i\sqrt{v}]$, and let $z_0 = \sqrt{u} + i\sqrt{v} - a_0 \in B$. As z_0 is not a Gaussian rational, the Hurwitz continued fraction expansion of z_0 is infinite. Let $\langle z_0, z_1, \ldots\rangle$ be the sequence of remainders, and $\langle a_0, a_1, \ldots\rangle$ the sequence of partial quotients, thus generated. Let $\langle p_0, p_1, \ldots\rangle$ and $\langle q_0, q_1, \ldots\rangle$ the sequence of numerators and denominators, generated by the Hurwitz algorithm on input $\sqrt{u} + i\sqrt{v}$. By convention, $p_{-1} = 1, q_{-1} = 0, p_0 = a_0$, and $q_0 = 1$. Let $w_0 = 0$, and for $j \geq 1$, let $w_j = q_{j-1}/q_j$. Equivalently, for $j \geq 1$, $w_j = 1/(a_j + w_{j-1})$.

We now consider a variant of the Hurwitz algorithm in which we keep track not only of the points $z_n = T^n z_0$, but of particular circles or lines on which z_n must lie. Our variant algorithm starts at an initial state $(z_0, C(-a_0, \sqrt{m})$. All states have the form (x, K), where K is either a circle passing through B of the form $(C(\gamma/d, \sqrt{m}/|d|))$ with $\gamma \in G^*$, $d \in \mathbb{Z}^*$, and $d \mid (|\gamma|^2 - m)$, and with radius $r = \sqrt{m}/|d|$ satisfying $1/(2\sqrt{2}) \leq r \leq \sqrt{m}$, or a line passing through B of the form $L(k/(2\delta), i\overline{\delta})$ where $k \in \mathbb{Z}$ and $|\delta|^2 = m$.

Let $\mathcal{T}(x, K) = (1/x - a, \phi_a(K))$ where $a = [1/x]$ and ϕ_a is the linear fractional map given by $\phi_a(z) = 1/z - a$.

Each statement below is readily verified and the details are left to the reader.

1. If $x \in K \cap B$, K is a circle of the form $C(\gamma/d, \sqrt{m}/|d|)$ with $d \in \mathbb{Z}^*, d \mid (|\gamma|^2 - m) \neq 0$, and $\sqrt{m}/d \geq 1/(2\sqrt{2})$, then $\mathcal{T}(x, K) = (x', K')$ where $x' \in K' \cap B$ and K' is another circle of the same form satisfying the same conditions.

2. If $x \in K \cap B$, K is a circle of the form $C(\gamma/d, \sqrt{m}/d)$ with $d \in \mathbb{Z}^*$, $|\gamma|^2 = m$, and $\sqrt{m}/d \geq 1/(2\sqrt{2})$, then $\mathcal{T}(x, K) = (x', K')$ with $x' \in B \cap K'$, and K' is the line $L((d - 2\Re(a\gamma))/(2\gamma), i\overline{\gamma})$, which passes through B.

3. If $x \in K \cap B$ and K is a line of the form $L(0, i\overline{\delta})$ with $|\delta|^2 = m$, then $\mathcal{T}(x, K) = (x', K')$ where $K' = L(-\Re(a\overline{\delta})/\overline{\delta}, i\delta)$.

4. If $x \in K \cap B$ and K is a line of the form $L(k/2\delta, i\overline{\delta})$ with $k \in \mathbb{Z}^*$, and $|\delta|^2 = m$ then $\mathcal{T}(x, K) = (x', K')$ where $x' \in K' \cap B$ and $K' = C\left(\frac{\overline{\delta} - ak}{k}, \frac{|\delta|}{|k|}\right)$, the new circle passes through 0 so that it has the required form, and the radius $|\delta|/|k| = \sqrt{m}/|k|$ is at least $1/\sqrt{2}$ because the diameter of K' is at least $\sqrt{2}$ because it is the reciprocal of the distance from 0 to the point on K nearest 0, and that nearest point lies within B and thus lies at a distance of less than $1/\sqrt{2}$ from 0.

There are only finitely many lines of the form $L(k/2\delta, i\overline{\delta})$ passing through B, with $|\delta|^2 = m$. There are only finitely many circles of the form $C(\gamma/d, \sqrt{m}/|d|)$ with $\gamma \in G$, $d \in \mathbb{Z}^*$, $d \mid (|\gamma|^2 - m) \neq 0$, and $\sqrt{m}/d \geq 1/(2\sqrt{2})$. We take $\mathcal{A}$ to be the set of all such lines and circles. ∎

We now turn to the topic of scaled errors. Let e_j be the scaled error for the jth convergent. Since $z_0 = (p_j + z_n p_{j-1})/(q_j + z_n q_{j-1})$, and since $\begin{vmatrix} p_{j-1} & p_j \\ q_{j-1} & q_j \end{vmatrix} = (-1)^j$, the scaled errors are also given by

$$e_j = \frac{(-1)^j \overline{q_j} z_j}{q_j + z_j q_{j-1}}.$$

Theorem 4. *Let u and v be positive integers. Let $\alpha = \sqrt{u} + i\sqrt{v}$. Let p, q be Gaussian integers with $|q|^2 \geq 9$, and assume $|p - q\alpha| \leq 2\sqrt{2}/|q|$. Let $e(q, p, \alpha)$ be the scaled error $e(q, p, \alpha) = \overline{q}(q\alpha - p)$. Then*

$$\Re(\overline{\alpha} e(q, p, \alpha)) = \frac{n}{2} + \delta$$

where $n = (|p|^2 - (u + v)|q|^2)$ is an integer satisfying $|n| \leq 8\sqrt{u + v}$ and $\delta \in \mathbb{C}$ with $|\delta| \leq 4/|q|^2$.

Remark 1. *In other words, the scaled errors lie almost, but not quite, on a finite set of parallel lines. The amount by which the scaled errors stand off from these lines is itself at most comparable to $|q|^{-2}$.*

Proof. Let $p = p_1 + ip_2$, $q = q_1 + iq_2$ with $p_1, p_2, q_1, q_2 \in \mathbb{Z}$. Let $c = p_1 q_1 + p_2 q_2$, $d = p_2 q_1 - p_1 q_2$, so that $c^2 + d^2 = |p|^2 |q|^2$. Let $\theta_1 + i\theta_2 = e(q, p, \alpha)$. Then the given conditions are then equivalent to

$$\begin{aligned} &(p_1 + ip_2)(q_1 - iq_2) = |q|^2(\sqrt{u} + i\sqrt{v}) + \theta_1 + i\theta_2, \\ &e(q, p, \alpha) = (c - |q|^2\sqrt{u}) + i(d - |q|^2\sqrt{v}) = \theta_1 + i\theta_2, \\ &\theta_1^2 + \theta_2^2 \leq 8. \end{aligned}$$

Thus, $\overline{\alpha} e(q, p, \alpha) = (\theta_1\sqrt{u} + \theta_2\sqrt{v}) + i(\theta_2\sqrt{u} - \theta_1\sqrt{v})$, and

$$\Re(\overline{\alpha} e(q, p, \alpha)) = \theta_1\sqrt{u} + \theta_2\sqrt{v} = (c\sqrt{u} - |q|^2 u) + (d\sqrt{v} - |q|^2 v).$$

Turning this around, we have $c = |q|^2\sqrt{u} + \theta_1$, $d = |q|^2\sqrt{v} + \theta_2$, and so

$$|p|^2|q|^2 = c^2 + d^2 = |q|^4(u + v) + 2|q|^2(\theta_1\sqrt{u} + \theta_2\sqrt{v}) + (\theta_1^2 + \theta_2^2).$$

Hence,

$$(\theta_1\sqrt{u} + \theta_2\sqrt{v}) = \frac{1}{2}(|p|^2 - (u + v)|q|^2) - \frac{\theta_1^2 + \theta_2^2}{2|q|^2}.$$

With $n = |p|^2 - (u + v)|q|^2$, we note that

$$\begin{aligned} |n| &\leq |p^2 - (\sqrt{u} + i\sqrt{v})^2 q^2| = |p - \alpha q||2\alpha q + (p - \alpha q)| \\ &\leq \frac{2\sqrt{2}}{|q|}\left(2\sqrt{u + v}|q| + \frac{2\sqrt{2}}{|q|}\right) \leq (4\sqrt{2}\sqrt{u + v} + 8/|q|^2) \leq 8\sqrt{u + v}. \end{aligned}$$

□

This takes us back to the question of just what *is* typical of the orbits $T^n z$?

7 The Gauss-Kuz'min Density for the Hurwitz Algorithm

There can be more than one measure on a set that is invariant under a self-mapping T. There can, in principle, be more than one that is absolutely continuous with respect to Lebesgue measure.

We are certainly not exempt from the first possibility, as there are periodic orbits and uniformly weighted discrete measures on these are invariant under T. But the second pathology does not exist.

Theorem 5. *There is a unique measure μ on B that is absolutely continuous with respect to Lebesgue measure and invariant under T. This measure has a density function ρ; the density is continuous except perhaps along the arcs $|z \pm 1| = 1$, $|z \pm i| = 1$, and $|z \pm 1 \pm i| = 1$. It is real-analytic on an each of the 12 open regions into which B° is dissected by these arcs. The mapping T on B is* mixing *with respect to μ: for Lebesgue-measurable $A_1, A_2 \subseteq B$,*

$$\lim_{n\to\infty} \mu(T^{-n}A_1 \cap A_2) = \mu(A_1)\mu(A_2).$$

The expected symmetry obtains: $\mu(iA) = \mu(A)$ and $\mu(\overline{A}) = \mu(A)$ for all measurable A.

The general plan of the proof is this: We show that if f is a continuous probability density function on B and X a random variable on B with density f, then there exists a positive integer n, depending on f, and $\epsilon > 0$, such that the density of $T^n X$ is greater than ϵ everywhere in B°.

The theory of positive operators [4] can then be applied. The result we use is an extension of the Perron-Frobenius theorem to certain positive operators. (The classical Perron-Frobenius theorem states that if M is a square matrix with nonnegative entries, and if there exists n such that all entries of M^n are positive, then the largest eigenvalue of M is positive and has multiplicity 1, and the corresponding eigenvector has positive entries.) What we first need is a Banach space $\mathcal{V}$ of functions on B and a positive compact linear operator L on $\mathcal{V}$ such that if a random variable X has density $f \in \mathcal{V}$ then the density of TX is Lf. Then we need an element z_0 of $\mathcal{V}$ with the property that for all nonzero positive $f \in \mathcal{V}$, there exists an n, and positive constants c_1 and c_2, such that $c_1 z_0 \le L^n f \le c_2 z_0$. The operator L is then said to be u_0-positive, and as a consequence, the operator L has a decomposition as $L = \lambda P + N$, where P is a positive projection of $\mathcal{V}$ onto a one-dimensional subspace of $\mathcal{V}$, λ a positive real number, and N an operator of spectral radius less than λ such that $PN = NP = 0$.

In our specific circumstances, some care must be given to the selection of the Banach space $\mathcal{B}$ in which our transfer operator will lie. We recall that $T : B\backslash\{0\} \to B$, with $Tz = 1/z - [1/z]$. We want $\mathcal{B}$ to consist of reasonably nice functions, and we want the transfer operator L_T for T to be compact. Since the formula for the transfer operator is basically set in stone by the requirement that it *be* the transfer operator for T, our only latitude is in the selection of $\mathcal{B}$. We use the ideas of O. Bandtlow and O. Jenkinson from their recent work [1]. Their work features a single compact connected subset X of $\mathbb{R}$ and a map $T: X \to X$ that is "real-analytic full branch expanding with countably many branches." In our setting, it might at first seem as though we could take $X = B$ and use T as is, but unfortunately, the part about "full branch" fails. The conclusion they reach fails as well, at least on the numerical evidence, which should not be surprising as the premises were not met. The invariant density for our T is, on the numerical evidence, discontinuous along the arcs through B of the eight circles that dissect it into 12 distinct sectors.

Nevertheless, their work can be adapted to the situation at hand. The trick is to stitch together $\mathcal{B}$ as the Cartesian product of 12 distinct subsidiary Banach spaces $\mathcal{B}_1 \ldots \mathcal{B}_{12}$, and to break L_T up into a matrix of 144 compact operators carrying the various $\mathcal{B}_j$ to each other.

First, we describe the zones into which the numerical evidence prompts us to expect we must dissect B. Our zones B_j are connected, compact subsets of B with disjoint interiors, whose union is $\overline{B}$. We take the view that B is a subset of $\mathbb{R}^2$ rather than a subset of $\mathbb{C}$, we set $R: \mathbb{R}^2 \to \mathbb{R}^2, R(x, y) = (-y, x)$ (a counterclockwise rotation through an angle of $\pi/2$ radians), and we set

$$
\begin{aligned}
&B_1 = \{(x, y) \mid 0 \le x \le 1/2, x^2 + (y-1)^2 \ge 1, \text{and } x^2 + (y+1)^2 \ge 1\},\\
&B_2 = \{(x, y) \mid x^2 + (y-1)^2 \le 1, (x-1)^2 + y^2 \le 1, \text{and}\\
&\quad (x-1)^2 + (y-1)^2 \ge 1\},\\
&B_3 = \{(x, y) \mid x \le 1/2, y \le 1/2, \text{ and } (x-1)^2 + (y-1)^2 \le 1\},\\
&B_4 = RB_1,\ B_5 = RB_2,\ B_6 = RB_3,\ B_7 = -B_1,\ B_8 = -B_2,\\
&B_9 = -B_3,\ B_{10} = -B_4,\ B_{11} = -B_5,\ B_{12} = -B_6.
\end{aligned}
$$

We take $U = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 < 1\}$ to be the open unit ball in $\mathbb{C}$, and, reserving the choice of ϵ but with $0 < \epsilon < 1/10$, we set $D_k = B_k + \epsilon U$, $1 \le k \le 12$. Thus, D_k is an ϵ-neighborhood of B_k, and all elements of D_k are close to a real pair $(x, y) \in B_k$.

Let $\psi: \mathbb{C}^2 \backslash \{(z, \pm iz) \mid z \in \mathbb{C}\} \to \mathbb{C}^2$ be given by

$$\psi(z_1, z_2) = \left(\frac{z_1}{z_1^2 + z_2^2}, \frac{-z_2}{z_1^2 + z_2^2} \right).$$

Remark 2. *Of course this is a thinly disguised translation of the mapping $1/z$ acting on $\mathbb{C}$. But we need a map on pairs of complex numbers that tracks what happens to a single complex number in terms of the real and imaginary parts.*

Let $G' \subset \mathbb{Z} + i\mathbb{Z} = \{a_1 + ia_2 \mid a_1, a_2 \in \mathbb{Z}, a_1 + ia_2 \ne 0, \pm 1, \pm i\}$, and for $1 \le k \le 12$, let $G_k = \{a \in G' \mid \psi(a + B_k) \subseteq \overline{B}.\}$. Then $G_1 = G' \backslash \{-1 \pm i, -2\}$, $G_2 = G' \backslash \{-1 \pm i, -2, -2i\}$, and $G_3' = G' \backslash \{-1 \pm i, 1-i, -2, -2i, -2-i, -1-2i\}$. Together with $G_{k+3} = iG_k$, this describes G_k, $1 \le k \le 12$.

Now T maps the arcs partitioning B into themselves or the boundary of B, so for each $a \in G'$, (abusing terminology by equating a with $(a_1, a_2) = (\Re a, \Im a)$), either $\psi(a + B_k) \subseteq \mathbb{R}^2 \backslash B^\circ$, or for some j, $1 \le j \le 12$, $\psi(a + B_k) \subset B_j$.

Let $G_{j,k} = \{a \in G_k \mid \psi(a + B_k \subset B_j\}$, and for $z \in D_k$, let $T_a(z) = \psi(a_1 + z_1, a_2 + z_2)$. At this point we need a lemma.

Lemma 2. *For ϵ sufficiently small, and $a \in G_{j,k}$, $T_a(B_k + \epsilon U) \subset B_j + \frac{4}{5}\epsilon U$.*

(That is, with an appropriate choice of ϵ, T_a maps D_k compactly into D_j, and T_a takes B_k into B_j.)

Proof. It will suffice to show that for $a \in G_k$ and $z \in B_k + \epsilon U$, $\|\psi'\| < 2/3$. Here, ψ' is the matrix derivative of ψ, and for $w \in \mathbb{C}^2$, $\|w\| = \sqrt{|w_1|^2 + |w_2|^2}$. Now for $a \in G_k$ and $z \in \epsilon U + B_k$, there exists $s \in a + B_k$ and $\zeta \in \epsilon U$ such that $z = s + \zeta$.

Consider $1/\overline{B}$ as a subset of $\mathbb{R}^2$. It is the exterior of the union of the open disks about $(\pm 1, 0)$ and $(0, \pm 1)$ of radius 1. Thus if $(a_1, a_2) \in G_k$ and $(s_1, s_2) \in (a_1, a_2) + B_k \subset 1/\overline{B}$, then $s_1^2 + s_2^2 \geq 2$.

It follows that for $(\zeta_1, \zeta_2) \in \epsilon U$ and $(s_1, s_2) \in 1/\overline{B}$,

$$\begin{aligned}\left|(s_1 + \zeta_1)^2 + (s_2 + \zeta_2)\right|^2 &\geq (s_1^2 + s_2^2) - 2\left|s_1\zeta_1 + s_2\zeta_2\right| - \epsilon^2 \\ &\geq (s_1^2 + s_2^2) - 2\epsilon(s_1^2 + s_2^2)^{1/2} - \epsilon^2 \geq \|s\|^4(1 - 2\epsilon).\end{aligned}$$

Now with $z = s + \zeta$ with $s \in 1/\overline{B}$ and $\zeta \in \epsilon U$, $\psi'(z) = (z_1^2 + z_2^2)^{-2} M_z$, where $|(z_1^2 + z_2^2)|^2 \geq (1 - 2\epsilon)\|s\|^4$ and where

$$M_z = \begin{pmatrix} z_2^2 - z_1^2 & -2z_1z_2 \\ 2z_1z_2 & z_2^2 - z_1^2 \end{pmatrix},$$

we have

$$M_z = M_s + 2\begin{pmatrix} s_2\zeta_2 - s_1\zeta_1 & -s_1\zeta_2 - s_2\zeta_1 \\ s_1\zeta_2 + s_2\zeta_1 & s_2\zeta_2 - s_1\zeta_1 \end{pmatrix} + M_\zeta.$$

But the entries of $M_{s,\zeta}$ are each bounded in absolute value by $\epsilon(s_1^2 + s_2^2)^{1/2}$ by Cauchy's inequality, so $\|M_{s,\zeta}\| \leq 2\epsilon(s_1^2 + s_2^2)$ and similarly $\|M_\zeta\| \leq 2\epsilon^2$. From this it follows that for ϵ sufficiently small, $\|M_z\| \leq \|M_s\| + 3\epsilon\|s\|$. But by direct calculation, we see that $\|M_s\| = \|s\|^2$, so

$$\|M_z\| \leq \tfrac{4}{5}(1 - 2\epsilon)\|s\|^4 \leq \tfrac{4}{5}|z_1^2 + z_2^2|^2.$$

Thus $\|\psi'\| \leq 4/5$ for $z \in 1/\overline{B} + \epsilon U$. Since T_a mapped B_k into B_j, and since excursions from B_k in the input to T_a are damped by a factor of 4/5 or less in the output, the lemma is proved. ☐

Now consider the space $H^\infty(D_k)$ of bounded holomorphic functions from D_k into $\mathbb{C}$, equipped with the sup norm. Let $\mathcal{B}_k$ be the subspace of $H^\infty(D_k)$ consisting of elements that take B_k into $\mathbb{R}$, (equipped with the same norm.) Let $D_j'' = \cup_{a \in G_{j,k}} T_a D_k$. As we have seen in our lemma above, D_j'' fits inside a compact subset of D_j. We take D_j' to be a connected, open subset of D_j containing D'' and contained in a compact subset of D_j. Within $H^\infty(D_j')$ we take $\mathcal{B}_j'$ to be the subspace consisting of all $f \in H^\infty(D_j')$ such that f is real on $D_j' \cap B_j$.

For $a \in G_{j,k}$, $f \in H^\infty(D'_j)$ and $z \in D_k$ let

$$\begin{aligned} w_a(z) &= ((a_1 + z_1)^2 + (a_2 + z_2)^2)^{-2}, \\ L_a f &= w_a(z) f(T_a(z)), \\ L_{j,k} &= \sum_{a \in G_{j,k}} L_a. \end{aligned}$$

Lemma 3. *$L_{j,k}$ is a bounded linear operator from $H^\infty D'_j$ into $H^\infty D_k$, it is compact, and it maps $\mathcal{B}'_j$ into $\mathcal{B}_k$.*

Proof. The proof of the first claim goes exactly as in prop. 2.3 of [1], which (since this paper has not yet appeared) we quote with a few minor changes to adapt it to our slightly different circumstances. We take $\mathcal{L}$ to be $L_{j,k}$. In their work, a single domain serves both as D_k and D'_j, so all mention of either of these is our own adaptation; a verbatim quotation would have had D in place of both. We also use w_a in place of their w_i, and our set of subscripts is $G_{j,k}$ rather than $\mathbb{N}$.

> Let $S := \sup_{z \in D_k} |w_a(z)| < \infty$. First we show that $\mathcal{L}$ maps $H^\infty(D'_j)$ to $H^\infty(D_k)$. Fix $f \in H^\infty(D'_j)$. If
>
> $$g_k(z) = \sum_{|a| \leq k, a \in G_{j,k}} w_a(z) f(T_a(z))$$
>
> for $k \in \mathbb{N}$, then $g_k \in H^\infty(D_k)$. Since
>
> $$|g_k(z)| \leq \sum_{|a| \leq k, a \in G_{j,k}} |w_a(z)| \|f(T_a(z)\| \leq S \|f\|_{H^\infty(D'_j)} \qquad (*)$$
>
> for all $k \in \mathbb{N}$, we see that the sequence $\{g_k\}$ is uniformly bounded on D'_j. Moreover, $\lim_{k\to\infty} g_k(z) =: g(z)$ exists for every $z \in D_k$. By Vitali's convergence theorem (see e.g. [6], Chapter 1, Prop. 7) g_k thus converges uniformly on compact subsets of D_k. Hence, g is analytic on D_k. Moreover, by (*) we see that $|g(z)| \leq S\|f\|$ for any $z \in D_k$, so $g \in H^\infty(D_k)$. Thus, $\mathcal{L}f = g \in H^\infty(D_k)$.
>
> In order to see that $\mathcal{L}$ is continuous we simply note that
>
> $$\begin{aligned} \|\mathcal{L}f\|_{H^\infty(D_k)} &= \sup_{z \in D_k} \left| \sum_{a \in G_{j,k}} w_a(z) f(T_a(z)) \right| \\ &\leq \sup_{z \in D_k} \sum_{a \in G_{j,k}} |w_a(z)| |f(T_a(z)| \\ &\leq S \|f\|_{H^\infty(D'_j)}. \end{aligned}$$

We next introduce their 'canonical embedding' operator J, defining J to be the operator which takes any element of $H^\infty(D_j)$ and regards it as an element of

$H^\infty(D'_j)$. This, as they note, is, by Montel's theorem [6] Chapter 1, Prop. 6, a compact operator from $H^\infty(D'_j)$ into $H^\infty(D_j)$. But $L_{j,k} = \mathcal{L} \circ J$, and the composition of a bounded operator with a compact operator is compact. This proves the second claim. The third claim is obvious: if $z = (x, y) \in \mathbb{R}^2$, then $w_a(z) \in \mathbb{R}$ and $T_a(z) \in \mathbb{R}^2$, so for $f \in \mathcal{B}'_j$, $L_{j,k} f \in \mathcal{B}_k$. □

We are now at last in a position to define our overall transfer operator L. We take $\mathcal{B} := \mathcal{B}_1 \times \cdots \times \mathcal{B}_{12}$ to be our Banach space, equipped with the norm $\|f\| = \max_{1 \le j \le 12} \|f_j\|$, and we take L_T to be the operator on $\mathcal{B}$ given by

$$L_T \begin{pmatrix} f_1 \\ \vdots \\ f_{12} \end{pmatrix} = \begin{pmatrix} L_{1,1} & \cdots & L_{1,12} \\ \vdots & \ddots & \vdots \\ L_{12,1} & \cdots & L_{12,12} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_{12} \end{pmatrix}.$$

This L_T is compact since its components are compact.

Remark 3. *One point worth noting is that for $z = (x, y) \in \mathbb{R}$, $w_a(z) = ((a_1 + x)^2 + (a_2 + y)^2)^{-2} = |(a_1 + ia_2) + (x + iy)|^{-4}$. Thus if we had defined w_a as a map on $\mathbb{C}$ rather than on $\mathbb{C}^2$, then it would not have been holomorphic, and the arguments of [1] could not have been used. Another point is that this formulation of L_T opens the way for effective computation of the invariant density ρ, which will be a suitably normalized multiple of the one-dimensional eigenspace of L_T corresponding to the dominant eigenvalue 1. The operators $L_{j,k}$ will have matrices with respect to the expansion of f in a two-dimensional power series, and these matrices can be truncated with controllable errors, so that L_T becomes a computationally tractable object. A final point is that by taking full advantage of symmetry we can work with just $\mathcal{B}_1 \times \mathcal{B}_2 \times \mathcal{B}_3$.*

At this point, we may as well regard L_T as an operator acting on a space of real-valued functions of $B \subseteq \mathbb{C}$, since the details of just which functions count as elements of our Banach space, and just how we coped with the difficulties presented by the boundaries of the B_k, are no longer relevant.

Having now constructed $\mathcal{B}$ and L_T, we next set about demonstrating that L_T, seen as an operator on $\mathcal{B}$, has the needed positivity properties.

What are these? We need the following, in the topology induced by our norm $\|\cdot\|$ on $\mathcal{B}$. There is a *positive cone* P with the following properties, some intrinsic to P and others related to L_T. We write $f \ge g$ if $f - g \in P$.

1. P is closed under addition and scaling by positive numbers.

2. $P^\circ \ne \emptyset$.

3. P is a closed subset of $\mathcal{B}$.

4. For all $f \in \mathcal{B}$, there exist $f_1, f_2 \in P$ such that $f = f_1 - f_2$. (P is *reproducing*.)

5. L_T maps P into P.

6. There exists $u_0 \in P^\circ$ such that if $f \neq 0$ and $f \in P$ then there exist $n \in \mathbb{Z}^+$ and $c_1, c_2 \in \mathbb{R}^+$ such that $c_1 u_0 \leq L_T^n f \leq c_2 u_0$. ($L_T$ is u_0-positive.)

We define P to be the set of all $f \in \mathcal{B}$ such that each of the f_k is nonnegative on B_k. Clearly P° is nonempty, and P is closed in $\mathcal{B}$. Any element of $\mathcal{B}$ can be written as $f_1 - f_2$ with $f_1 = f + C$ where $C = 2\max_{j=1}^{12} \|f\|$ and $f_2 = C$. Clearly f maps P into P. We claim that the element $f \in \mathcal{B}$ given by $f_k \equiv 1$ can serve as u_0, but this is far from clear. The reason it is *true* is that, given any open disk E in any of the B_k, there exists an open squarish (bounded by arcs of circles that meet at right angles) subset Q of E, and a positive integer n, such that the image under T^n of Q is B°. This we prove by demonstrating the existence of such a domain Q that does not, during iteration of T, touch the boundaries of B until the end, when it covers B. To this end, we first demonstrate that given $x \in \widetilde{B}$ and $\epsilon > 0$, there exists a Gaussian rational $r' \in \widetilde{B}$ and a neighborhood A of r', such that the orbit of r' under T stays inside $\widetilde{B}$ until the end, when, as it eventually must, it reaches zero and stops.

Using the existence of r', we then show that there exists also r, another Gaussian rational near x, a positive integer n, and a neighborhood A_r of r and contained within A, with the following properties:

1. $T^k r \in \widetilde{B}$ for $1 \leq k < n$.
2. $T^n r = 0$.
3. If $a_1, a_2 \ldots$ are given by $a_k = [1/T^{k-1}r]$, and if ψ_1 is the linear fractional map $z \to 1/z - a_1$, and $\psi_k(z) = 1/\psi_{k-1}(z) - a_k$, then for $1 \leq k < n$, ψ_k takes A_r into the same one of the B_j to which $T_k r$ belongs, while the restriction of ψ_n to A_r is a conformal bijection from A_r to B°.

We begin our proof that such an r' exists by considering an auxiliary $s \in \widetilde{B}$ with $|s - x| < \epsilon/2$, and the histories of various elements of a small neighborhood of s under iteration of not only T, but all the variants of T got by truncating to the right, or top, of B instead of to the bottom and left as we have been in the habit of doing. We do this by means of a tree, which we shall call the *disk partition tree* of s, in which the vertices are ordered pairs $(P, \mathbf{a})$. The first entry is a 'triangle' in $\mathbb{C}$ with one vertex at s, bounded by arcs of a circle, and contained in the disk $|s - x| < \epsilon/2$. The second entry is a list $\mathbf{a}$ of Gaussian integers, all in G'. For vertices at distance j from the root of the tree, the triangles P are disjoint and fit together to cover an open disk $D_j \subset D_0$ centered on s, apart from their shared boundaries and from the perimeter of D_j. The lists $\mathbf{a}_P$ paired with these 'pizza slices' all have length j, and an edge extends down the tree from $(P, \mathbf{a}_P)$ to $(Q, \mathbf{a}_Q)$ if and only if $\mathbf{a}_Q$ is an extension of $\mathbf{a}_P$ and $Q \subseteq P$. The key point of the construction is that we arrange matters so that each slice P at depth j is carried intact to a region inside B° by all iterates of T of depth 1 through j . We now turn to the details.

Let $s_0 = s$. Since $s \in \widetilde{B}$, $1/s_0 \pmod{G} \in B^\circ$, though it could happen that $1/s_0 \pmod{G} \notin \widetilde{B}$. For however long it may be that this does not happen, we construct a list of further s_j, and a nested list D_j of open disks about s_0, starting

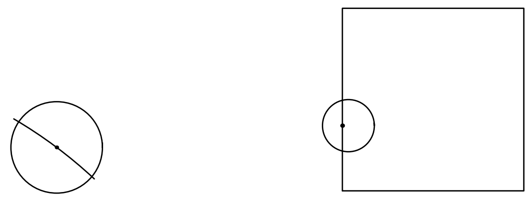

Figure 10: D_{k-1} cut by ψ^{-1}(edge) and $\psi[\mathbf{a}](D_{k-1})$ cut by ∂B.

with $D_0 = \{z \in \widetilde{B} : |z - s| < \epsilon/4\}$. We put $a_1 = [1/s_0]$ and $s_1 = 1/s_0 - a_1$, and we take D_1 to be a (possibly smaller) disk about s_0 such that $1/D_1 - a_1 \subseteq B^\circ$. We continue the construction taking $a_2 = [1/s_1]$, $s_2 = 1/s_1 - a_2$, and $D_2 \subseteq D_1$ so that

$$\frac{1}{\frac{1}{D_2} - a_1} - a_2 \subseteq B^\circ,$$

and so on. Eventually, (unless we are lucky) there will be a first integer k such that $T^j s_0 \in \widetilde{B}$ for $1 \leq j < k$ but $1/T^{k-1}s_0$ has either its real part, or its imaginary part, on the boundary of some translate of B by a Gaussian integer.

To track what now happens, we bring in some linear fractional mappings defined in terms of a list $\mathbf{a}$ of Gaussian integers. If $\mathbf{a} = (a_1, \ldots, a_k)$, let $\psi[a, j]$ be the linear fractional given by

$$\psi[\mathbf{a}, j](z) = \begin{cases} \frac{1}{z} - a_1, & j = 1 \\ \frac{1}{\psi[\mathbf{a}, j-1](z)} - a_j, & 1 < j \leq k \end{cases}.$$

Let $\psi[\mathbf{a}] = \psi[\mathbf{a}, k]$. Thus if we put $\psi_g := z \to 1/z - g$, then $\psi[\mathbf{a}, j] = \psi_{a_j} \circ \psi_{a_{j-1}} \cdots \circ \psi_{a_1}$. By our construction, if $\mathbf{a} = (a_1, \ldots, a_{k-1})$ then for $1 \leq j \leq k-1$, $\psi[\mathbf{a}, j]$ maps D_k conformally into B°, but $s_{k-1} = T^{k-1}s_0$ lies on the boundary of B. Let $a_k = [1/s_k]$, and consider the inverse image under $\psi[(a_1, \ldots, a_k), k]$ of the line (or, in the exceptional case that $s_k = (-1-i)/2$, two lines) in $\mathbb{C}$ and bounding B on which s_k lies. This line, (or these lines), will cut D_k into two or four pieces. We take the disks D_j small enough that these circles, inverse images of one of the lines bounding B, meet each other only at the center, and so that they all exit all the D_j.

Now $\psi[(a_1, \ldots, a_k)]$ maps one of the parts into which this arc cuts D_k conformally into B°, and carries the other part to the exterior of B. But $\psi[(a_1, \ldots, a_{k-1}, 1 + a_k)]$ carries that part of D_k conformally into B°. In the exceptional case that $T^k s_0 = -\frac{1}{2}(1+i)$, we would see instead a circle on the left about s_0, roughly quartered, mapped by ψ to a disk enclosing, (though typically not centered at) $-\frac{1}{2}(1+i)$. One of the quarters would map to inside B°, while the others would be mapped there by a modification of ψ corresponding to a change in the last entry in $\mathbf{a}$. Note that this exceptional case can only occur at stage n, because $T(-\frac{1}{2}(1+i)) = 0$.

Both in the main case, and in this special case, D_k has been partitioned into a finite number of 'slices', together with an equal number of bounding arcs each running along some circle from the center s_0 to the rim of D_k, and for each of the slices P, there is a corresponding list $\mathbf{a}_P$ of Gaussian integers, and a corresponding linear fractional map $\psi[\mathbf{a}_P]$, that maps $P \subset D_0$ conformally into B°. This begins the construction of a list of similar partitions $\mathcal{P}_j$ of neighborhoods D_j of s_0.

The vertices of the tree are pairs $(P, \mathbf{a})$. Edges join $(P, \mathbf{a})$ and $(Q, \mathbf{a}^*)$ if $\mathbf{a}^*$ is an extension of $\mathbf{a}$ to a list one entry longer, and in this case, $Q \subseteq P$.

Remark 4. *Normally, $Q \subset P$, as we shall have to shrink our neighborhoods of s_0 as construction proceeds. Any mapping $z \to 1/z - a$ is expanding when acting on B, and since we want the image of D_j under a composition of linear fractionals of this sort to be a subset of B°, we must make it small enough that the expansions to come do not cause it to overflow these bounds.*

The first $k-1$ stages have an open disk D_j centered on s_0, associated with an integer list $\mathbf{a}_j$ of length j, each an extension of the previous, and with the property that for $1 \leq \ell \leq j$, $\psi[\mathbf{a}_j, \ell]$ maps D_j conformally into B°. Beginning with D_k, and continuing until such time as $T^n s_0 = 0$, we have disks D_j, and associated partitions $\mathcal{P}_j$ of D_j, such that for each slice $P \in \mathcal{P}_j$, there is a list $\mathbf{a}_P = (a_1, \ldots, a_j)$ of Gaussian integers such that for $1 \leq \ell \leq j$, $\psi[\mathbf{a}_P, \ell]$ maps P conformally into B°, is injective on D_j, and the inverse image under $\psi[\mathbf{a}_P, \ell]$ of the lines $\Re(z) = \pm\frac{1}{2}$ and $\Im(z) = \pm\frac{1}{2}$ bounding B° either miss D_j entirely (if $T^j s_0 \notin \partial B$), or cut through D_j in a single arc through s_0 (if $T_j s_0$ lies on just one of these bounding lines), or quarter D_j, if $T_j s_0 = -\frac{1}{2}(1+i)$. In any case, we take D_j small enough that all these arcs enter and exit D_j, and none of them meets any of the others except at s_0.

Assuming we have carried this construction to depth m, with $k \leq m < n$, we now detail how it is extended to depth $m+1$. Each $P \in \mathcal{P}_m$ is a 'triangle' bounded by three circular arcs. Two of these arcs have the form of the preimage under $\psi[\mathbf{a}_P, m]$ of a segment of one of the lines bounding B°, and s_0 lies at the intersection of these, while the third arc is part of the rim of D_m. For $1 \leq \ell \leq m$, if $z \in P$, then $T^\ell z = \psi[\mathbf{a}_P, \ell](z) \in B^\circ$.

Remark 5. *Once we hit the first occasion on which $T^k s_0 \in \partial B$, the future orbits of s_0 under the various $\psi_P = \psi[\mathbf{a}_P]$ are confined to the union of the boundary of B and the internal boundaries of the B_k, as this set is closed under T and its variants that round to the right or up.*

Now consider the region

$$R_P = \frac{1}{\psi_P(P)} - \left[\frac{1}{\psi_P(s_0)}\right].$$

If the $T(\psi_P(s_0)) \in B^\circ$, then there is a new, possibly shrunken, neighborhood D_{m+1} of s_0 such that, with $a' = [1/\psi_P(s_0)]$, we have $(1/\psi_P(z)) - a' \in B^\circ$ for all $z \in D_{m+1} \cap P$. We won't need to split P, though we may have to trim it back towards its vertex s_0.

If, on the other hand, $b = 1/\psi_P(s_0) - a' \in \partial B$, we introduce a new arc cutting D_m through s_0. This arc is the inverse image under $z \to 1/\psi_P(z) - a'$ of the line bounding B° on which b lies. Say that an arc α *cuts* P if it meets P in any neighborhood of s_0. Our new arc may or may not cut P. If it does not cut P, we take a suitably shrunken portion of P about s_0 to be one of the pieces of $\mathcal{P}_{m+1}$, and we extend the Gaussian integer list $\mathbf{a}$ by appending a', to provide the pair $(P, \mathbf{a}')$ say. If it does cut P, we take suitably shrunken portions of the two sectors into which the portion of P near s_0 has been cut by our arc, to be the two pieces in $\mathcal{P}_{m+1}$ corresponding to P.

One of these pieces will be associated with the same extension of $\mathbf{a}$ we used before, while for the other, we extend instead by $a'+1$, if b was on the left edge of B, or by $a'+i$, if b was on the bottom edge of B. We now further truncate our slices, if necessary, to a radius which is the minimum of the radii of the pieces we got during the construction of $\mathcal{P}_{m+1}$. The union of these pieces, together with their internal boundaries with each other, is now a pie-slice partition of an open disk D_{m+1} centered on s_0, and for each $P \in \mathcal{P}_{m+1}$, there is a corresponding Gaussian integer list $\mathbf{a}$, an extension of the one that was associated with the parent P in $\mathcal{P}_m$, such that $\psi[\mathbf{a}, \ell]$ maps P into B° for $1 \leq \ell \leq m+1$.

We now claim that for all $P \in \mathcal{P}_{n-1}$, $1/\psi[\mathbf{a}_P](s_0) \in G'$. That is, the variants of the algorithm all run to completion on Gaussian rational input in the same number of steps. The reason for this is symmetry. The group S of rigid motions taking $\overline{B}$ onto $\overline{B}$ is generated by $\lambda_1(z) = \overline{z}$ and $\lambda_2(z) = iz$. We consider the equivalence relation $\sim$ on $\mathbb{C}$ determined by $u \sim v$ if $u = \lambda v$ for some $\lambda \in S$.

If $z \sim w$ then $1/z \sim 1/w$, and if $z, w \in \overline{B}$ and $z \sim w$, and if $a, b \in G'$ such that $1/z - a \in \overline{B}$ and $1/w - b \in \overline{B}$, then $1/z - a \sim 1/w - b$. Initially, $s_0 \sim s_0$. If $\psi[\mathbf{a}_1, \ell](s_0) = s_l \sim \psi[\mathbf{a}_2, \ell](s_0) = s_l'$ for all pairs $(\mathbf{a}_1, \mathbf{a}_2)$ of lists of length ℓ occurring in the disk partition tree of s_0, then $1/s_\ell \sim 1/s_{\ell'}$, and now for any of the choices a^* by which $\mathbf{a}_1$ may be extended,the various $1/s_\ell - a^*$ are equivalent under $\sim$ to each other, and to the complex numbers that may arise out of $1/s_{\ell'}$ by extending $\mathbf{a}_2$ with a^{**}, say. (Another way to say this is to say that if we identified all elements of a given equivalence class under $\sim$, all the variants of T would condense to one.)

At the end, when $s_{n-1} \in 1/G'$, all other $s \sim s_{n-1}$ also belong to $1/G'$, so T maps them to zero as well. Therefore, all branches of our disk partition tree for s_0 have the same depth.

Now for each $(P, \mathbf{a}_P)$ of depth n in our disk partition tree,

$$\psi[\mathbf{a}, n](z) = \left(\frac{1}{z} - a_n\right) \circ \left(\frac{1}{z} - a_{n-1}\right) \cdots \circ \left(\frac{1}{z} - a_1\right)$$

takes s_0 to 0, and takes P on an orbit that at each stage remains inside B°.

Since these slices P partition the final disk D_n about s_0 into a finite number of slices, and since the sum of the angles at s_0 of the slices is 2π, there must be at least one slice that occupies a positive angle at s_0. Let P' be such a slice, and $\mathbf{a}'$ its Gaussian integer list. The image under $\psi[\mathbf{a}']$ of P is a 'triangle' in B° with its vertex $\psi[\mathbf{a}'](s_0) = 0$ with a positive angle at 0. For all $z \in P'$, $T^n z = \psi[\mathbf{a}'](z)$. Our disk partition tree has now served its purpose.

There are Gaussian integers $g \in G'$ such that $1/g \in P'$. Any neighborhood E of $1/g$ contained in P' will be taken conformally under $z \to 1/z - g$ to a neighborhood $\widetilde{E}$ of 0. We take another Gaussian integer h such that $1/(h+\overline{B}) \subseteq \widetilde{E}$, and such that $1/(h+\overline{B})$ is also a subset of some particular B_r. Let $\widetilde{\mathbf{a}}$ be the list got by appending g and then h to $\mathbf{a}'$. Now let $r = \psi[\mathbf{a}']^{-1}(1/(g+1/h)) = \psi[\widetilde{\mathbf{a}}]^{-1}(0)$, and let

$$\widetilde{D} \subset D_n = \psi[\mathbf{a}']^{-1} \circ \left(\frac{1}{z+g}\right) \circ \left(\frac{1}{z+h}\right) B^\circ.$$

Iterating T takes $\widetilde{D}$ conformally to various squarish neighborhoods inside B°, until T^{n+1} takes it to the squarish region $(1/(h+B^\circ)) \subset \widetilde{E}$, and finally T^{n+2} carries it via $z \to 1/z - h$ onto B°.

A bit more is true. The orbit of $\widetilde{D}$ never meets any of the internal boundaries separating the various B_k until the last step, when it is splashed across B°. This is because if, at any earlier point any part of the orbit had touched an internal boundary, the subsequent iterations of T on that point would have also been in the union of these internal boundaries with the boundary of B itself, and yet, at the stage immediately prior to the final one, no point of $T^{n+1}\widetilde{D}$ belongs to any of these boundaries.

Now that we have our magic r, we are in a position to complete the proof that L_T is u_0-positive with $u_0 \equiv 1$. If $f \neq 0 \in \mathcal{V}$, and $f \geq 0$, then there exists $s \in \widetilde{B}$ such that $f(s) > 0$. We choose an open disk D_0 around s such that $D_0 \subset B_k$ where B_k is the one containing s, and such that for $z \in D_0$, $f(z) > \frac{1}{2}f(s)$. Now for $z \in \widetilde{B}$, $(L_T^{n+2}f)(z)$ is a sum of terms, among which (keeping in mind that $[c_1, c_2, \ldots, c_k] = 1/(c_1 + 1/(c_2 + \ldots + 1/c_k)))$ is the term

$$\prod_{j=1}^{n+2} |[a_j, \ldots, a_{n+2} + z]|^4 \, f([\langle a_{n+2}, a_{n+1}, \ldots, a_2, a_1 + z\rangle]).$$

But for arbitrary $z \in B^\circ$, the argument of f lies inside D_0, and the product is bounded below by a positive constant on B°, so this term is bounded below by a positive constant on $\widetilde{B} \subset B^\circ$. Since L_T is a bounded linear operator, we are almost done. We have proved most of the claims in our Gauss-Kuz'min theorem.

For the symmetry claims, we first note that $\rho(z) = \rho(\overline{z})$. Suppose not. Let $\rho_1(z) = \rho(z) - \rho(iz)$. Then $L_T\rho_1 = \rho_1$. But $\int_{\widetilde{B}} \rho_1 = 0$, so ρ_1 must be identically zero on $\widetilde{B}$. For mirror symmetry, let $\rho_2(z) = \rho(z) - \rho(\overline{z})$. Then

$$\begin{aligned} L_T\rho_2(z) &= \sum_{g \in G'} |g+z|^{-4} \left(\rho(1/(g+z)) - \rho(1/(\overline{g+z}))\right) \\ &= \rho(z) - L_T\rho(\overline{z}) = \rho(z) - \rho(\overline{z}) = \rho_2(z). \end{aligned}$$

As before, $\int_{\widetilde{B}} \rho_2 = 0$, so ρ_2 is the zero element of $\mathcal{V}$. This completes the proof of our Gauss-Kuz'min theorem.

References

[1] Bandtlow, O. and Jenkinson, O. Invariant measures for real analytic expanding maps (to appear).

[2] Hurwitz, A. (1888). Über die Entwicklung Complexer Grössen in Kettenbrüche, *Acta Math.* **11**, pp. 187-200.

[3] Hurwitz, J. (1895) Über eine besondere Art der Kettenbruch-entwicklung complexer Grössen, Thesis, Universität Halle-Wittenberg, Halle, Germany.

[4] Krasnoselskii, M. (1964). Positive solutions of operator equations. P. Noordhoff, Groningen.

[5] Lagarias, J. (1994). Geodesic multidimensional continued fractions, *Proc. London Math. Soc.* **3**, 69, pp. 464-488.

[6] Narasimhan, R. (1971). Several Complex Variables, University of Chicago Press, Chicago.

[7] Schmidt, A. (1975). Diophantine approximation of complex numbers. *Acta Math.* **134**, pp. 1-85.