The Riemann Mapping Theorem

Christopher J. Bishop

C.J. Bishop, Mathematics Department, SUNY at Stony Brook, Stony Brook, NY 11794-3651

E-mail address: bishop@math.sunysb.edu

1991 *Mathematics Subject Classification.* Primary: 30C35, Secondary: 30C85, 30C62

Key words and phrases. numerical conformal mappings, Schwarz-Christoffel formula, hyperbolic 3-manifolds, Sullivan's theorem, convex hulls, quasiconformal mappings, quasisymmetric mappings, medial axis, CRDT algorithm

The author is partially supported by NSF Grant DMS 04-05578.

Abstract. These are informal notes based on lectures I am giving in MAT 626 (Topics in Complex Analysis: the Riemann mapping theorem) during Fall 2008 at Stony Brook. We will start with brief introduction to conformal mapping focusing on the Schwarz-Christoffel formula and how to compute the unknown parameters. In later chapters we will fill in some of the details of results and proofs in geometric function theory and survey various numerical methods for computing conformal maps, including a method of my own using ideas from hyperbolic and computational geometry.

Contents

Preface 1

Chapter 1. Introduction to conformal mapping 3
1. Conformal and holomorphic maps 3
2. Möbius transformations 19
3. The Schwarz-Christoffel Formula 23
4. Crowding 30
5. Power series of Schwarz-Christoffel maps 32
6. Harmonic measure and Brownian motion 41
7. The quasiconformal distance between polygons 50
8. Schwarz-Christoffel iterations and Davis's method 58

Chapter 2. The Riemann mapping theorem 69
1. The hyperbolic metric 69
2. Schwarz's lemma 72
3. Square roots and logarithms 73
4. Morera's theorem and uniform convergence 74
5. Equicontinuity and compactness 75
6. The Poisson integral formula 76
7. A proof of Riemann's theorem 78
8. Koebe's method 79
9. Caratheodory's Theorem 85
10. Schwarz reflection 88
11. The maximum principle 89
12. Existence of Schwarz-Christoffel parameters 90
13. Maps to a rectangle 92

Chapter 3. Representing conformal maps 97

1. The Carleson decomposition 97
2. An expansion around the singularities 105
3. Gauss-Jacobi quadrature 114
4. The fast Fourier transform 127
5. Fast power series mainulations 129

Chapter 4. Some geometric function theory 137
1. Conformal modulus 137
2. Modulus and cross ratio; rectangles revisited 139
3. Pfluger's theorem and Beurling's estimate 141
4. Logarithmic capacity 144
5. Symmetry and Modulus 151
6. The distortion theorems 155
7. The Voronoi diagram 159
8. Convergence of Kakutani's method 161

Chapter 5. Quasiconformal Mappings 163
1. Compactness of K-quasiconformal maps 163
2. Quasi-isometries 165
3. Quasisymmetric maps 166
4. BiLipschitz maps 166
5. The Beltrami equation 169

Chapter 6. Schwarz-Christoffel iterations 171
1. The space of n-tuples 171
2. Davis's iteration 172
3. Newton's method 173
4. Broyden updates 173

Chapter 7. Tree-of-disk maps 175
1. The general set up 175
2. Triangulations 176
3. Delaunay triangulations and CRDT 183
4. The CRDT iteration 190
5. The medial axis 191

6. Formulas for the ι map 203
7. ι decreases length 212
8. Uniform bounds for tree-of-disk maps 213
9. The factorization theorem and Brennan's conjecture 219

Chapter 8. Domes and scaling 223
1. The dome of a domain 223
2. The Sullivan-Epstein-Marden theorem 228
3. The retraction map onto the dome 232
4. The gap-crescent decomposition for finitely bent domains 236
5. Angle scaling 239
6. Angle scaling is QC continuous 242
7. Angle scaling and Davis' method 246

Chapter 9. Linear methods 247
1. A linear algebra glossary 247
2. Iterative metods for linear systems 253
3. Symm's method 262
4. The Kerzman-Stein formula 274
5. The fast multipole method 280
6. Computing the Beurling transform 284

Chapter 10. The conjugation operator 293
1. Harmonic conjugates 293
2. Theodorsen's method 301
3. Fornberg's method 304
4. Wegman's method 308
5. Comparing wegman's and Fornberg's methods 309

Chapter 11. Higher dimensions 311
1. Liouville's theorem 311
2. Hamilton's theorem 311
3. Spectral geometry 311

Chapter 12. Higher conenctivity 313

1. The uniformization theorem 313
2. Koebe's theorem 313
3. Koebe's conjecture 313
4. Slit mappings 313

Chapter 13. Circle packings 315
1. Definitions 315
2. The Perron method 315
3. The hexagonal packing is rigid 315
4. Packing maps converge to conformal maps 315

Chapter 14. Conformal Welding 317
1. The fundamental theorem 317
2. Koebe's theorem and conformal welding 317
3. Marshall's Zipper algorithm 317
4. SLE 317

Chapter 15. The Schwarz-Christoffel formula (again) 321
1. Circular-arc polygons 321
2. Multiple connected domains 321
3. Black box solvers 321

Chapter 16. Conformal mapping in linear time 323
1. The idea 323
2. Thick and thin parts of a polygon 323
3. Arches 323
4. Building approximate bending laminations 323
5. Angle scaling is continuous 323
6. The algorithm 323

Chapter 17. Conformal maps and martigales 325
1. The Bloch space and Nehari's theorem 325
2. Bloch functions and Bloch martingales 325
3. Radial limits of conformal maps 325
4. Makarov's upper bound 325

5. The law of the iterated logarithm 325

Appendix A. Some domains used in the text 327

Appendix B. Some *Mathematica* code 333

Appendix C. Bits and pieces 345
1. Alternative definitions of quasiconformality 345
2. The Hardy-Littlewood maximal theorem 346
3. The distortion theorems 349
4. Extremal problems in geometric function theory 352
5. The strong law of large numbers 354

Appendix D. Background material 357
1. Real Analysis 357
2. Topology 357

Appendix. Bibliography 359

Preface

These are very rough and very incomplete notes related to Riemann mapping theorem, although my hope is to make them less rough and more complete as time goes on. They are written with three goals:

(1) give an introduction to the mapping theorem appropriate to advanced undergraduates and beginning graduate students,

(2) briefly describe various numerical schemes for computing conformal maps and give some of the required background to implement these methods (Gauss quadrature, FFTs, fast multipole, numerical linear algebra, ...),

(3) serve as an introduction to some of my own work relating hyperbolic and computational geometry to conformal mappings.

There are various excellent books on the Riemann mapping theorem, harmonic measure and geometric function theory, such as those of Pommerenke and Garnett-Marshall, and I have tried to stay away from topics covered in those sources. Mostly, these notes contain material which I learned in the last few years and which I probably should have known much earlier. Very little here is new, and some of it is very well known, but I hope there is some value in putting everything together.

As of this writing, the first three chapters form an introduction to the Riemann mapping theorem, the Schwarz-Christoffel formula and a few methods for computing conformal maps onto polygons. I have used this material for a semester long undergraduate seminar (where the last few weeks were devoted to students reseaching and presenting additional material). These chapters require a lot of polishing, but the content is more or less complete.

The next few chapters are intended to discuss iterative methods for computing the parameters in the Schwarz-Christoffel formula. I choose to measure the accuracy of the current guess using the quasiconformal distortion needed to map the guessed n-tuple to the correct one, so there will be a chapter introducing quasiconformal maps

(especially self-maps of the disk). Then a chapter on Davis's method and the CRDT method of Driscoll and Vavasis, followed (I hope) by a method of my own based on the medial axis of the domain.

Later chapters will discuss other methods dues to Symms, Fornberg, Wegman, Stein and Kerzman. These exist in rough form. Eventually I also hope to have descriptions of the circle packing approach (covered in greater detail by Stephenson's book) and Marshall's Zipper algorithm I would also like to have an introduction to conformal welding, the uniformization theorem and some higher dimensional results. However, none of this material is written yet, although I have optimistically included these headings in the table of contents.

I apologize for the numerous mistakes currently in the notes and its various other shortcomings (e.g., no references yet). Please feel free to inform me of particularly appalling oversights, such as not mentioning your obviously relevent work (or mentioning your work, but not your name). Other comments, suggestions and questions are also welcome.

Chris Bishop
Stony Brook, NY
January, 2010

CHAPTER 1

Introduction to conformal mapping

In this chapter we introduce conformal maps with an emphasis on the Schwarz-Christoffel formula. We discuss several ideas including Möbius transformations, conformal invariants, crowding, domain decompositions and quasiconformal maps which will be explored in greater depth in later chapters.

1. Conformal and holomorphic maps

A conformal map between planar domains is a C^1, orientation preserving diffeomorphism which preserves angles. Write $f(x, y) = (u(x, y), v(x, y))$. We can compute it derivative matrix

$$Df = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}.$$

Since f preserves orientation and angles, the linear map represented by this matrix must be an orientation preserving Euclidean similarity. Thus it is a composition of a dilation and rotation and must have the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

which implies

$$u_x = v_y, \quad u_y = -v_x.$$

These are known as the Cauchy-Riemann equations. Thus f is conformal if it is C^1 diffeomorphism which satisfies the Cauchy-Riemann equations.

The simplest examples are the Euclidean similarities, and indeed, these are the only examples if we want maps $\mathbb{R}^2 \to \mathbb{R}^2$. However, if we consider subdomains of $\mathbb{R}^2$, then there are many more examples. The celebrated Riemann mapping theorem says that any two simply connected planar domains (other that the whole plane) can be mapped to each other by a conformal map. We will give a more precise statement of this later and will eventually give a proof of the result, but for the present we introduce some notation and a few more examples.

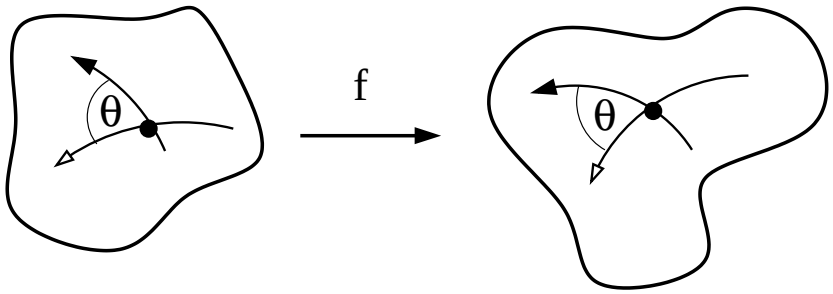


FIGURE 1. A conformal map preserves angles (and orientation).

After the linear maps, the next easiest conformal maps are quadratic polynomials. If we take

$$f(x,y) = (u(x,y), v(x,y)) = (x^2 - y^2, 2xy),$$

then we can easily check that

$$Df(x,y) = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix},$$

so the Cauchy-Riemann equations are satisfied. The map is not conformal on the plane since $f(-x,-y) = f(x,y)$ is 2-to-1 for $(x,y) \neq (0,0)$ and Df vanishes at the origin. However, it is a conformal map if we restrict it to a domain (an open, connected set) where it is 1-to-1, such as the open square $[0,1]^2$. The map sends this square conformally to a region in the upper half-plane. See Figure 5. Note that angles are doubled at the origin; we do not require that a conformal mapping of a domain preserve angles at boundary points and this map does not.

By this point, anyone who has had a course in complex analysis will have recognized the map f as complex squaring. We identity $\mathbb{R}^2$ with the complex numbers $\mathbb{C}$ by writing a real 2-vector (x,y) as a single complex number $z = x + iy$. The complex numbers form a field under the usual addition

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

and multiplication defined using the relation $i^2 = -1$ as follows

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1 x_2 + i x_1 y_2 + i x_2 y_1 + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \end{aligned}$$

Complex squaring is then

$$z^2 = (x+iy)^2 = (x^2 - y^2) + i2xy,$$

which is the map described earlier.

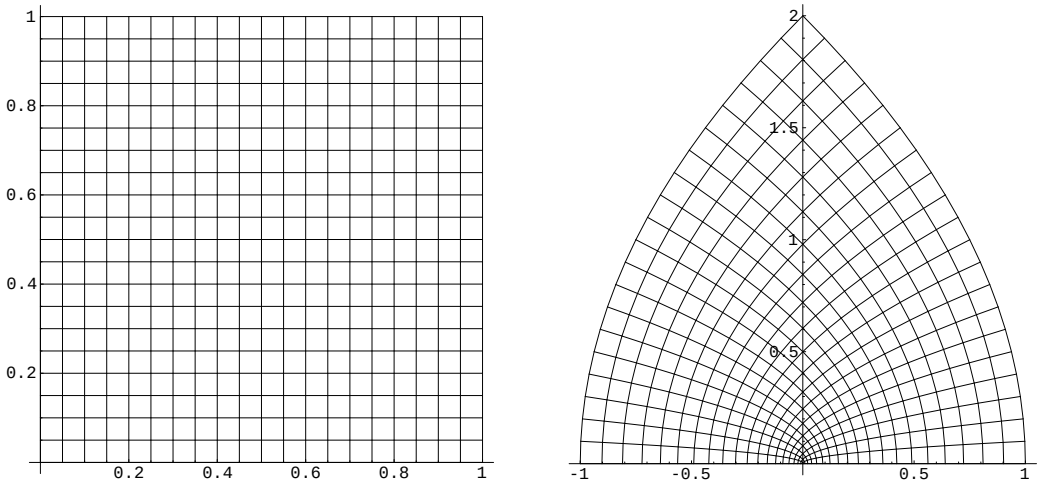


FIGURE 2. This illustrates the map $z \to z^2$ or $(x,y) \to (x^2-y^2, 2xy)$. The top left shows a grid in the square $[0,1]^2$. The top right shows the image under squaring map.

Complex multiplication is easier to understand in polar coordinates. Let $r = |z| = \sqrt{x^2+y^2}$ denote the distance from z to the origin and let $\theta = \arg(z)$ be the angle so that $x = r\cos\theta, y = r\sin\theta$. Note that if θ is a possible value of $\arg(z)$, then so is $\theta + 2\pi n$ for any integer n. In order to make $\arg(z)$ a function, we need to restrict to a single value, so we often choose $\theta \in (-\pi, \pi]$. This is the principal branch of arg and is denoted Arg(z). . Note that it has a jump discontinuity along the negative real axis. It is often convenient to choose other branches of arg which have discontinuities along a different ray, or possibly a curve connecting 0 to ∞. Given any simply connected domain $\Omega \subset \mathbb{C}$ which does not contain 0, we can always choose a continuous branch of $\arg(z)$ that is defined in Ω.

LEMMA 1. *Suppose Ω is a simply connected plane domain which does not contain the origin. Then there is a continous branch of* $\arg(z)$ *defined on Ω, i.e., there is a continuous function $f(z)$ so that* $\exp(\log|z| + if(z)) = z$.

PROOF. This "proof" is simply a reference to a standard result about topology. Consider $e^{it} : \mathbb{R} \to \mathbb{T}$ as a covering map. Note that $g(z) = e^{i \arg(z)}$ is a continuous map $\Omega \to \mathbb{T}$. Since Ω is simply connected, there is a lifting of g to a map $f : \Omega \to \mathbb{R}$, i.e., a map so that $g(z) = e^{if(z)}$. Thus f is the desired branch. See Chapter 5 for more about covering maps and some standard references. □

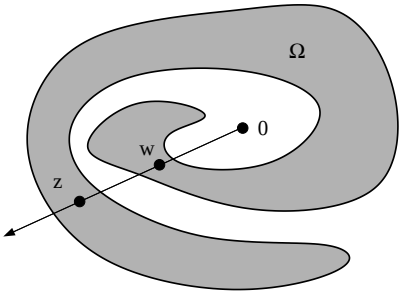


FIGURE 3. The points z and w lie on the same ray through the origin, but a continuous branch of arg on Ω will give z a value 2π larger than the value for w.

Define

$$e^z = e^{x+iy} = e^x(\cos y + i \sin y)$$

and

$$\log(z) = \log|z| + i \arg(z), \text{ if } z \neq 0.$$

The exponential functions satisfies the Cauchy-Riemann equations and the partials are never zero, so this function is conformal on any domain where it is 1-to-1. It is not 1-to-1 on the whole plane because $e^{z+2\pi i} = e^z$; each point except the origin has infinitely many preimages arranged along a vertical line. Each vertical line is mapped to a circle centered at the origin and teach horizontal line is mapped to a ray from 0 to ∞. See Figure 5. The logarithm is a branch of the inverse of this map; it sends rays to horizontal lines and circular arcs centered at the orgin to vertical lines.

A complex number z can be written as $z = re^{\imath\theta}$ where $r = |z|$ and $\theta = \arg(z)$. This is the polar coordinates form of a complex number. When we multiply two complex numbers the absolute values multiply and the arguments add, i.e.,

$$z_1 z_2 = (r_1 e^{i\theta_1})(r_2 e^{i\theta_2}) = r_1 r_2 e^{i(\theta_1+\theta_2)}.$$

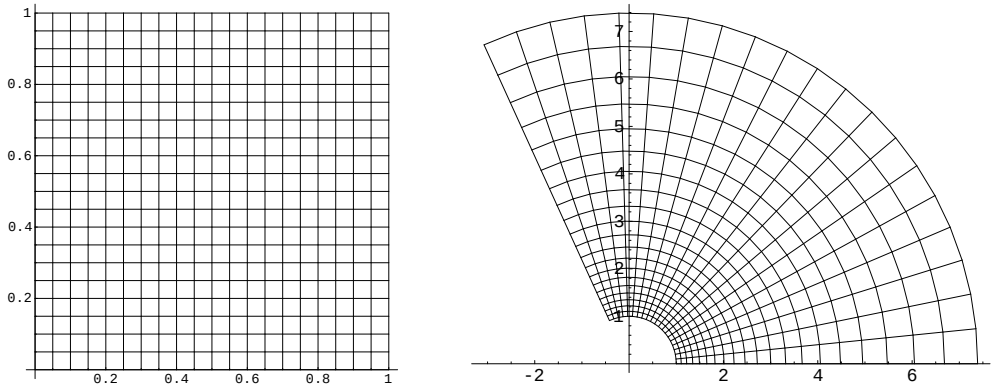


FIGURE 4. The same square grid of $[0, 2]^2$ and its image under e^z.

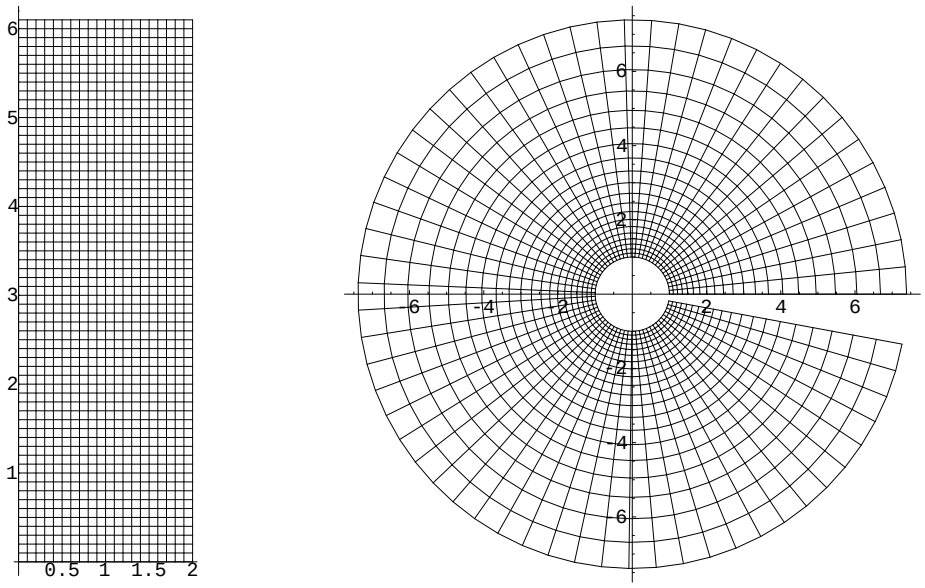


FIGURE 5. This illustrates the exponential map again. We take the image of $[0, 2] \times [0, 6]$. The line at height 2π will be mapped into the positive real axis. The top edge of the grid is just below this, so the image stops just before it reaches the axis.

This explains why the angle doubles at the origin in Figure 5. If we consider the maps z^3 and $z^{1/2}$, then angles at the origin will multiply by 3 and $\frac{1}{2}$ respectively, as

shown in Figure 5. In general we define

$$z^\alpha = e^{\alpha \log z} = e^{\alpha(\log|z| + i\arg(z))} e^{\alpha(\log|z| + i\mathrm{Arg}(z) + 2\pi i\mathbb{Z})}.$$

If α is an integer then this the various possible valuesof $\arg(\alpha)$ all give the same value of z^α. If $\alpha = p/q$ then there are q possible different values. Otherwise, z^α has infinitely many possible values. Moreover, some caution is needed when applying the rules of exponents. Consider

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i \cdot i = -1.$$

The problem is that $\sqrt{1}$ and $\sqrt{-1}$ each have two possible values and by choosing the wrong we can arrive at an apparent contradiction.

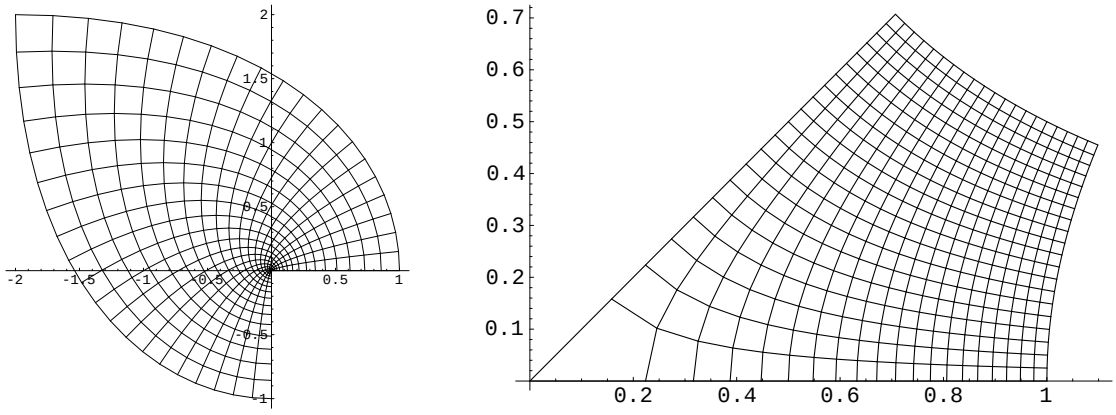


FIGURE 6. The images of $[0,1]^2$ under z^3 and $z^{1/2}$. These are all conformal maps of the square, but are not conformal at the origin (which is a boundary point).

A complex function of a complex variable is differentiable if

$$f'(z) = \lim_{h\to 0} \frac{f(z+h) - f(z)}{h},$$

exists. Here h can approach the origin in any way whatsoever. Two special ways of approaching are along the real or imaginary axes, which lead to the equations

$$f'(z) = \lim_{h\to 0} \frac{u(x+h,y) + iv(x+h,y) - u(x,y) - iv(x,y)}{h} = u_x + iv_x,$$

$$f'(z) = \lim_{h\to 0} \frac{u(x,y+h) + iv(x,y+h) - u(x,y) - iv(x,y)}{ih} = -iu_x + v_x.$$

Cauchy-Riemann equations. Conversely, is these partials exist, are continuous in a neighborhood of z and satisfy the Cauchy-Riemann equations, then f' exists and

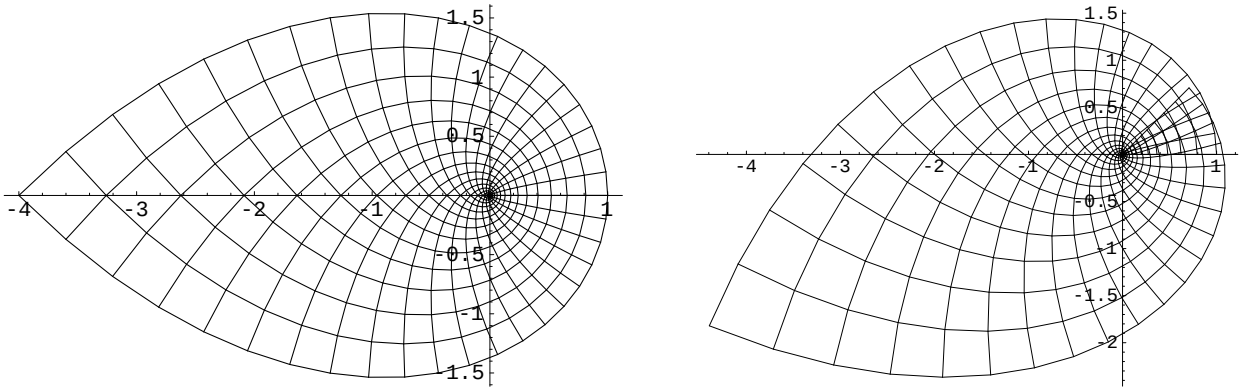


FIGURE 7. The images of $[0,1]^2$ under z^4 and $z^{4.5}$. In the first the segment $[0,16]$ is part of the boundary, not the interior of the dommain. For powers > 4, the map is no longer 1-to-1 and the image intersects itself.

equals $u_x + iv_x$. Continuity is required because examples like $xy/(x^2+y^2)$ show that a function can have partial derivatives at 0, but not even be continuous there. But if the partials exist and are continuous in neighborhood of a point, then results from calculus imply it is approximated by the linear map Df, i.e., if $h = s + it$, then

$$f(z+h) - f(z) = (u_x s + u_y t) + i(v_x s + v_y t) + o(|h|) = (u_x + iv_x)h + o(|h|), \tag{1}$$

which implies f is differentiable with derivative $u_x + iv_x$.

This is the first time we have used the "little-oh" notation, so perhaps we should explain it. The term $o(|h|)$ refers to term which is going to zero faster than $|h|$ as $|h| \to 0$. Equation (1) means that for every $\epsilon > 0$ there is a $\delta > 0$ so that if $|h| \le \delta$, then

$$|f(z+h) - f(z) - (u_x + iv_x) \cdot h| \le \epsilon h.$$

However, it is quicker and more convenient to write (1). Note that $o(1)$ stands for a term that tends to zero as the relevant parameter tends to its limit. Thus $o(h) = o(1) \cdot h$. The "big-Oh" notation $O(1)$ stands for a term that remains bounded as the relevant parametet tends to its limit. For example, $O(|x|)$ as $|x| \to \infty$ stands for a term that is bounded by $C|x|$ fore some fixed $C < \infty$ as $|x|$ grows.

The reader should now check that $(e^z)' = e^z$ and $(\log z)' = 1/z$ (on an region where log is defined and continuous). The usual rules of differentiation hold:

$$(f+g)' = f' + g', (fg)' = f'g + fg', (f/g)' = (f'g - fg')/g^2, (f \circ g)' = (f' \circ g)g'.$$

which implies the complex derivative exists. Therefore polynomials and rational functions are differentiable (at least at points where we don't divide by zero).

A power series is an infinite series of the form

$$\sum_{n=0}^{\infty} a_n(z-z_0)^n.$$

we say it converges at z if the sequence of partial sums has a finite limit, i.e.,

$$f(z) = \lim_{N\to\infty} \sum_{n=0}^{N} a_n(z-z_0)^n,$$

exists. Such a series obviously converges at $z = z_0$. More generally, there is a radius of convergence R ($R = 0, \infty$ are possible) where

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n},$$

and the series converges inside $\{z : |z - z_0| < R\}$ and diverges in $\{z : |z - z_0| > R\}$. The series might or might not converge at various boundary points of the disk; this depends on the particular coefficients. When a power series converges, it defines a continuous function on the open disk of convergence and this function is complex differentiable. If it is also 1-1, then it is a conformal map of the disk. More surprisingly, if f is a conformal map of an open disk, then f has a power series converging to it in this disk. We will prove this in the next few pages, but first need to introduce complex integrals and the Cauchy integral theorem.

A curve is continuous map $\gamma : [a, b] \to \mathbb{C}$. We also call the compact set $\Gamma = \gamma([a, b])$ a curve, although technically this should be called the trace of γ. A curve γ is rectifiable if there is a $M < \infty$ so that

$$\sup_{\mathcal{P}} \sum_{k=1}^{n} |\gamma(x_{k-1}) - \gamma(x_k)| \le M,$$

for every finite ordered set $\mathcal{P} = \{a = x_1 < \cdots < x_{n+1} = b\} \subset [a, b]$. The smallest such upper bound is the length of γ, denoted $\ell(\gamma)$. $\mathcal{P}$ is called a parition of $[a, b]$ and $\|\mathcal{P}\| = \max_k |x_{k+1} - x_k|$ denotes the size of the largest gap between consequtive points.

Felix Klein was quoted in [] as saying "Everyone knows what a curve is, until he has studied enough mathematics to become confused through the countless number of possible exceptions." Figures 5 and 5 show two such possible exceptions.

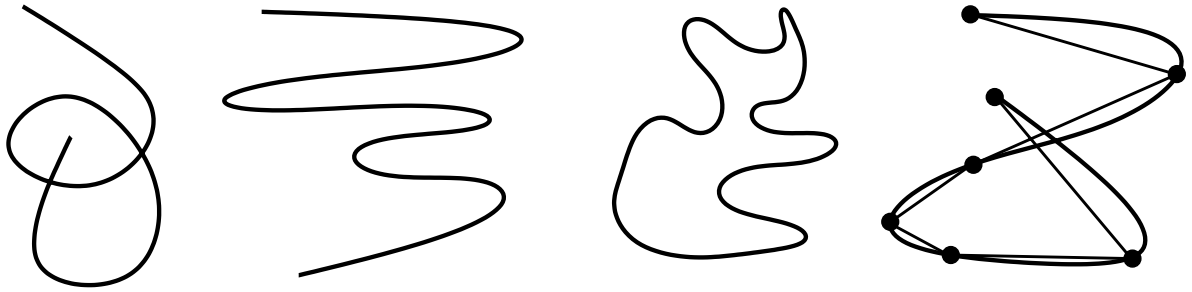


FIGURE 8. A curve is continuous image of a closed interval. A Jordan curve is a 1-to-1 image (no self-intersecting), and a closed Jordan curve has $\gamma(b) = \gamma(a)$. A rectifiable curve is one where inscribed polygons have uniformly bounded length.

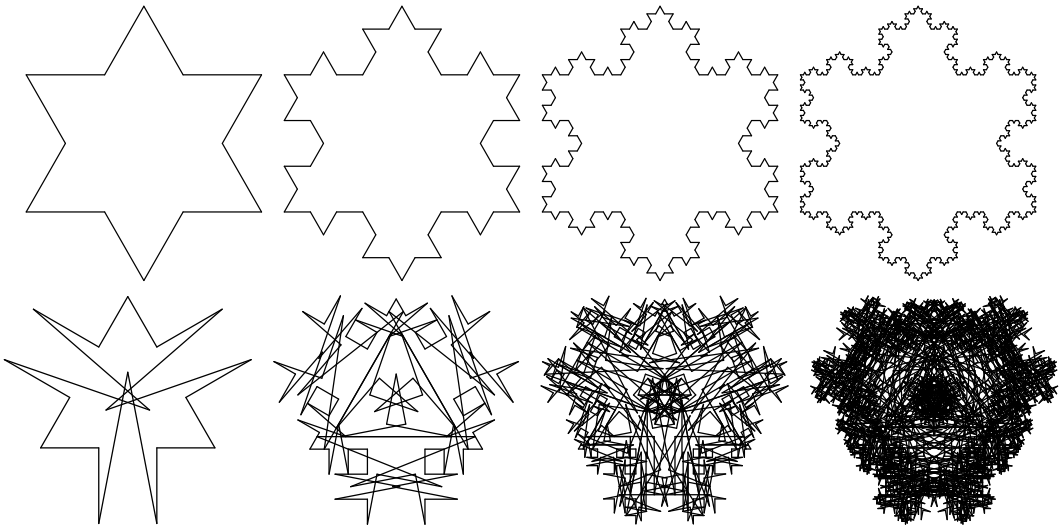


FIGURE 9. The top row shows four generations of the construction of the von Koch snowflake, a closed Jordan curve that is not rectifiable. The bottom row shows four generations of a variation of the snowflake. In this case the limiting curve covers an open set, i.e., is a type of Peano curve.

If γ maps into a domain Ω and f is a continuous function on Ω, we define the integral of f along γ as

$$\int_\gamma f(z)dz = \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} f(\gamma(x_k))(\gamma(x_{k+1}) - \gamma(x_k)),$$

where the limit is taken over paritions as the maximum gap tends to zero. If f is continuous and γ is rectifiable, then it is easy to see that this limit exists. If γ is

piecewise C^1, we sometimes write $\int_\Gamma f(z)dz$, instead of $\int_\gamma f(z)dz$. This is permissible since one can show that two different parameterizations of Γ (with the same starting and ending points) give the same integral. The following estimate is obvious, but very useful:

LEMMA 2. $\int_\gamma f(z)dz \leq \max_\gamma |f| \cdot \ell(\gamma)$.

If γ is differentiable, this definitions agrees with the idea of a line integral in calculus:

LEMMA 3. *If γ is a C^1 curve then*

$$\int_\gamma f(z)dz = \int_a^b f(\gamma(x))\gamma'(x)dx,$$

(where γ' is interpreted as a complex number instead of a 2-vector).

PROOF. The two integrals are limits of Riemann sums of the form

$$\sum f(z_k)(z_{k+1} - z_k), \quad \sum f(z_k)\gamma(x_k)(x_{k+1} - x_k),$$

and since $(z_{k+1} - z_k) = (x_{k+1} - x_k)\gamma'(x_k) + o(1)$, the result follows. □

An important consequence is the following equality: if $\gamma(t) = re^{it}$ maps $[0, 2\pi]$ onto a circle of radius r, then

$$\int_\gamma \frac{1}{z} dz = 2\pi i.$$

To see this, we simply compute the left hand side as

$$\int_0^{2\pi} \frac{1}{r} e^{-it} rie^{it} dt = i \int_0^{2\pi} dt = 2\pi i.$$

LEMMA 4. *Suppose f is continuous and $f = F'$ for some complex differentiable F on Ω. Then we claim that*

$$\int_\gamma f(z)dz = F(\gamma(b)) - F(\gamma(b)).$$

PROOF. To prove this consider a partition $\mathcal{P} = \{x_1 < \dots x_{n+1}\}$ and let $z_k = \gamma(x_k)$. Since γ has compact image and f is continuous, it is uniformly continuous on $\Gamma = \gamma([a, b])$. Hence

$$|f(z) - f(w))| \leq o(1)$$

unifromly as $\|\mathcal{P}\| \to 0$ and thus

$$F(z) = F(w) + (z-w)f(w) + o(|z-w|),$$

for every $w \in \Gamma$. Since γ is uniformly continuous, if the gaps in our partition are small enough then $\gamma([x_{k+1}mx_k]) \subset D(z_k, r)$. Thus

$$\begin{aligned}\int_\gamma f(z)dz &= \sum_{k=1}^{n} f(\gamma(x_k))(z_{k+1}-z_k) + o(1)\\ &= \sum_{k=1}^{n}[F(z_{k+1})-F(z_k)] + o(\sum_k |z_{k+1}-z_k|) + o(1)\\ &= F(z_{n+1}) - F(z_1) + \o(\ell(\gamma)) + o(1),\end{aligned}$$

which gives the desired equality. In particular if γ is a closed curve, i.e., $\gamma(a) = \gamma(b)$, then $\int_\gamma f(z)dz = 0$. □

LEMMA 5. *If f is conformal on a domain Ω and γ is a closed rectifiable curve in $D(z,r) \subset \Omega$ then $\int_\gamma f(z)dz = o(r\ell(\gamma))$.*

PROOF. Since constants and linear functions of z are derivatives of other functions, the integral of one of these around a closed curve is zero. So if $\gamma \subset D(z_0, r)$ and

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + o(|z-z_0|),$$

then

$$\begin{aligned}\int_\gamma f(z)dz &= f(z_0)\int_\gamma dz + f'(z_0)\int_\gamma (z-z_0)dz + \int_\gamma o(|z-z_0|)dz\\ &= 0 + 0 + o(r)\ell(\gamma).\end{aligned}$$

□

We say a domain Ω is decomposed in subdomians $\{\Omega\}$ if each $\Omega_k \subset \Omega$, they are pairwise disjoint and $\Omega = \cup_k \overline{\Omega_k} \cap \Omega$. Right now, we are only interested in decompositions of peicewise C^1 domains into finitely many piecewise C^1 subdomains, as illustrated in Figure 10. Note that if f is continuous on the closure of Ω, then

$$\int_{\partial\Omega} f(z)dz = \sum \int_{\partial\Omega_k} f(z)dz,$$

becuase each arc of $\partial\Omega_k$ which is interior to Ω is also a boundary arc of another domain, but with the opposit orientation. Thus these integrals cancel. The only

parts of the integrals on the right that don't cancel are the one on the boundary of Ω, and they sum to the integral on the left. Technically, we should write $\int_\gamma$ instead of $\int_{\partial\Omega}$ where γ is a parameterization of $\partial\Omega$, but the integral is independent of the particular parameteriation, so this abuse of notation is reasonable.

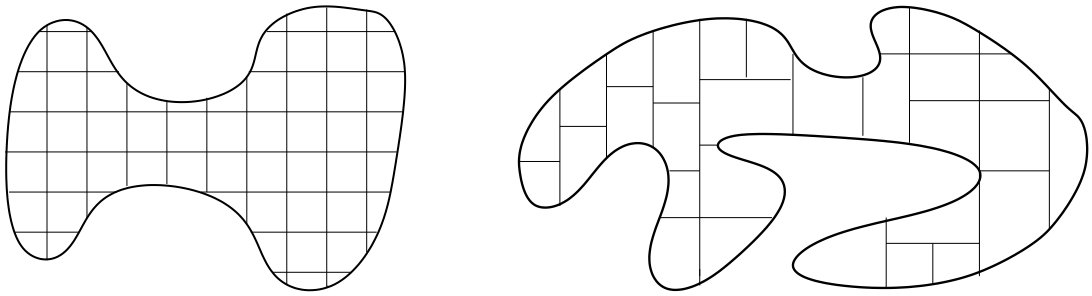


FIGURE 10. Decomposing a domain means breaking it into smaller domains. We will only be interested in decompositions of piecewise smooth domains into finitely many piecewise smooth domains with small diameter. For a piecewise C^1 curve this can be done by intersecting with a standard ϵ-grid, although more irregular decompositions are also allowed.

LEMMA 6. *Suppose $\gamma \subset \Omega$ is a closed Jordan curve which bounds a region which can be decomposed into $O(\epsilon^{-2})$ many regions each with boundary length $\leq \epsilon$. Suppose f has a continuous complex derivative on Ω. Then $\int_\gamma f(z)dz = 0$.*

PROOF. The integral around the boundary of each subpiece is $o(\epsilon^2)$ and there are $O(\epsilon^{-2})$. Thus the sum of these integrals tends to zero as $\epsilon \to 0$. The sum is always equal to $\int_\gamma f(z)dz$, so the lemma is proved. □

LEMMA 7. *Suppose f is holomorphic on a neighborhood of a closed disk $D = D(0,r)$ and $|z| < r$. Then*

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-z}dw,$$

where $\gamma(t) = z_0 + re^{it}$ maps $[0, 2\pi]$ to the boundary of D

PROOF. Choose ϵ so $0 < \epsilon < (r - |z|)/2$ and let S be a line segment connecting the circles $C_\epsilon = \{w : |z-w| = \epsilon\}$ to $C_r = \{w : |w| = r\}$. Let γ be the curve consisting of four parts: traversing C_r one time in the counterclockwise direction, along S from

C_r to $C\epsilon$, around C_ϵ once clockwise and finally along S from C_ϵ to C_r. This curve bounds a slit annulus and can clear be decomposed as in Lemma 6, so the integral of f around γ is 0. The integrals over the two line segments cancel, so the integrals over the inner and outer circle also cancel. We rewrite the inner integral as

$$\int \frac{f(w)}{w-z}dw = \int \frac{f(z)-f(0)}{w-z}dw + \int f(0)\frac{dw}{w-z}.$$

By an earlier computation the second integral on the right is $2\pi i f(0)$. The first is bounded by $\sup_{|w|=\epsilon}|f(w)-f(0)|$ which tends to zero as $\epsilon \to 0$. This proves the lemma. □

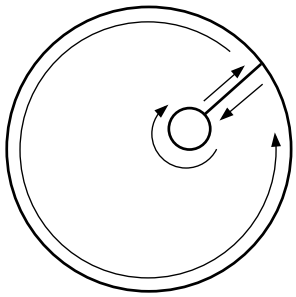


FIGURE 11. The curve used in the proof of Lemma 7.

If we differentiate both sides of the Cauchy integral formula we get

$$f'(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{(w-z)^2}dw.$$

Pulling the differentiation inside the integral is justified because

$$\begin{aligned}\lim_{h\to 0}\frac{f(z+h)-f(z)}{h} &= \frac{1}{2\pi i}\lim_{h\to 0}\int_\gamma f(w)\frac{1}{h}[\frac{1}{w-z-h}-\frac{1}{w-z}dw \\ &= \frac{1}{2\pi i}\int_\gamma f(w)\lim_{h\to 0}\frac{1}{(w-z)(w-z-h)}dw \\ &= \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{(w-z)^2}dw\end{aligned}$$

and the convergence is uniform on γ (and hence the integral of the limit is the limit of the integrals). This formula for f' implies

COROLLARY 8 (The Cauchy Estimate). *If f is holomorphic on $D(z,r)$ and $|f|$ is bounded by M on the disk, then $|f'(z)| \le M/r$.*

This will be useful in Chapter 5 when we want to show that holomorphic maps into the unit disk form an equicontinuous family. Similar estimates can be proven for higher derivatives of f. One of the most important consequences is:

THEOREM 9 (Liouville's Theorem). *If f is bounded and holomorphic on the whole plane then f is constant.*

PROOF. By Cauchy's formula

$$|f'(z)| = |\int_0^{2\pi} f(z + Re^{i\theta}) \frac{iRe^{i\theta}d\theta}{(Re^{i\theta})^2}| \leq \max_{\mathbb{C}} |f| \frac{2\pi}{R} \to 0,$$

as $R \to \infty$. Thus $f' = 0$ everywhere. □

In particular, there cannot be any conformal map of the whole plane to the unit disk, for this would define a non-constant, bounded holomorphic function on the plane. Thus when Riemann's theorem states that every simply connected region, **except the plane**, can be conformally mapped to the disk, it is making the sharpest possible claim.

Next we want to compute power series for holomorphic functions. The most important example is the geometric series

$$\sum_{n=0}^{\infty} z^n = 1 + z + z^2 + z^3 + \cdots = \frac{1}{1-z}.$$

The derivation is exactly the same as for real numbers in calculus. Suppose

$$1 + z + \cdots + z^n = S.$$

Then

$$z + z^2 + \ldots z^{n+1} = zS,$$

so subtracting gives

$$1 - z^{n+1} = S - zS = S(1 - z),$$

$$S = \frac{1 - z^{n+1}}{1 - z},$$

$$1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}. \tag{2}$$

For $|z| < 1$ we have $z^{n+1} \to 0$, which proves the result. Also note that if we differentiate (2) we get

$$0 + 1 + 2z + \cdots + nz^{n-1} = \frac{-(n+1)z^n(1-z) + (1-z^{n+1})}{(1-z)^2}. \tag{3}$$

For $|z| < 1$ the terms z^n and z^{n+1} tend to zero, so

$$\sum_{n=1}^{\infty} nz^{n-1} = 1 + 2z + 3z^2 + 4z^3 + \cdots = \frac{1}{(1-z)^2}.$$

The closely related function

$$\frac{z}{(1-z)^2} = z + 2z^2 + 3z^3 + \ldots,$$

is called the Koebe function and has an important place in the history of geometric function theory.

LEMMA 10. *If f is holomorphic on a neighborhood of the closure of $D = D(0,r)$ then f has a power series centered at 0 that converges in D.*

PROOF. Let γ be the curve that traverses ∂D once in the counterclockwise direction. Let $z \in D$. Then

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-z}dw \\ &= \frac{1}{2\pi i}\int_\gamma f(w)\frac{1}{w}[\sum_{n=0}^{\infty}(\frac{z}{w})^n]dw \\ &= \sum_{n=0}^{\infty}[\frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w^{n+1}}dw]z^n \\ &= \sum_{n=0}^{\infty} a_n z^n. \end{aligned}$$

The infinite sum and integral can be exchanged because the sum is absolutely and unifromly convergent for $|z| < |w|$ (see Appendix D. Note that $|a_n| \le \max_\gamma |f| \cdot r^{-n-1}$, so the radius of convergence is $\ge r$. □

LEMMA 11. *If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is defined by a convergent power series in $D(0,r)$, then f is holomorphic and $f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}$ in $D(,r)$*

PROOF. We compute the derivative by taking quotients

$$\begin{aligned} f'(z) &= \lim_{h\to 0}\frac{1}{h}[f(z+h)-f(z)] \\ &= \lim_{h\to 0}\frac{1}{2\pi i h}\int_\gamma \frac{f(w)}{w-z}-\frac{f(w)}{w-z-h}dw \\ &= \lim_{h\to 0}\frac{-1}{2\pi i}\int_\gamma \frac{f(w)}{(w-z)(w-z-h)}dw \\ &= \frac{-1}{2\pi i}\int_\gamma \frac{f(w)}{(w-z)^2}dw, \end{aligned}$$

where we have used uniform convergence to justify passing the limit through the integral. Next use (3).

$$\begin{aligned} f'(z) &= \frac{1}{2\pi i}\int_\gamma f(w)\frac{1}{w}[\sum_{n=1}^{\infty} n(\frac{z}{w})^{n-1}]dw \\ &= \sum_{n=1}^{\infty} n[\frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w^{n+1}}dw]z^{n-1} \\ &= \sum_{n=1}^{\infty} n a_n z^{n-1}. \end{aligned}$$

□

From the power series formula one can derive $a_n = \frac{f^{(n)}(0)}{n}$, where $f^{(n)}$ denotes the nth derivative of f and $n = 1\cdot 2\cdot 3\cdots n$. Some important examples are

$$e^z = \sum_{n=1}^{\infty}\frac{1}{n}z^n,$$

$$(1+z)^\alpha = \sum_{n=0}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n}z^n.$$

Thus every holomorphic function on the unit disk has a power series expansion and hence every conformal map does. While its easy to determine which power series correspond to holomorphic functions ($\limsup |a_n|^{1/2} \geq 1$) it is probably impossible to give a concise characterization of the series corresponding to 1-to-1 holomorphic functions (e.g., conformal maps). One of the most famous problems in complex analysis was the Beierbach conjecture that if f is 1-1 and holomorphic on $\mathbb{D}$ with $|f'(0)| = 1$, then

$$|a_n| \leq n.$$

Sharpness is shown by the Koebe function mentioned earlier. The conjecture was proven in 1984 by Louis deBrange in a technical tour-de-force, later simplified by other authors. Similar questions for related collection of maps still remain open.

2. Möbius transformations

A linear fractional transformation (or Möbius transformation) is a map of the form $z \to (az+b)/(cz+d)$. This is a 1-1, onto, holomorphic map of the Riemann sphere $\mathbb{S} = \mathbb{C} \cup \{\infty\}$ to itself. The non-identity Möbius transformations are divided into three classes. Parabolic transformations have a single fixed point on $\mathbb{S}$ and are conjugate to the translation map $z \to z+1$. Elliptic maps have two fixed points and are conjugate to the rotation $z \to e^{it}z$ for some $t \in \mathbb{R}$. The loxodromic transformations also have two fixed points and are conjugate to $z \to \lambda z$ for some $|\lambda| < 1$. If, in addition, λ is real, then the map is called hyperbolic.

Given two sets of three distinct points $\{z_1, z_2, z_3\}$ and $\{w_1, w_2, w_3\}$ there is a unique Möbius transformation that sends $w_k \to z_k$ for $k = 1, 2, 3$. This map is given by the formula

$$\tau(z) = \frac{w_1 - \zeta w_3}{1 - \zeta},$$

where

$$\zeta = \frac{(w_2 - w_1)}{(w_2 - w_3)} \frac{(z - z_1)(z_2 - z_3)}{(z - z_3)(z_2 - z_1)}.$$

A Möbius transformation sends the unit disk 1-1, onto itself iff it is if the form

$$z \to \lambda \frac{z - a}{1 - \bar{a}z},$$

for some $a \in \mathbb{D}$ and $|\lambda| = 1$. In this case, any loxodromic transformation must actually be hyperbolic.

Given four distinct points a, b, c, d in the plane we define their cross ratio as

$$\mathrm{cr}(a, b, c, d) = \frac{(d - a)(b - c)}{(c - d)(a - b)}.$$

Note that $\mathrm{cr}(a, b, c, z)$ is the unique Möbius transformation which sends a to 0, b to 1 and c to ∞. This makes it clear that cross ratios are invariant under Möbius transformations; that $\mathrm{cr}(a, b, c, d)$ is real valued iff the four points lie on a circle; and is negative iff in addition the points are labeled in counterclockwise order on the circle.

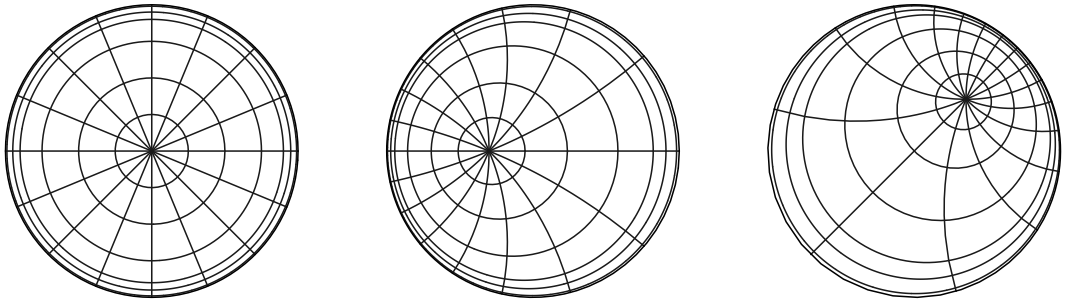


FIGURE 12. A polar grid in the disk and some images under Möbius transformations that preserve the unit disk.

Möbius transformations form a group under composition. If we identity the transformation $(az+b)/(cz+d)$ with the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then composition of maps is the same as matrix multiplication. For any non-zero λ, the translations $(\lambda az+\lambda b)/(\lambda cz+\lambda d)$ are all the same, but correspond to different matrices. We can choose one to represent the transformation, say the one with determinate 1, and this identifies the group of transformations the the group $\mathrm{SL}(2,\mathbb{C})$ of two by two matrices of determinate 1. (If $ad=bc$, then

$$\frac{az+b}{cz+d} = \frac{adz+bd}{cdz+d^2} = \frac{bcz+bd}{cdz+d^2} = \frac{b}{d}\frac{cz+d}{cz+d} = \frac{b}{d},$$

is constant and not a Möbius transformation.

The mapping

$$z \to \frac{az+b}{cz+d},$$

can be written as a composition of the maps

$$z \to cz+d,$$

$$z \to \frac{1}{z},$$

$$z \to \frac{a}{c} + \frac{bc-ad}{c}z,$$

which equivalent to claiming

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} (bc-ad) & a \\ 0 & c \end{pmatrix}.$$

Either claim follows by a direct computation. The linear maps have the property that circles map to circles an lines map to lines. The inversion also has this property, although it may interchange the two types of sets. The equation

$$x^2 + y^2 + \alpha x + \beta y + \gamma = 0 \tag{4}$$

defines a circle in the plane, depending on the choice of α, β, γ. If we set $z = x+iy \neq 0$ and $\frac{1}{z} = u + iv$, then

$$u = \Re(\frac{x - iy}{x^2 + y^2}) = \frac{x}{x^2 + y^2},$$
$$v = \Im(\frac{x - iy}{x^2 + y^2}) = \frac{-y}{x^2 + y^2},$$
$$x = \frac{u}{u^2 + v^2},$$
$$y = \frac{-v}{u^2 + v^2},$$

so (4) becomes

$$\frac{u^2}{(u^2 + v^2)^2} + \frac{v^2}{(u^2 + v^2)^2} + \frac{\alpha u}{(u^2 + v^2)^2} + \frac{-\beta v}{(u^2 + v^2)^2} + \gamma = 0.$$

After simplifying this becomes

$$\frac{1}{(u^2 + v^2)^2} + \frac{\alpha u}{u^2 + v^2} + \frac{-\beta v}{u^2 + v^2} + \gamma = 0,$$

$$1 + \alpha u - \beta v + \gamma(u^2 + v^2) = 0,$$

which is the equation of a circle or line (depending on whether $\gamma \neq 0$ or $\gamma = 0$). Thus $z \to \frac{1}{/}z$ sends a circle not passing through the origin to a circle and a circle that does pass though 0 to a line (which is the same as a circle passing through ∞). Thus we have shown

LEMMA 12. *Möbius transformations map circles to circles, assuming the convention that lines are considered as circles through infinity.*

The reflection through a circle $|z - c| = r$ is defined by $\arg(w^* - c) = \arg(w - c)$ and $|w - c| \cdot |w^* - c| = r^2$. Möbius transformation preserve reflections, i.e., if τ is a linear fractional transformation that send circle (or line) C_1 to circle (or line) C_2 then pairs of symmetric points for C_1 are mapped by τ to symmetric points for C_2.

LEMMA 13. *Every Möbius transformation can be written as a even number of compositions of circle and line reflections.*

The proof is left to the reader.

In higher dimensions, reflections through planes and spheres still makes sense. In this case, Möbius transformations are defined as the group generated by any even number of compositions of such maps (even so that the result is orientation preserving).

Each reflection in a line can be extended to a reflection across a plane in 3-space that is perpendicular to $\mathbb{R}^2$. Similarly, any circle reflection in the plane can be extended to a reflection through a sphere in 3-space. From this it is possible to show that every Möbius transformation has a unique extension to a conformal map of $\mathbb{S}^2$ to itself.

The plane can be identified with a 2-sphere minus a point via stereographic projection. Möius transformations can be considered as mappings of the sphere to itself. To be more concrete, we consider the unit sphere $S^2 = \{(x, y, zz) : x^2 + y^2 + z^2 = 1\}$ in $\mathbb{R}^3$ and let $N = (0, 0, 1)$ denote the "north pole". Then $S^2 \setminus \{N\}$ is topolgocially a plane and the correspondence can be made explicit by joining each point of the (x, y)-plane to N by a straight line in $\mathbb{R}^3$. This line hits S^2 at some point $(u, v, t) \neq N$ and the map $(x, y) \to (u, v, t)$ is called the stereographic projection of the plane onto a sphere. We can easily compute formulas for this map. Set $r = \sqrt{x^2 + y^2}$ and $\rho = \sqrt{u^2 + v^2}$. Then we have $t^2 + \rho^2 = 1$ and $(r - \rho)/r = t$. Solving for t gives

$$t = \frac{r - 1}{r + 1},$$

which implies

$$u = \frac{x}{r\rho} = \frac{x}{r\sqrt{1 - t^2}},$$
$$v = \frac{y}{r\rho} = \frac{y}{r\sqrt{1 - t^2}}.$$

We leave it to the reader to check that circle or lines in the (x, y)-plane map to circles on S^2.

In 3 dimensions and higher, these are the only conformal maps. By a theorem of Liouville, any conformal map from a domain $\Omega \subset \mathbb{R}^3$ into $\Omega' \subset \mathbb{R}^2$ must the restriction of a Möbius transformation. This is not at all elementary. For one proof

(assuming the map is at least C^2, see "Inversion theory and conformal mappings" by [?]. The result is still true if we assume only C^1, but even harder to prove.

3. The Schwarz-Christoffel Formula

The Schwarz-Christoffel formula gives a formula for the Riemann map of the disk onto a polygonal region Ω: if the interior angles of P are $\alpha\pi = \{\alpha_1\pi, \ldots, \alpha_n\pi\}$, then

$$f(z) = A + C \int^z \prod_{k=1}^{n} (1 - \frac{w}{z_k})^{\alpha_k - 1} dw,$$

where $\{z_1, \ldots, z_n\}$ are the points that map to the vertices of the polygon (and will be called the prevertices or conformal prevertices or z-parameters). See e.g., [?], [?], [?]. The interior angles of an n-gon sum to $(n-2)\pi$, which implies $\sum_k \alpha_k = -2$.

On the half-plane the formula is

$$f(z) = A + C \int \prod_{k=1}^{n} (w - z_k)^{\alpha_k - 1} dw.$$

In the case of the half-plane, there is a special boundary point, namely ∞. We assume this point is mapped to the last vertex, v_n, of the polygon, then the Schwarz-Christoffel formula can be written as

$$f(z) = A + C \int \prod_{k=1}^{n-1} (w - z_k)^{\alpha_k - 1} dw.$$

The formula was discovered independently by Christoffel in 1867 [?] and Schwarz in 1869 [?], [?]. For other references and a brief history see Section 1.2 of [?]. It is also possible to formulate it with other base domains, such as an infinite strip (see [?]). See [?] for a version involving doubly connected polygonal regions. There are also versions for domains other than polygons, e.g., circular arc polygons as in [?], [?]. In this case, we get a simple formula for the Schwarzian derivative of the conformal map, but it involves unknown parameters with no obvious geometric interpretation.

LEMMA 14. *With notation as above* $\sum_{k=1}^{n} (\alpha_k - 1) = -2$.

PROOF. The interior angles of an n-gon sum to $(n-2)\pi$, so

$$\sum_{k=1}^{n} (\alpha_k - 1) = \frac{1}{\pi}(n-2)\pi - n = -2.$$

□

If we apply a Euclidean similarity to a polygon, the interior angles do not change. Thus the α parameters do not change. Such a mapping also leaves the z parameters unchanged. Thus the maps for different but similar polygons differ only by the constants A and C. Changing the first translates the image and the changing the second alters the size and orientation.

LEMMA 15. *The function $\prod_{k=1}^{n-1}(w-z_k)^{\alpha_k-1}$ is a non-vanishing holomorphic function on the upper half plane which extends continuously to each component of $\mathbb{R}\setminus\{z_k\}$ and has constant argument on each such component.*

PROOF. Suppose α and c are real numbers and consider $f(w) = (w-c)^\alpha$. If f is not an integer, then this is not a single valued holomorphic function on the whole plane. To make it holomorphic we need to remove a branch cut from c to ∞ and define a single valued branch on the remaining domain. For the Schwarz-Christoffel formula we want the integrand to be holomorphic in the upper half-plane, so we can choose any branch cut in the lower half-plane, and we choose a branch of $(w-c)^\alpha$ which is positive if w is real and $w > c$. If we do this then $\arg((w-c)^\alpha)$ piecewise constant on the real line with a jump discontinuity at c. It has value 0 to the right of c and value $\alpha\pi$ to the right of c.

When we multiply the various terms in the integrand of the SC formula, the arguments add. Thus the argument of the integrand is piecewise constant with jump at z_k parameter of size $\theta_k = \pi\alpha_k$, i.e.,

$$\arg(f') = \arg(C) + \sum_{k=1}^{n}(\alpha_k - 1)\arg(w - z_k).$$

The image of the segment $I_k = [z_k, z_{k+1}]$ thus has constant argument (i.e., it lies in a line segment) and the angle between the images of I_k and I_{k+1} is θ_k.

The first claim is obvious since it is a product of non-vanishing holomorphic functions. □

Since f' is bounded, except possibly at the z-parameters, its integral f is well defined and has a continuous extension to the boundary except at these points (in fact, it has a holomorphic extension across the complementary intervals). It will have a continuous extension to a z-parameters z_k if $|f'(z)| = O(|z-z_k|^\beta)$ for some $\beta > -1$, so that f' is integrable on the boundary in a neighborhood of z_k. This happens as long

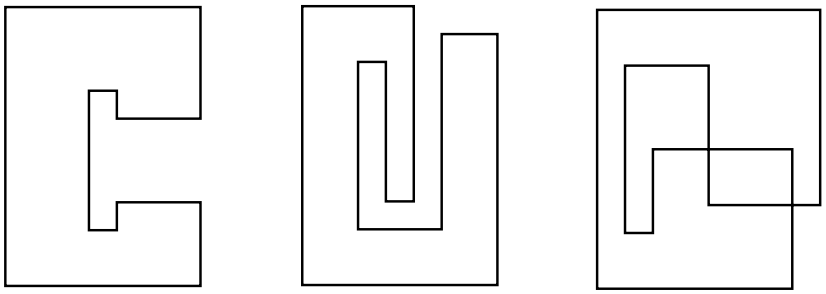


FIGURE 13. Different possible SC images using the same angles but different z-parameters.

as $\alpha_k > 0$ which happens iff $\theta_k > 0$. This is always the case for bounded polygons. It can happen that $\theta_k = 0$ for some unbounded polygons, e.g., an infinite strip can be considered as an unbounded polygon with two interior angles of size zero. In this case the Schwarz-Christoffel formula predicts that the map from the disk to the strip with prevertices $0, \infty$ should be given by

$$f(z) = A + C \int w^{-1} dw = A + C \log z,$$

which is correct. Moreover the maps fails to continuous exactly at the parameter value $z_1 = 0$ where the integrand has a pole of order -1.

In general, if we have two infinite edges which tend to infinity in directions $\theta_1 < \theta_2 \leq \theta_1 + 2\pi$ we define the corresponding interior angle at ∞ to be $\theta_1 - \theta_2$. This is in the interval $[-2\pi, 0]$ (including both endpoints, unlike the finite case). If we mapped the unbounded polygon by a Möbius transformation so that ∞ is mapped to the finite point (so the edges of the image are now circular arcs and not necessarily line segments), then this is the same as the negative of the interior angle at the image vertex. With this convention for the angle at ∞, the Schwarz-Christoffel formula can be extended to handle unbounded polygons (e.g., see []), but in these notes we will concentrate on the bounded case.

Our remarks so far prove the following.

LEMMA 16. *The SC formula for parameters $\{\alpha_k\}$ (with $\sum_{k=1}^n \alpha_k = -2$) and $z_1 < \ldots z_n$ defines a locally 1-1 holomorphic function which extends continuously to the boundary and maps each parameter interval 1-1 onto a line segment. If the map*

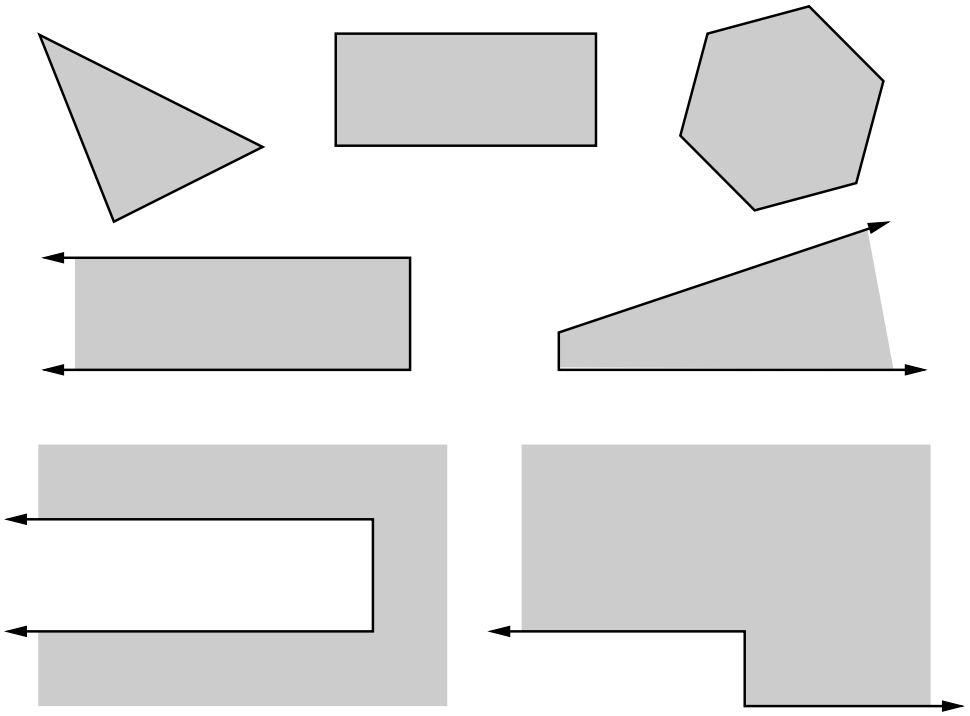


FIGURE 14. Some examples of regions for which the Schwarz-Christoffel formula gives explicit maps (i.e., we don't need to solve for the z-parameters). Triangles have only three prevertices which can be placed wherever we like. Similarly for the unbounded regions shown (assuming the angle at ∞ is is interpreted as discussed in the text). Because of the symmetries, prevertices for regular n-gons can be taken to be the roots of unity. Also because of symmetries the pervertices for a rectangle can be placed at $\pm\exp(\pm i\theta)$ where θ depends on the eccentricity e of the rectangle. This relationship is given by an explicit infinite product, which will be discussed in Section 5.

is globally 1-1 on the boundary then it defines a conformal map from the half-plane to a polygon with interior angles $\{\theta_k\}$

The last claim follows because this is a general property of holomorphic maps: if they are 1-1 on the boundary, they must be 1-1 on the interior.

The only remaining question is whether every bounded polygon can occur as the image of such a map, i.e., given the polygon, can we find parameter values so that the SC formula gives a map to the polygon? We shall see that this is the case later using the Riemann mapping theorem.

First, however, we want to check that the observations made above for the half-plane also apply to the Schwarz-Christoffel formula for the disk.

For the disk formula, there is a similar argument. The term

$$(1 - \frac{w}{z}) = \frac{z - w}{z} = -\frac{z - w}{0 - z},$$

has an argument that equals the angle between the vector $z - w$ and $0 - w$, which is the same as $(\pi - \psi)/2$ where ψ is the angle between z and w (see Figure 15. This means that as we move around the unit circle with $w = \exp(i\theta)$, the function $u(\theta) = \arg(1 - \frac{w}{z})$ has constant derivative $1/2$ except at the point z where the function jumps from $\pi/2$ to $-\pi/2$.

When raise this term to a power $(1 - w/z)^\alpha$, the argument changes with derivative $\alpha/2$ except for a jump of size $\alpha\pi$. When we take the product

$$\prod_{k=1}^{n} (1 - w/z_k)^{\alpha_k - 1},$$

the arguments sum, and so the argument of the product increases with derivative

$$\sum_{k=1}^{n} \frac{1}{2}(\alpha_k - 1) = \frac{1}{2}(-2) = -1,$$

except at the points $\{z_k\}$ where there are jumps of size $\pi(\alpha_k - 1)$ Thus the argument of

$$\prod_{k=1}^{n} (1 - w/z_k)^{\alpha_k - 1} dw,$$

is constant on the arcs between parameter values. Thus the Schwarz-Christoffel formula maps onto a polygonal region with the correct angles.

For maps onto a polygon there are three unknown SC-parameters, but a Möbius transformation of the disk can map three distinct points on the circle to any other three distinct points, so any triple will do. The only difference between the different choices is where the origin will map to. We are free to choose the harmonic measures of the three sides of the triangle any way we wish (as long as they sum to 1) and Schwarz-Christoffel formula gives the corresponding map. See Figure 16.

Quadrilaterals are the first case where we have a non-trivial parameter problem to solve. We have to find the correct 4-tuple on the unit circle and every 4-tuple is determined by its cross ratio $P \in (0, \infty)$ (up to Möbis equivalence). In the special case of rectangles, the domain is also determined by a single number, the eccentricty

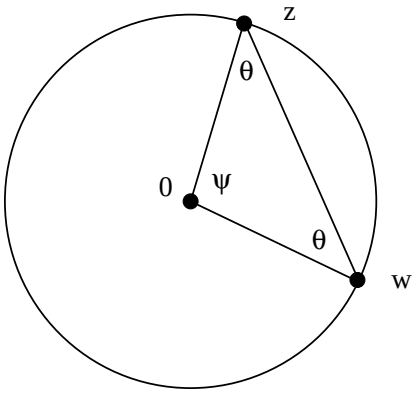


FIGURE 15. Since $0, w, z$ form an isosceles triangle, $\psi + 2\theta = \pi$, or equivalently, $\theta = (\pi - \psi)/2$.

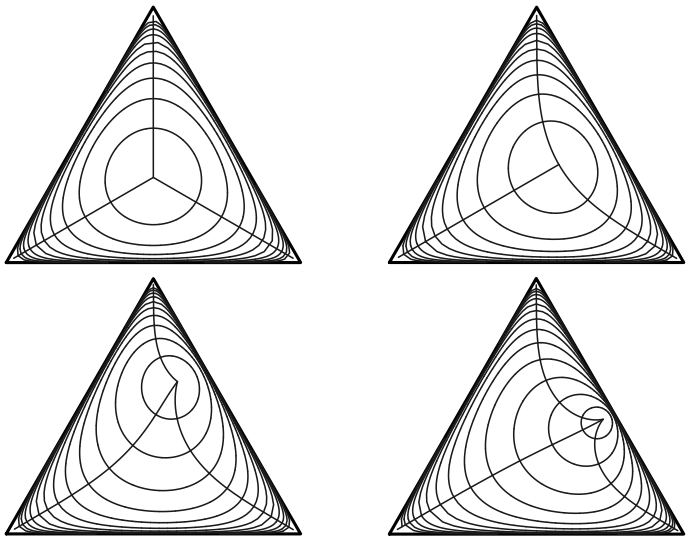


FIGURE 16. The Schwarz-Christoffel map onto a triangle has no parameters to solve for; any three distinct points give a map onto the triangle, but with the origin going to different points

R (up to Euclidean similarity). Thus there we expect a 1-1 correspondence between cross ratios of 4-tuple on $\mathbb{T}$ and eccentricities $0 < R < \infty$. Suppose Ω is a rectangle with vertices at $\{0, R, R+i, i\}$ and we have a conformal map of Ω to the upper half-plane that sends the vertices to $\{0, P, 1, \infty\}$. Alternatively, we might send the points to $\{Q, 0, 1, \infty\}$, $\{\infty, 0, M, 1\}$ or $\{0, 1, N, \infty\}$. It easy to see that

$$P(R) = M(\frac{1}{R}), Q(R) = \frac{P(R)}{P(R)-1}, N(R) = \frac{1}{P(R)},$$

so it is enough to calculate any one of these functions. For example $M(R)$ is given by

$$M = \exp(-\pi R)\frac{1}{16}\prod_{n=1}^{\infty}\left(\frac{1+\exp(-2n\pi R)}{1+\exp(-(2n-1)\pi R)}\right)^8. \tag{5}$$

We will give a proof of this later (see Chapter 5).

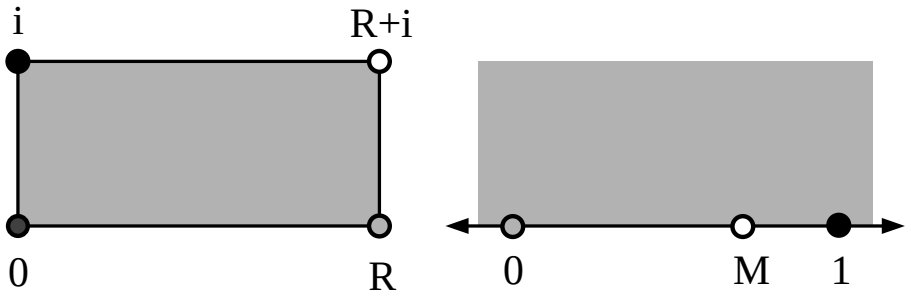


FIGURE 17. We assume f maps a rectangle to the upper half-plane with the vertices mapped as shown.

For $R > 0$ the infinite product converges and for R large (say $R \geq 1$) we have

$$\prod_{n=1}^{\infty}\left(\frac{1+\exp((1-2n)\pi R)}{1+\exp((-2n)\pi R)}\right)^8 = 1 + 8e^{-2\pi R} + O(e^{-4\pi R}).$$

Thus for $R \geq 1$, (equivalently $2 \geq 1$), we have

$$\log\left(\frac{1}{M}\right) = \pi R - \log 16 + 8e^{-\pi R} + O(e^{-2\pi R}),$$

which implies

$$M \simeq \exp(-\pi R).$$

If we take the 4-tuple to be $\{w, -\bar{w}, -w, \bar{w}\}$ where $w = e^{i\theta}$ is the first quadrant, then the cross ratio is easily computed to be

$$P = \tan^2(\theta),$$

or $\theta = \arctan(\sqrt{P})$. So if we want to compute the conformal map onto a $1 \times R$ rectangle, we compute M by (14), then compute θ as above and use the four points given. To find R given M, we can use a secant method for find a root of $M(R) = M_0$.

There is no simpler formula for the inverse of the Schwarz-Christoffel map but the inverse for a particular point can be computed either by using a Newton iteration

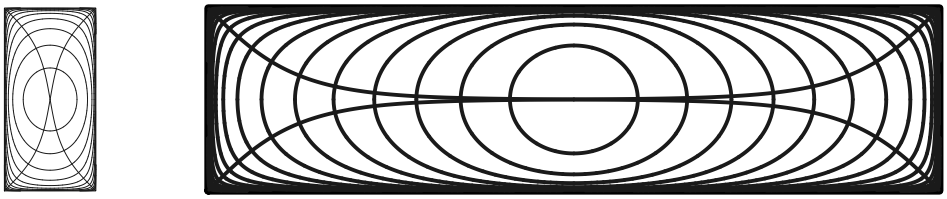


FIGURE 18. Rectangles plotted using the Schwarz-Christoffel formula and the relation between eccentricity and cross ration for $R = \frac{1}{2}$ and $R = 4$.

on the forward map, or numerical solution of the initial value problem

$$\frac{dz}{dw} = \frac{1}{f'(z)}, z(w_0) = z_0.$$

The Newton iteration is faster, but requires a good initial guess. Solving the IVP numerically is generally more reliable but slower, according to [**?**].

4. Crowding

Formula **??** also illustrates one of the main problems with numerical conformal mapping: crowding. The map of the disk onto a $1 \times R$ rectangle uses a 4-tuple on the circle with a cross ratio of $\approx \exp(-\pi R)$. If we take symmetric points $\{w, -\bar{w}, -w, \bar{w}\}$ with $w = e^{i\theta}$, for a moderate $R = 20$, this means

$$\theta \approx 5.1579 \times 10^{-28} = .00000000000000000000000000051579.$$

The separation between w and $\bar{w}$ is only about twice this size, which is smaller than machine precision on most computers. Thus, unless we take special care, a computer may think the parameters are $\{1, -1, -1, 1\}$ (which can be interpreted as the parameters for an infinite strip).

The connection between harmonic measure and Brownian motion gives us a good way to getting a "feel" for what the harmonic measure should look like. Consider the infinite strip in Figure 5 which has been divided into squares and start a a Brownian a the center of one of these squares. By symmetry it has an equal chance of first hitting any of the sides of the square and hence has a 1/2 probability of hitting the top or bottom of the strip before leaving the square. If it does it the left or right side of the square, then it has less than 1/4 chance of hitting a different dashed segment before running into the edge of the strip. Thus is has a less that 2^{-2n} chance of hitting n distinct dashed lines. Thus the harmonic decays exponentially as we travel down a

strip. A similar argument shows that in a $1 \times R$ rectangle, the harmonic measure of the short sides is at most $O(2^{-R})$ with respect to the center. Thus at least one side has harmonic measure this small, regardless of where we choose the Brownian motion to begin. This means that any conformal map of the rectangle to the disk must send two vertices to within distance 2^{-R} of each other. For polygons with long, narrow channels, this means that not all prevertices may be distinct in machine precision.

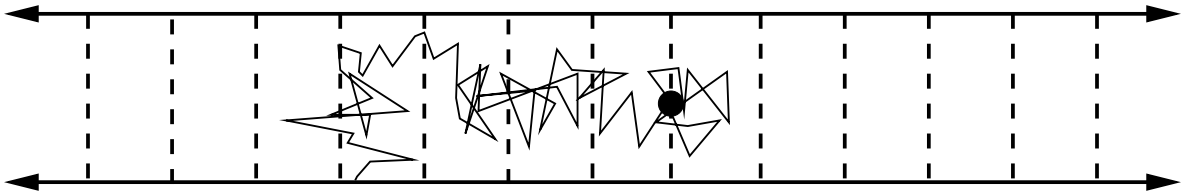


FIGURE 19. Brownian motion explains why harmonic measure decays exponentially fast in a strip. A motion starting on any of the dashed line has probability $\leq 1/4$ of hitting each of the adjacent dashed lines (with equality only if it starts at the midpoint).

In fact, the situation is a bit worse than indicated above, since the upper bound is not sharp. We can actually compute a conformal map from the disk to the strip $(0,1) \times (-\infty, \infty)$ as

$$z \to \frac{1}{\pi} \log(i \frac{z-1}{z+1}).$$

which shows that points with $|z-1| \approx \epsilon$ are mapped to points with $\Re(w) \approx \frac{1}{\pi} \log \frac{1}{\epsilon}$. Thus the short sides of a $1 \times R$ rectangle actually have harmonic measure $\approx \exp(-\pi R)$ with respect to the center. Even for $R = 10$, this is $\approx 2 \times 10^{-14}$. Thus we would have to start about a trillion random walks at the center of a 1×20 rectangle to expect to hit the short sides even once. Thus our method of estimating z-parameters using random walks is not practical in general.

The crowding phenomenon is the source of many of the difficulties in numerical conformal mapping. Roughly it says that conformal maps from a domain into a disk can undergo exponential compression, so that points that are well separated in the domain become identified in the disk (at least with finite precision). The inverse map from the disk to the domain is, in a loose sense, not even well defined numerically. The problem is that there is no choice of "center" for a polygon from which all the sides look about the same size (in the sense of harmonic measure) or even within

several orders of magnitude of the same size. we shall see later (e.g. Section 5) that one way around this is to compute conformal maps with respect to several different centers such that any small part of the domain "looks uncrowded" from some center point.

For the present we will simply avoid computing any examples where the separation of the SC-parameters is too small. Later we will present methods for dealing with such domains.

5. Power series of Schwarz-Christoffel maps

Note that using the general form of the binomial theorem,

$$(1+z)^p = \sum_{k=0}^{\infty} \frac{p(p-1)\cdots(p-k+1)}{k!} z^k,$$

we can easily compute power series for these functions in disks away from the singularities. For example, suppose we take $n = 4$, $\alpha = \{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\}$ and parameters $\mathbf{z} = \{1, i, -1, -i\}$. The Schwarz-Christoffel formula gives

$$\begin{aligned} f'(w) &= \prod_{k=1}^{n} (1 - \frac{w}{z_k})^{\alpha} \\ &= (1-w)^{-1/2}(1+iw)^{-1/2}(1+w)^{-1/2}(1-iw)^{-1/2} \\ &= (1-w^2)^{1/2}(1+w^2)^{-1/2} \\ &= (1-w^4)^{-1/2} \\ &= 1 + \frac{1}{2}w^4 + \frac{3}{8}w^8 + \frac{5}{16}w^{12} + \frac{35}{128}w^{16} + \dots \end{aligned}$$

so

$$f(z) = z + \frac{1}{10}z^5 + \frac{3}{72}z^9 + \frac{5}{208}z^{13} + \frac{35}{2176}z^{17} + \dots.$$

This series is plotted for various truncations in Figure **??**.

If we change the SC-parameters by a Möbius transformation, the image has the same shape, but the origin is mapped to a different point. Figure 21 illustrates this, In the first case, the Möbius transformation is symmetric with respect to the real axis so the four new parameters are as well, and hence the image domain is also symmetric. In the second example, the parameters are no longer symmetric, so we might not expect the image to be symmetric either (but since the angles are the same, it will still be a square), However, the figure is not rotated. This is becuase

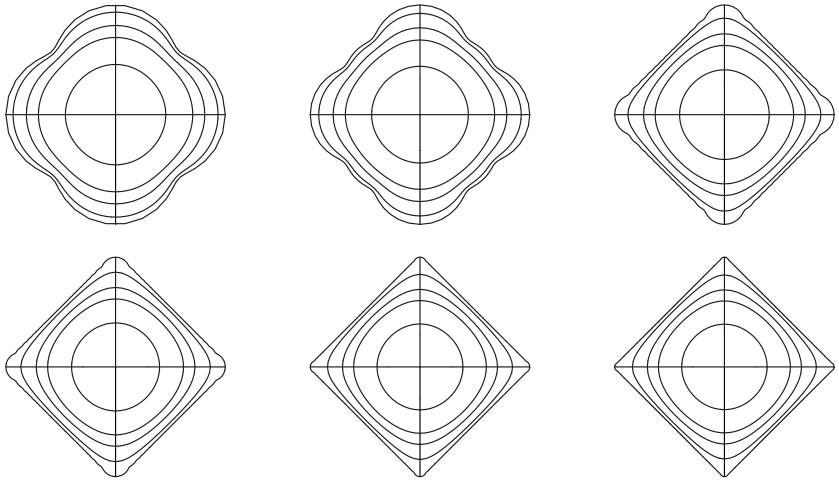


FIGURE 20. The image of the unit disk under different truncations of the power series for the conformal map onto a square. The truncations are at $n = 5, 10, 50, 100, 500, 1000$.

we choose our branches of $(1 - \frac{w}{z_k})^{\alpha_k - 1}$ to be real on the reals axis; thus the image edge containing the image of 1 will always be vertical (or have a fixed angle with the vertical if 1 is a SC-parameter).

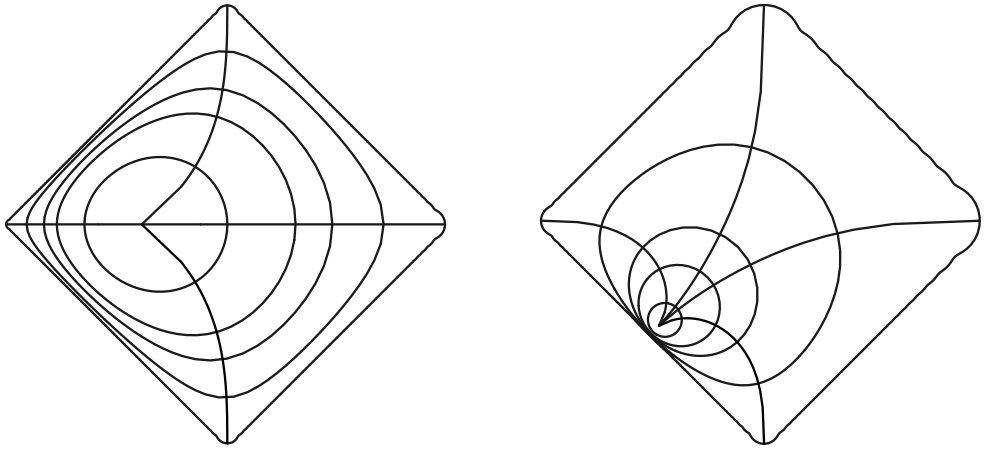


FIGURE 21. The same Schwarz-Christoffel map as in Figure 20, except that we have moved the SC-parametes by a Möbius transformation. In the first case we used $z \to (z - \alpha)/(1 - \bar{\alpha}z)$ with $\alpha = -1/2$ and in the second with $\alpha = -.9(1 + i)/\sqrt{2}$. Note, that this also effects the sharness of the corners, since some parameter values are now closer together.

The power series for more general Schwarz-Christoffel maps can be computed by finding the Taylor series for each term $(1+(\frac{-w}{z_k})^{\alpha_k-1}$ and then multiplying the series using the standard formula

$$(\sum_{n=0}^{\infty} a_n z^n)(\sum_{n=0}^{\infty} b_n z^n) = \sum_{n=0}^{\infty}(\sum_{k=0}^{n} a_k b_{n-k})z^n.$$

and then integrating term-by-term using

$$\int \sum_{n=0}^{\infty} c_n z^n = \sum_{n=0}^{\infty} \frac{c_n}{n+1} z^{n+1}.$$

For example, we we take

$$\alpha = \{\frac{1}{2}, \frac{3}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}\frac{1}{2}, \frac{1}{2}\},$$

and eight equally spaced SC-parameters, we get

$$\begin{aligned} f(z) \;=\; & z + -0.333333iz^4 + 0.0555556z^{10} - 0.0454545iz^{12}0.0220588z^{18} \\ & -0.0197368iz^{20} + 0.0125z^{26} \\ & -0.0115741iz^{28} + 0.00828598z^{34} - 0.0078125iz^{36} \\ & +0.00600229z^{42}, 0, -0.00572311iz^{44} + \ldots. \end{aligned}$$

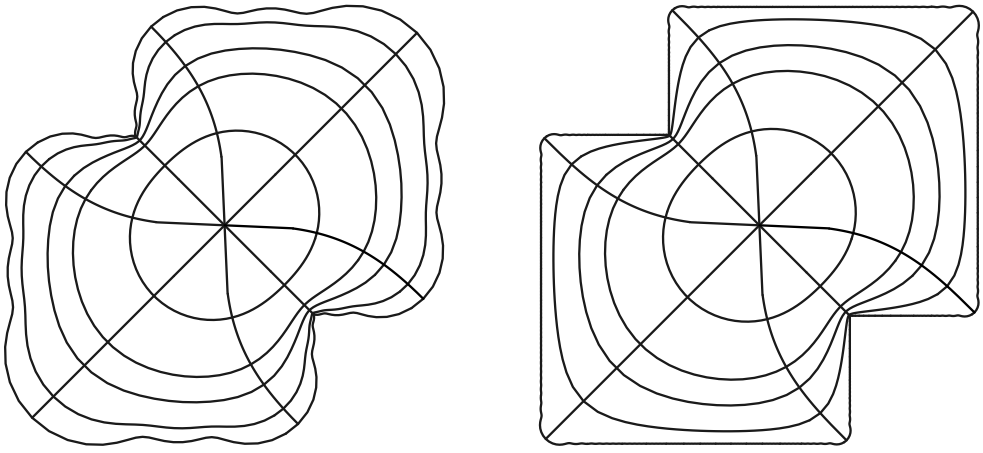


FIGURE 22. Truncations of a Schwarz-Christoffel map onto an 8-gon truncated at $n = 50$ and 500. Every edge has equal harmonic measure. Every edge has equal harmonic measure.

Even though every conformal map has a power series expansion in the disk, this expansion may not be a computationally effective way to represent the map. In the

following figures I show the image of the unit disk under truncations of the power series that represent conformal maps onto rectangles. These maps have the form

$$f(z) = \int_0^z \prod (1 - \frac{w}{e}^{-i\theta_k})^{-.5} dw,$$

where $\theta_k \in \{-\epsilon, \epsilon, \pi - \epsilon, \pi + \epsilon\}$. The form of this equation we be explained later, when we discuss the Schwarz-Christoffel formula. The parameter ϵ represents the probability that a random path started at the center of the rectangle will hit one of the two shorter sides before it hits one of the two longer sides. This determines the eccentricity of the rectangle, although the explicit relationship is a complicated infinite product and will be discussed later. The main point is that the smaller ϵ is, the longer the rectangle will be and the more terms will be needed to represent the map onto the rectangle accurately. In Figure 23, we take $\epsilon = 1$ and show truncations for $n = 5, 10, 50, 100, 500, 1000$. Even by $N = 100$ the the shape of the rectangle is clear. In Figure 24 we take $\epsilon = .001$ which corresponds to a much longer rectangle. In this case, we do not see the corners clearly even for 1000 terms of the power series. This example shows that the degree of approximation of a power series to a conformal map depends on the shape of the image domain. If there are thin corridors or "hard-to-reach" corners, then extremely high degree approximations may be needed. Later we will investigate ways to represent conformal maps that are accurate with much smaller storage.

The bad news is that this series converges rather slowly. If f is holomorphic on a disk $D = D(z_0, r)$, then f has a convergent power series

$$f(z) \quad = \quad \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

on this disk. Since this series converges, the terms must tend to zero, hence they must be bounded and so

$$\limsup_n |a_n| t^n < \infty,$$

for every $t < r$. In fact, if f extends continuously to the boundary of the disk then we can apply Cauchy integral formula and do a little better

$$|a_n| = |\frac{|f^{(n)}(z_0)|}{N!} = |\frac{1}{2\pi} \int_{w:|w-z_0|=r} \frac{f(w)}{(w - z_0)^{n+1}} dw| \leq \frac{1}{2\pi r^n} \max_{\partial D} |f(z)|.$$

If f is a conformal map from the disk onto a polygon there is at least one boundary point where $|f'|$ blows up to infinity (because there is at least one vertex which has interior angle $< \pi$). Therefore the radius of convergence for the power series for both f and f' has radius exactly one and the convergence will be slow near the boundary. A single singularity on the boundary can cause the power series to converge slowly every on the boundary, even at points where the function itself has an analytic convergence across the boundary. The best known example of this is the geometric formula

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \dots,$$

which has only one singularity on the unit circle, is analytic elsewhere in the plane but the power series diverges everywhere on the unit circle.

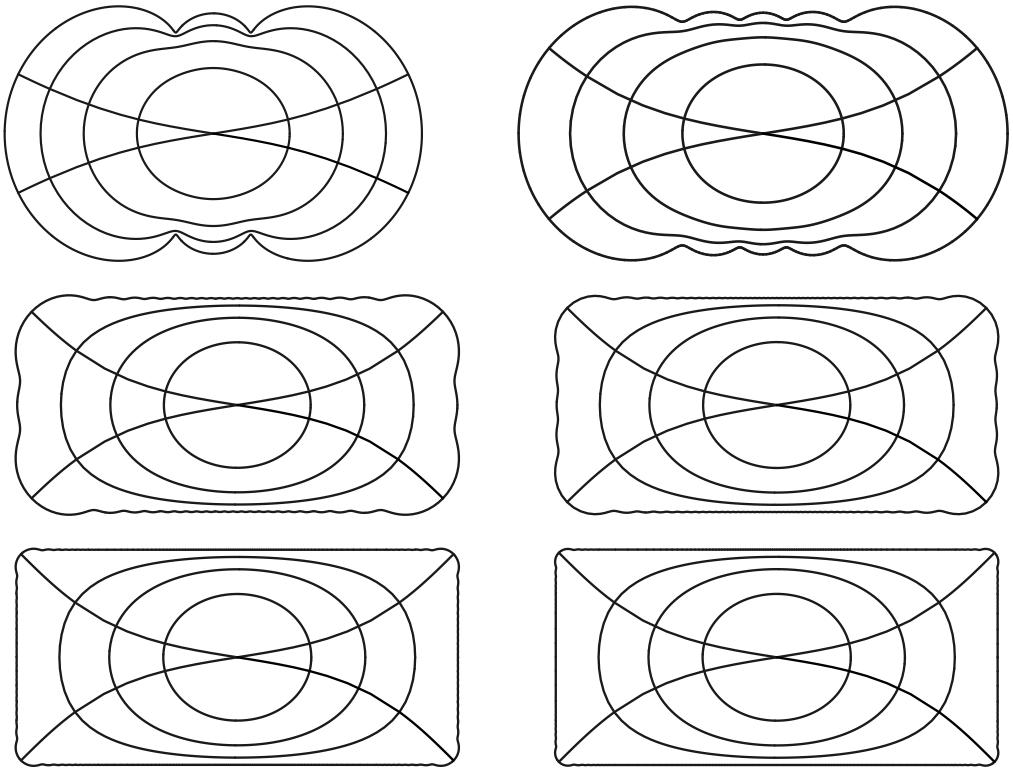


FIGURE 23. The image of the unit disk under different power series. Each one is a truncation of the infinite power series for the Schwarz-Christoffel map from the unit disk to a rectangle (chosen so the short edges have one tenth the harmonic measure of the longer sides). The truncations are at $n = 5, 10, 50, 100, 500, 1000$.

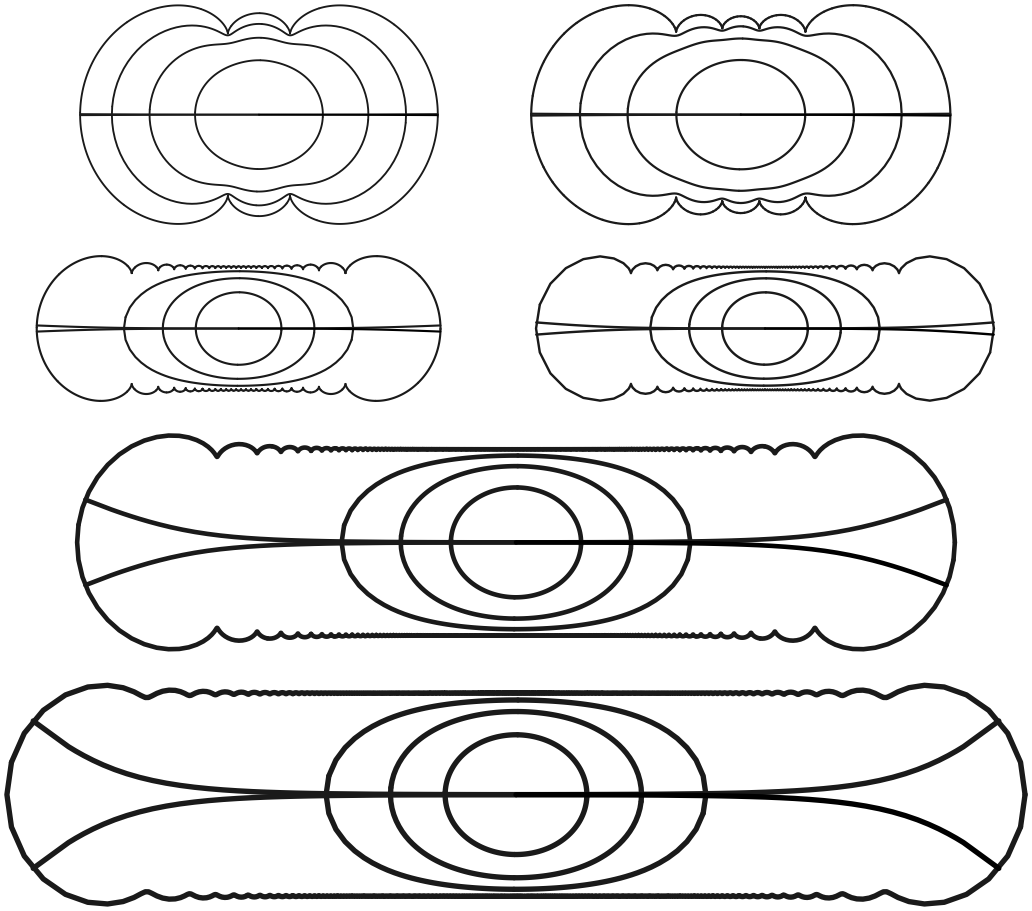


FIGURE 24. This is the same as the previous figure except that the target polygon has been changed so that the short sides have probability .001. The truncations are $n = 5, 10, 50, 100, 500, 1000$. This make the rectangle longer and requires a higher degree truncation to achieve the correct shape.

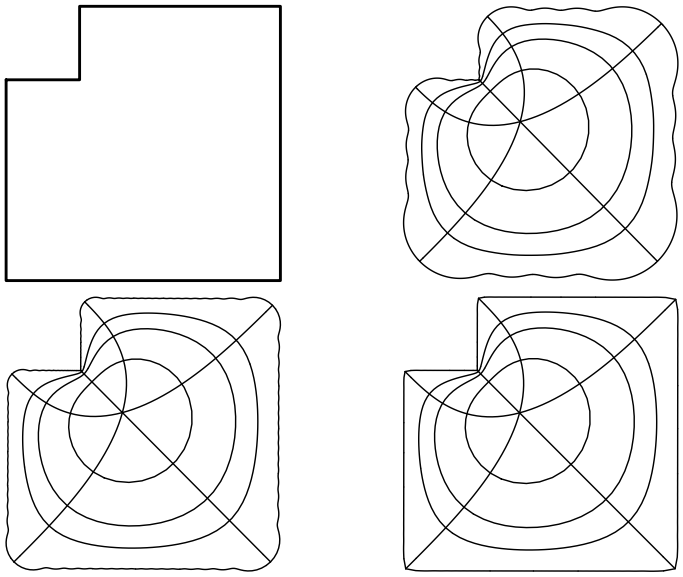


FIGURE 25. In the upper left is the target "L"-shaped polygon. The parameters for Schwarz-Christoffel are taken to be equidistributed in this example. The next three figures show the images of the disk under the power series for the map with truncations at order $20, 100, 1000$.

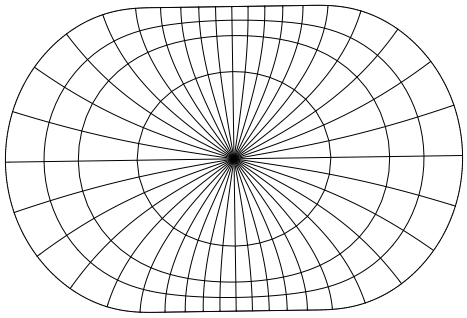


FIGURE 26. If the boundary is smooth then the power series is a better approximation. Here is a C^1 domain sampled at 40 boundary points to give a polygon. We then plotted the result using 20 terms of the power series. The domain is a square with half-disks attached to opposite sides.

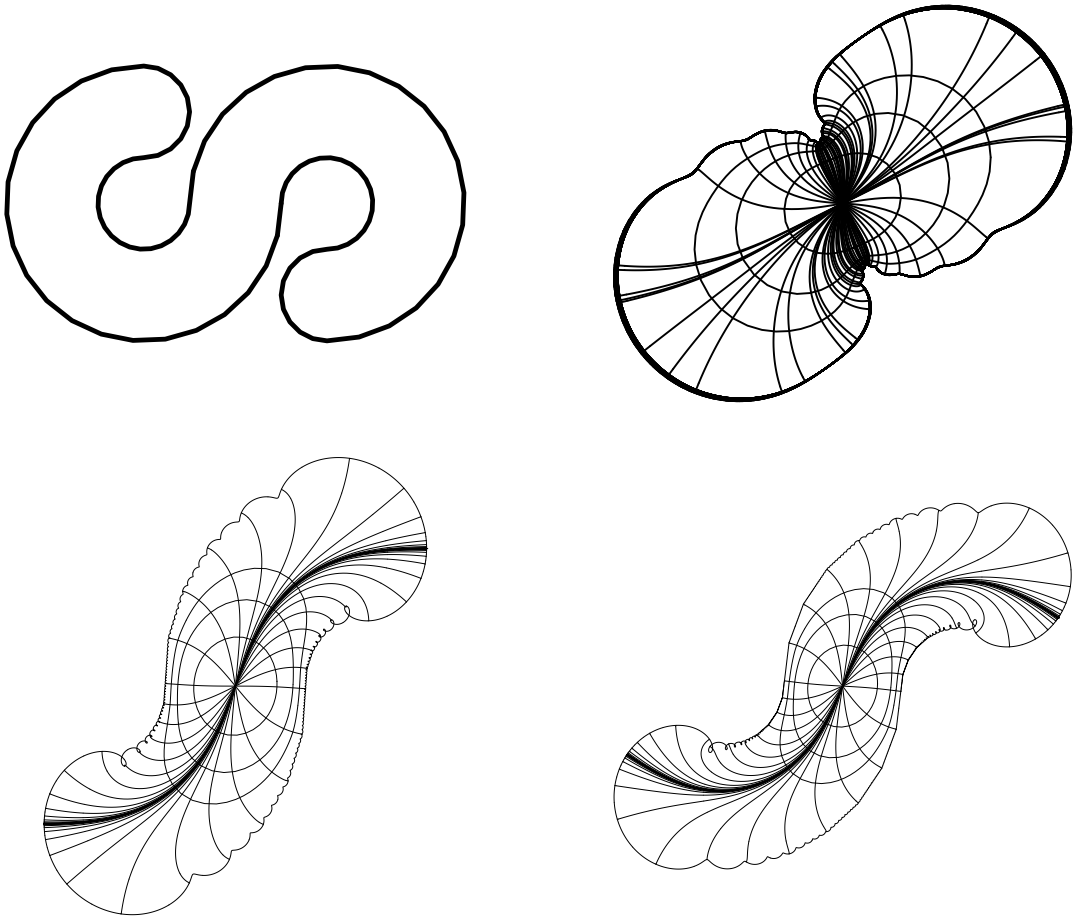


FIGURE 27. Even if the domain has a reasonable smooth boundary, crowding can still be a problem for power series. Here is an "S" shaped region and power series approximations with 100, 200 and 2500 terms.

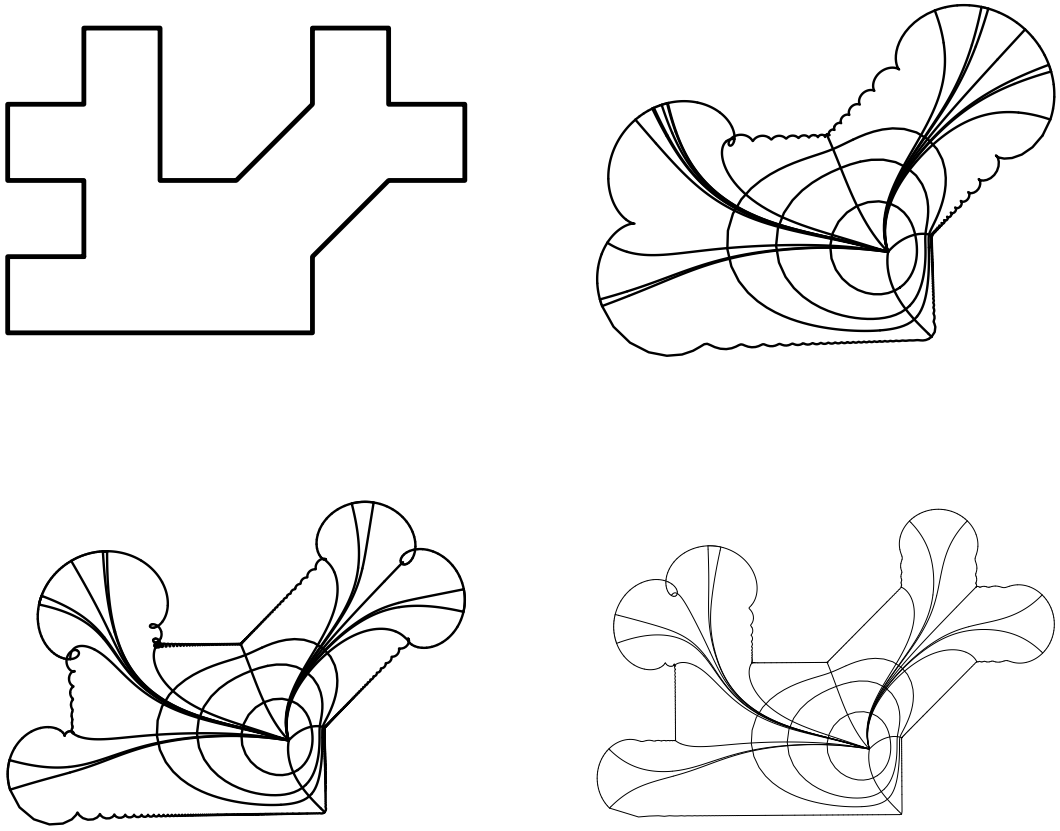


FIGURE 28. Again in the upper left is the target polygon. The next three figures show the images of the disk under the power series (centered at 0) for the Schwarz-Christoffel map with truncations of orders $100, 500, 2500$. In this example, the choice of SC-parameters is not obvious; the plots were made using parameters values found by Davis's method, which we will describe later.

6. Harmonic measure and Brownian motion

Suppose Ω is a simply connected domain. Choose a base point $z_0 \in \Omega$ and a conformal map $f : \mathbb{D}\Omega$. Assume for the moment that $\partial\Omega$ is a Jordan curve and that f has a 1-1, continuous extension to the boundary (see Caratheodory's theorem 5). Define the harmonic measure of $E \subset \partial\Omega$ with respect to the point z_0 as

$$\omega(z_0, E, \Omega) = \frac{1}{2\pi}|f^{-1}(E)|.$$

This does not depend on the particular choice of f since any two maps sending 0 to z_0 differ by a rotation of the disk.

If Ω is bounded by a polygon, then the harmonic measure of each side is determined by the spacing of the Schwarz-Christoffel parameters. Thus finding a set of parameters is equivalent to computing the harmonic measure of the sides of the polygon.

Brownian motion is the rigorous version of the idea of a "continuous random walk" in the plane. One can think of this as a limit of a random walk on an ϵ-grid as $\epsilon \to 0$. The important thing is that Brownian motion is conformally invariant; i.e., the image of Brownian motion under a conformal map is Brownian motion on the image domain. The harmonic measure $\omega(z, E, \Omega)$ is the probability that a Brownian motion started from z will first hit $\partial\Omega$ in the set E. On the disk, this is just normalized length measure on the boundary.

Brownian motion is a continuous version of a random walk. That is, it is a stochastic process $B(t)$, $t > 0$ such that

(1) Increments are independent: if $t_0 < t_1 < \ldots t_n$, then the random variables $B(t_0)$, $B(t_1) - B(t_0), \ldots B(t_n) - B(t_{n-1})$ are independent.
(2) Increments are normally distributed: if $s, t \geq 0$ then

$$\text{Prob}(B(s+t) - B(s) \in A) = \int_A (2\pi t)^{d/2} e^{-|x|^2/2t} dx.$$

(3) With probability one $B(t)$ is a continuous function of t.

We can also think of d-dimensional Brownian motion as a probability measure on the set of continuous paths in $\mathbb{R}^d$, i.e., functions from $[0, \infty)$ into $\mathbb{R}^d$. This is called Weiner measure and we shall denote it by Prob. For example, using this notation

property (3) would be written

$$\text{Prob}(B(t) \text{ is continuous }) = 1.$$

An event which happens with probability one with respect to Weiner measure will be said to happen almost surely.

Brownian motion was first described mathematically by Einstein in 1905, but was first proven to exist by Weiner in 1923. Brownian motion can be considered as the continuous limit of a random walk on a square grid, as shown in Figure 29. In fact, this grid structure is not really needed; one can take Brownian motion to be the limit of many different discrete walks (which must have mean value zero at each step). For example, in Figure 30, we show it as a limit of a walk in a triangular grid (at each time we may step unit distance in any of six directions) and in Figure 31 we show a walk in which at each time we step unit distance in any direction (chosen uniformly and at random). The fact that Brownian motion is the limit of this last process as the step size decreases to zero, gives a heuristic reason for its conformal invariance: since conformal maps send infinitesimal circles to infinitesimal circles the "random direction" process should be mapped to another such process, except with variable size circles depending on the size of the derivative of the conformal map.

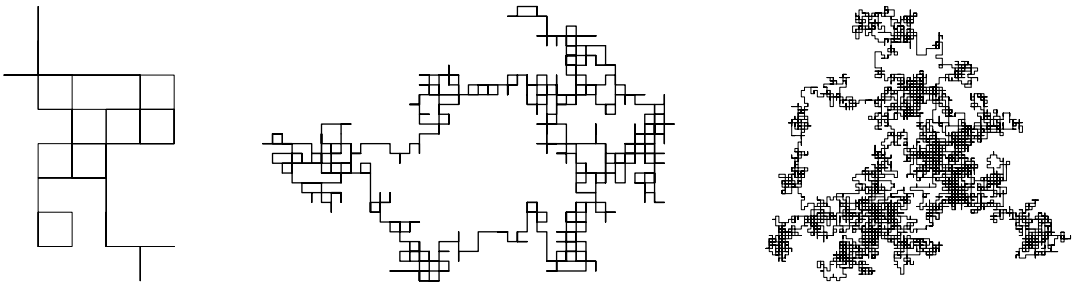


FIGURE 29. Random walks on a square grid with 100, 1000 and 10,000 steps.

Brownian motion itself is a rather technical process to deal with but there is a simpler process on Ω that has the same hitting distribution on the boundary and was introduced by Kakutani. Moreover, simulating a Brownian motion by a random walk on a very fine grid can take a long time, and Kakutani's process will speed this up. Starting at a point of $z\Omega$ choose a radius r so that $B(z,r) \subset \Omega$, e.g., take $r\lambda \text{dist}(z, \partial\Omega)$ for some fixed $0 < \lambda \leq 1$. If we start a Brownian motion at z and wait

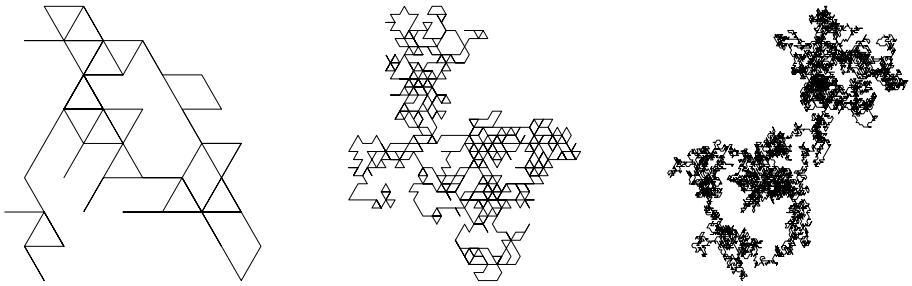


FIGURE 30. Random walks on a triangular grid with 100, 1000 and 10,000 steps.

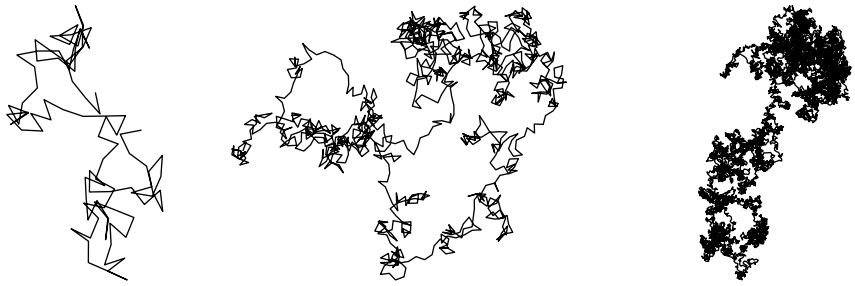


FIGURE 31. Random paths formed by stepping unit distance in a randomly chosen direction. The pictures show paths with 100, 1000 and 10,000 steps. Note that regardless of the discrete random walk, at large scales the results all look the same.

for the first contact with ∂B, this is the same as simply choosing a point at random on ∂B. Repeat this procedure to construct a sequence of points. This is the same as choosing a sequence along a single Brownian path with the selections becoming more frequent as the path approaches the boundary of the domain. With probability 1 the Brownian path hits the boundary of the domain and the sequence of points constructed must converge to the same boundary point. Thus the hitting probability of Kakutani's process is the same as for Brownian motion. See Figure **??**. Moreover, this process is faster to simulate. If we are simulating Brownian motion by a walk on an ϵ-grid, and the starting point is about unit distance from the boundary, then it takes about ϵ^{-2} steps to hit the boundary (or reach a grid point that is within ϵ on the boundary). On the other hand the Kakutani process will be within ϵ of the boundary in about $\log \frac{1}{\epsilon}$ steps. The main cost is to recompute the distance to the boundary each time (which is at worst an $O(n)$ computation in an n-gon and may

be faster if we are clever, e.g., precompute a Voronoi diagram and keep track of what cell we are in).

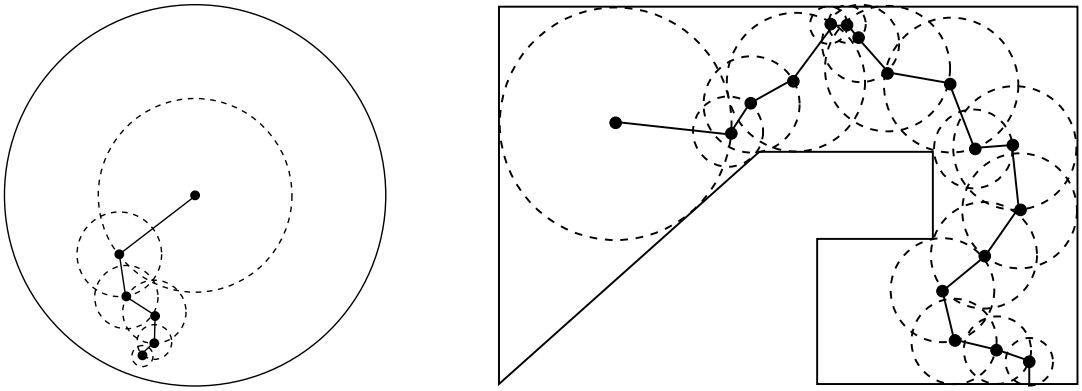


FIGURE 32. Examples of the paths in the Kakutani process with $\lambda = 1/2$ in the disk and $\lambda = 1$ in a polygon.

Therefore one way to approximate the z-parameters for a polygon is to choose a basepoint for the interior and simulate many Brownian paths starting from this point and keep a count of how many hit each edge. This count (divided by the total number of paths) gives the approximate harmonic measure of each edge and hence the approximate separation between the corresponding prevertices. The trouble with this method is that it is slow, very slow. In the best case, the number of sides of the polygon is small and all sides have comparable harmonic measure, i.e., they are all about equally likely to be hit. The bad news is that the hitting frequencies of random paths will converge to the actual harmonic measures with error that tends to zero like $n^{-1/2}$, e.g., about a million random walks are required to get three decimals of accuracy. The really bad news is that usually the sides do not all have large harmonic measure and if some sides have very small measure then we have to wait even longer for them to get hit frequently enough to estimate their measure. For example, consider the "L"-shaped polygon in Figure **??**. The vertices are $\{0, 2, 2 + 31, 1 + 3i, 1 + i, i\}$ and the starting point is $1.5 + i$. The left picture shows 10 sample paths of the Kakutani process, the center shows 100 paths and the right shows 1000 paths. Even after a 1000 sample paths, only one has managed to reach the top horizontal edge, and this polygon is by no means extreme.

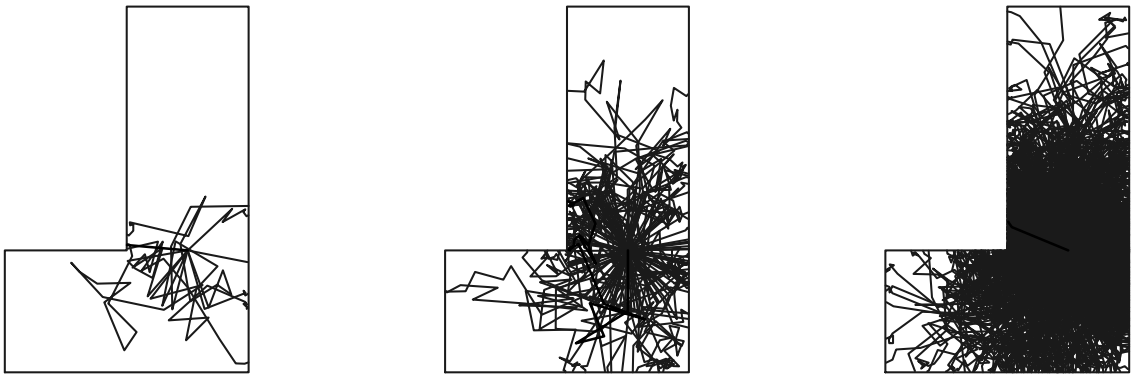


FIGURE 33. These show the polygon with 10, 100, and 1000 random walks. Note how hard it is for the remote edges to get hit. After 1000 attempts the top edge still has not been hit. If the smallest harmonic measure of any edge is ϵ the we expect to need $1/\epsilon$ to get even one hit on that edge.

Table 1 shows the number of hits per side for two experiments; one with 1,000 random paths and one with 10,000 random paths. In Figure 34, we show the images of the Schwarz-Christoffel maps when we use the parameters given by these experiments. The left picture is the target polygon, the center is uses the parameters from the 1,000 path experiment and the right uses 10,000 paths. The latter looks noticeable better, and we shall see later that it is indeed about 10 times better, according to a certain precise measure of the closeness of polygons. In Section 5 we shall discuss an improvement of this method, but using random walks to estimate harmonic measure should not be considered a really practical approach. We have introduced it because it is pretty, easy to understand and is guaranteed to produce the correct answer (eventually). Moreover, the intuition provided by thinking of harmonic measure (and hence the z-parameters) as the hitting probability of Brownian motion is invaluable and frequently leads us to quickly to the right answer (even if we later replace this intuition by a calculation or proof based on techniques such as extremal length, hyperbolic geometry or potential theory). We shall see one such case in the next section.

1	131	0.823097	0
2	516	3.24212	0.823097
3	1	0.00628319	4.06522
4	271	1.70274	4.0715
5	76	0.477522	5.77425
6	5	0.0314159	6.25177

1	1143	0.718168	0
2	4952	3.11143	0.718168
3	27	0.0169646	3.8296
4	2833	1.78003	3.84657
5	961	0.603814	5.62659
6	84	0.0527788	6.23041

TABLE 1. The left table was generated using 1,000 random walks and the right by 10,000 random walks. in each table the columns give, respectively, the side number, the number of hits on that side, the corresponding spacing between parameters and a choice of the parameters themselves. The vertices are $\{0, 2, 2+31, 1+3i, 1+i, i\}$ and the center point is $z = 1.5 + i$. In our numbering scheme the first edge is the horizontal edge on the bottom.

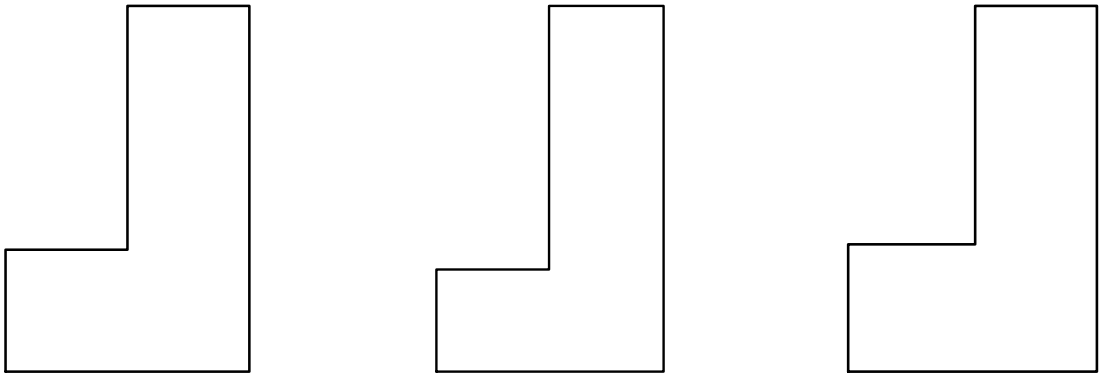


FIGURE 34. On the left is the target polygon, in the center the SC image derived from 1,000 random walks and on the right is the SC image using parameters derived from 10,000 random walks inside the polygon. The values of the parameters are given in Table 1.

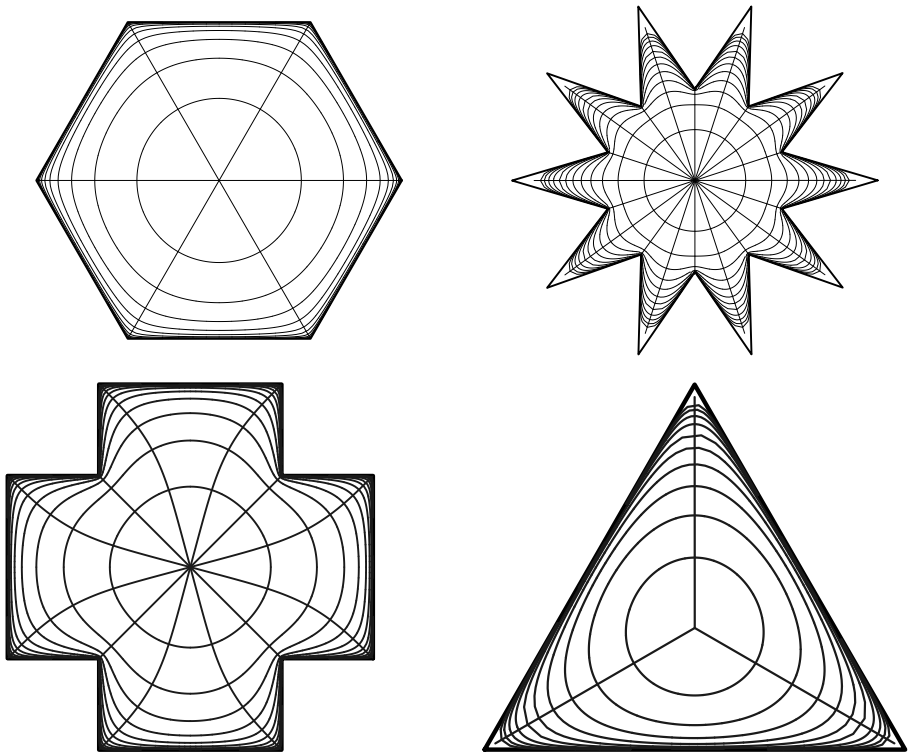


FIGURE 35. Polygons for which every edge has the same harmonic measure.

If a domain has a symmetry, then the conformal map should have a corresponding symmetry. For example, if the domain Ω is symmetric with respect to the real line, and we choose the conformal map $f : \mathbb{D} \to \Omega$ with $f(0)$ real and $f'(0) > 0$, then f will also be symmetric with respect to the real line, i.e., $f(\bar{z}) = \overline{f(z)}$. When a polygon has symmetries, then certain edges must have harmonic measures which agree (at least with respect to properly chosen base points) and this reduces the number of independent parameters we must solve for in the Schwarz-Christoffel formula.

For example, the domains in figure 5 have rotational and reflection symmetries that map any edge to any other edge, so every edge must have the same harmonic measure. Thus the parameters must be evenly distributed around the circle.

Even if the domain is not so symmetric that every edge has the same harmonic measure with respect to some point, there may still be able to group the edges into subcollections, so that each edge in a subcollection has the same harmonic measure. For example, the domain in Figure **??** is the second generation in the construction

of the von Koch snowflake and has 48 sides. However, the symmetries of the domain divide the edges into 4 classes of twelve sides each and the harmonic measure is the same for any two edges of the same class. Thus the problem of determining the z parameters is reduced from a 48-dimensional problem to 4 dimensions. To obtain Figure **??** we ran a random walk from the origin and recorded the class of the edge it stopped on. From this we estimated the harmonic measure of each class and gave each edge from a given class the same harmonic measure. Compare these results with Figure 36, where we used many more random walks to try to estimate the harmonic measure of each individual side.

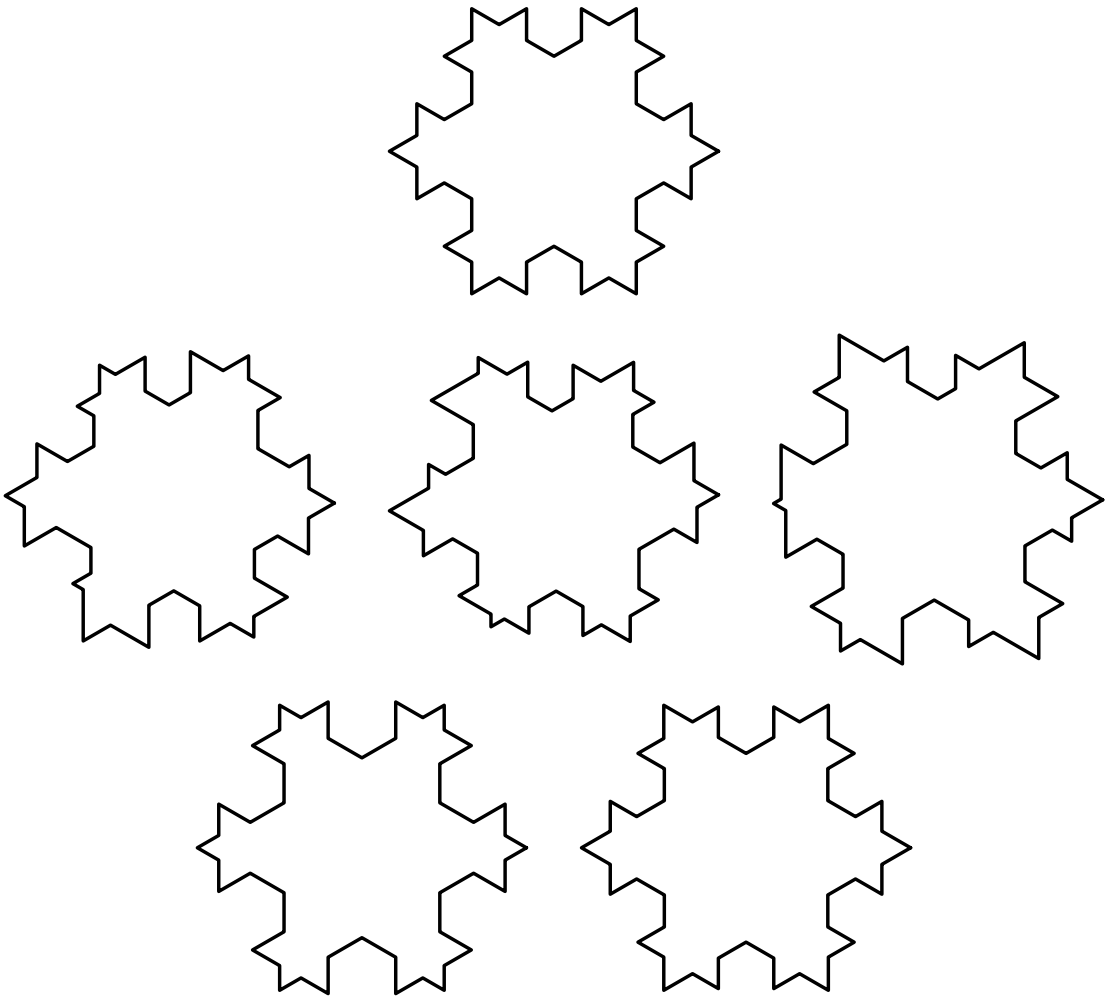


FIGURE 36. The top shows the target polygon: a second generation von Koch snowflake. The middle row shows three attempts to find the SC parameters using 1000 random walks where each of the 48 edges was considered separately. The bottom row uses the 12 fold symmetry of the polygon. The 48 edges are grouped into 4 collections; the number of hits in each collection is counted and divided by 12 to give the harmonic measure of each side. The left shows the result after 100 random walks and the right after 1000 random walks. See Figure 44 for a measure of how accurate these are.

7. The quasiconformal distance between polygons

When computing approximations for the z-parameters it would be nice to have a way of measuring how close our approximations are to correct parameters, assuming we can do this without actually knowing the correct parameters. Moreover, we should remember that there is more than one choice of "correct" answer, so that simply want

A standard way to measure the distance between sets is with the Hausdorff distance, which is defined as

$$d(E, F) = \inf\{\epsilon : E \subset F^\epsilon \text{ and } F \subset E^\epsilon\},$$

where $E^\epsilon = \{x : \text{dist}(x, E) < \epsilon\}$ is an ϵ neighborhood of E. However, for our purposes, we will often want to consider two regions that correspond under Euclidean similarities to be the same, so using the Hausdorff distance would be difficult to compute, even for polygons.

For polygons, the most obvious thing to do is compute the vector of sidelengths normalized by total length

$$\{\frac{|v_{k+1} - v_k|}{\sum_k |v_{k+1} - v_k|}\}$$

and compute the distance to the corresponding vector for the target polygon with respect to some norm on $\mathbb{R}^n$ (e.g., ℓ^1, ℓ^2, ℓ^∞).

The difficulty with this method is that when there are sides with many different length scales, the longer sides contribute much more to the distance than the short sides, e.g., see Figure **??**. Moreover, it is not clear that comparing the sides of the polygon in this way has a simple interpretation in terms of the geometry of the z-parameters.

We would like to have a distance based on comparing shapes of polygons, so that small distance means that the shapes are similar at all scales. Moreover, the distance should have a reasonable interpretation in terms of the z-parameters. The distance we will consider is based on the class of quasiconformal maps. Conformal maps have derivative maps that are Euclidean similarities. In particular the derivative maps send circle to circles. A K-quasiconformal map has a derivative which sends circles to ellipses of eccentricity at most K.

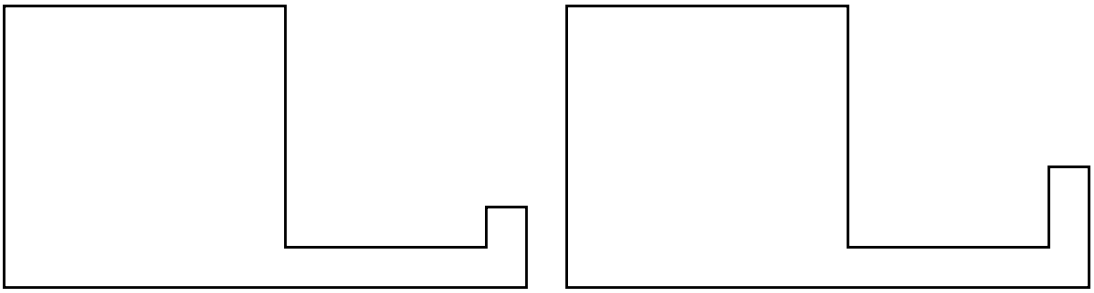


FIGURE 37. These two polygons only differ on a few short edges, so any distance based on normalized edge lengths with say these are close together. However, we want to define a distance that will place these about unit distance apart, recognizing that the "hooks" are significantly different shapes if we rescale them to unit size.

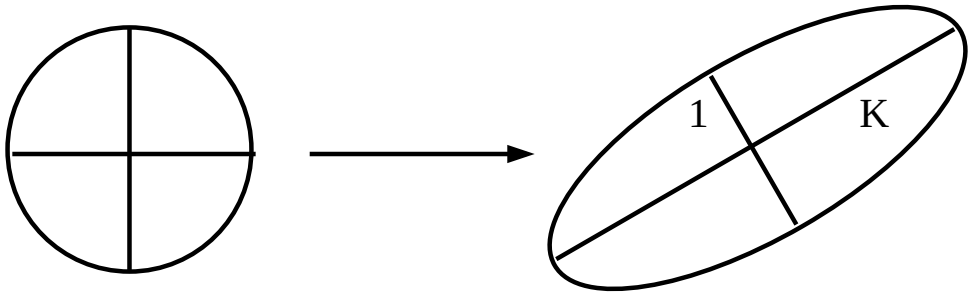


FIGURE 38. The derivative of a K-quasiconformal map sends circles to ellipses of eccentricy $\leq K$.

Suppose $f = (u, v)$ the derivative is given by

$$Df = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}.$$

The unit circle is mapped to an ellipse by this affine map and the eccentricity of the ellipse (the ratio of the length of major axis to the length of the minor axis) is $\sqrt{\lambda_1/\lambda_2}$ where λ_1, λ_2 are the eigenvalues of Df. The eigenvalues of this matrix are the roots of the equation

$$\det(DF - \lambda I) = \begin{pmatrix} u_x - \lambda & u_y \\ v_x & v_y - \lambda \end{pmatrix} = (u_x - \lambda)(v_y - \lambda) - u_y v_x = (u_x - \lambda)(v_y - \lambda) - u_y v_x.$$

We can use the quadratic formula to find the two roots $\lambda_1 \geq \lambda_2$ and then compute the eccentricy of the ellipse.

However, it is much more convenient to do this calculation in complex notation. Write

$$f_z = \frac{1}{2}(f_x - if_y) = \frac{1}{2}(u_x + v_y) + \frac{i}{2}(v_x - u_y),$$
$$f_{\bar{z}} = \frac{1}{2}(f_x + if_y) = \frac{1}{2}(u_x - v_y) + \frac{i}{2}(v_x + u_y).$$

Then

$$df = f_z dz + f_{\bar{z}} d\bar{z},$$

i.e., the tangent map to f is given by

$$z \to f_z z + f_{\bar{z}} \bar{z}.$$

If $|z| = 1$ then the largest its image can be is $|f_z| + |f_{\bar{z}}|$ and this occurs when the two components have the same argument (so the absolute values add), i.e., when

$$\arg(f_z z) = \arg(f_{\bar{z}} \bar{z}),$$
$$\arg(f_z) + \arg(z) = \arg(f_{\bar{z}}) - \arg z),$$
$$\arg z = \frac{1}{2}(\arg(f_{\bar{z}}) - \arg(f_z)).$$

Similarly, the minimal length the image of z can have is $|f_z| - |f_{\bar{z}}|$ when $f_z z$ points in the opposite direction from $f_{\bar{z}} \bar{z}$, and this occurs when

$$\arg(f_z z) = \pi + \arg(f_{\bar{z}} \bar{z}),$$
$$\arg z = \frac{\pi}{2} + \frac{1}{2}(\arg(f_{\bar{z}}) - \arg(f_z)).$$

Thus the image of the unit circle has major axis length

$$|f_z| + |f_{\bar{z}}|,$$

and the minor axis of length

$$|f_z| - |f_{\bar{z}}|,$$

and

$$(|f_z| + |f_{\bar{z}}|)(|f_z| - |f_{\bar{z}}|) = |f_z|^2 - |f_{\bar{z}}|^2 = u_x v_y - u_y v_x,$$

is the Jacobian of f. It is often convenient to write

$$\mu = \frac{K-1}{K+1} = f_{\bar{z}}/f_z,$$

and call this the dilatation of the map. Note that $|\mu| < 1$ if K is finite.

LEMMA 17. *Suppose T_1, T_2 are triangles with vertices $\{z_1, z_2, z_3\}$ and $\{w_1, w_2, w_3\}$ respectively (in the same orientation). Then the affine map $f : T_1 \to T_2$ defined by mapping $z_k \to w_k$ for $k = 1, 2, 3$ has complex dilatation*

$$\mu = \frac{a - b}{b - \bar{a}},$$

where $a = (z_3 - z_1)/(z_2 - z_1)$ and $b = (w_3 - w - 1)/(w_2 - w_1)$.

PROOF. We can map T_1 to the triangle T_1' with vertices $\{0, 1, a\}$ by the conformal map $z \to (z - z_1)/(z_2 - z_1)$ and map T_2 to the triangle T_2' with vertices $\{0, 1, b\}$ by $z \to (z - w_1)/(w_2 - w_1)$. Since these are conformal, they have complex dilatation 0. We can map T_1' to T_2' by the map of the form

$$z \to \alpha z + \beta \bar{z},$$

where we need $\alpha + \beta = 1$ and $\alpha a + \beta \bar{a} = b$. Solving gives $\beta = (a - b)/(a - \bar{a})$ and $\alpha = 1 - \beta = (b - \bar{a})/(a - \bar{a})$. Thus $\mu = \beta/\alpha = (a - b)/(b - \bar{a})$. □

Consider the map that stretches in the horizontal direction $f : (x, y) \to (ax, y)$ with $a > 0$. Then in complex notation this is

$$f(z) = \frac{a}{2}(z + \bar{z}) + \frac{1}{2}(z - \bar{z}) = \frac{a + 1}{2} z + \frac{a - 1}{2} \bar{z},$$

which has derivatives $f_z = \frac{1}{2}(a + i)$ and $f_{\bar{z}} = \frac{1}{2}(a - i)$. Thus

$$K = (|a + 1| + |a - 1|)/(|a + 1| - |a - 1|),$$

which equals a if $a \geq 1$ and equals $1/a$ if $0 < a < 1$. Thus K measures the amount of stretching.

If f is a C^1 map we define $K_f(z)$ and $\mu_f(z)$ applying the definitions above to the tangent map of f. A C^1 mapping on Ω is called a K-quasiconformal mapping if

$$\sup_{z \in \Omega} |K_f(z)| \leq K.$$

If $K = 1$, then the mapping is conformal. The function $\mu = \mu_f$ is called the Beltrami coefficient of f and satisfies the following composition laws:

$$\mu_{f^{-1}} \circ f = -(f_z / \overline{f_z})^2 \mu_f,$$

$$\mu_{g \circ f}(z) = (f_z(z)/\bar{f}_z(z)) \frac{\mu_g(f(z)) - \mu_f(z)}{1 - \mu_g(f(z)) \overline{\mu_f(z)}}.$$

We measure the distance between n-tuples, $\mathbf{z}, \mathbf{w} \in \mathbb{D}$, using the metric

$$d_{QC}(\mathbf{w}, \mathbf{z}) = \inf\{\log K : \exists\ K\text{-quasiconformal } h : \mathbb{D} \to \mathbb{D} \text{ such that } h(\mathbf{z}) = \mathbf{w}.\}$$

This metric is invariant under Möbius self-maps of the disk, which is natural, since we only expect to know the prevertices up to a Möbius transformation. Although the metric might seem a little awkward, we can often estimate this distance explicitly. In particular, when n-tuples are sets of z-parameters associated to two polygons we can estimate the distance between the n-tuples by finding maps between the polygons.

LEMMA 18. *If there is a K-quasiconformal map $\varphi : P_1 \to P_2$ sending vertices to vertices, then there is a K-quasiconformal map of the disk sending the SC-parameters for P_1 to the SC-parameters for P_2.*

PROOF. Take $f_2^{-1} \circ \varphi \circ f_1$ where f_k is the conformal map $\mathbb{D} \to P_k$ for $k = 1, 2$. □

Note that the composition of a quasiconformal and conformal map is again quasiconformal and with the same constant. In particular, if $f_1 : \mathbb{D} \to \Omega_1$ and $f_2 : \mathbb{D} \to \Omega_2$ are conformal maps, and $g : \Omega_1 \to \Omega_2$ is quasiconformal, then $G = f_2^{-1} \circ g \circ f_1 : \mathbb{D} \to \mathbb{D}$ is quasiconformal with the same constant as g. Consider the case then Ω_1, Ω_2 are bounded by polygons. Then any map between them which sends vertices to vertices corresponds to a map of the disk whose boundary extension sends one set of prevertices to the other. One simple case when there is an "obvious" vertex preserving map between the polygons is when the two polygons have equivalent triangulations. This means that there are cyclic labellings of the vertices of each polygon's vertices and triangulations of the polygons so that exactly the same set of triples of vertices are used. See Figure 5. In general, two n-gons need not have any equivalent triangulations, but we are mostly interested in the case when both polygons have the same set of interior angles and are close in some sense, so that we hope this does occur. (If one allows triangulations with Steiner points, i.e., points in the interior of the polygon, and not just the original vertices then any two n-gons have equivalent triangulations with at most $O(n)$ extra vertices. Moreover, one can add $O(n^2)$ Steiner points and obtain a triangulation that is equivalent to a certain triangulation depending on n but not on the particular polygon. See []).

Given two polygons with compatible triangulations we can explicitly compute the quasiconformal constant of the piecewise affine map which maps one triangulation to

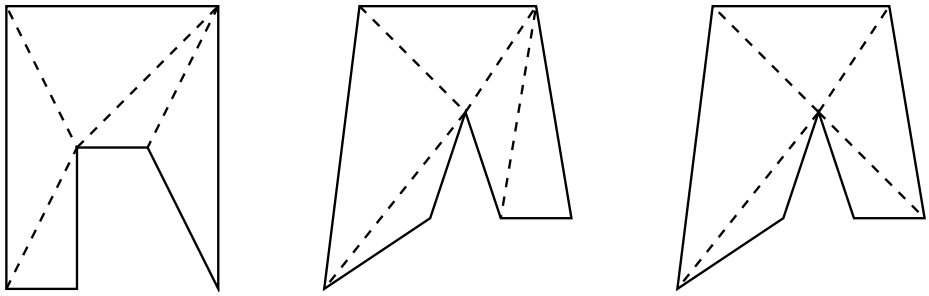


FIGURE 39. The left and center polygons have compatible triangulations but the one on the right is not compatible to either of these.

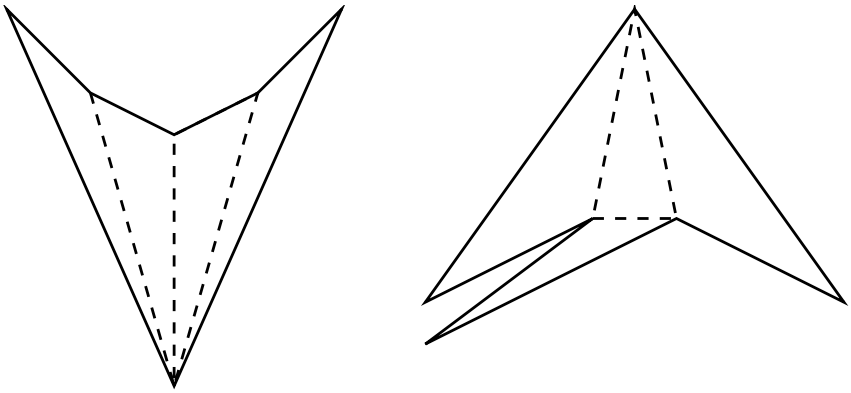


FIGURE 40. Each of these 8-gons has a single possible triangulation and they are not compatible. Thus general polygons need not have compatible triangulations if we do not allow Steiner points.

the other and this gives an upper bound for the best quasiconformal map sending Ω_1 to Ω_2 and preserving the vertices, and hence for the quasiconformal distance between the two sets of prevertices.

If P_1, P_2 are two polygons that have equivalent triangulations, then we can compute the quasiconstant for mapping each triangle for P_1 to the corresponding triangle for P_2. These individual affine maps form a global quasiconformal map from the interior of P_1 to the interior of P_2 whose quasiconstant is the maximum constant over all the triangle maps. See Figure 41.

One example of estimating the QC distance for two polygons is illustrated in Figure 41. A more interesting example is to consider the three polygons shown in Figure 5. This shows a target polygon on the left and two Schwarz-Christoffel images using 1000 and 10000 random walks to estimate the z-parameters. The polygon on

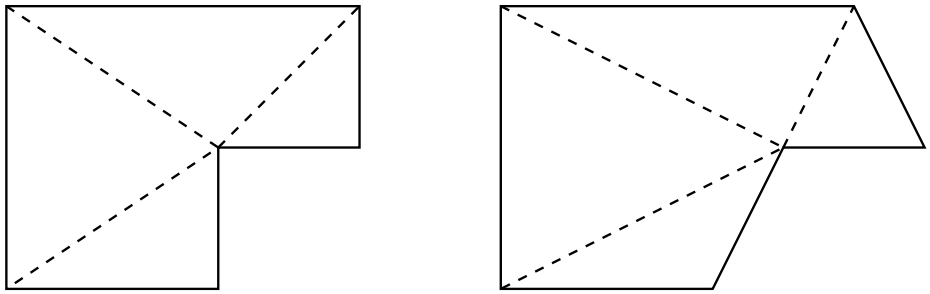


FIGURE 41. Two polygons with equivalent triangulations. The vertices for the first are $\{0, 3, 3+2i, 5+2i, 5+4i, 4i\}$, $\{0, 3, 4+2i, 6+2i, 5+4i, 4i\}$, and labeled counterclockwise starting at 0. The triangles (in terms of the vertex labels) are $\{1, 2, 3\}, \{1, 3, 6\}, \{3, 5, 6\}, \{3, 4, 5\}$. The distortion K for the four triangles is $1.64039, 1.33333, 1.64039, 1.64039$. The maximum of these is an upper bound for the QC-distance between the prevertices of the two polygons.

the right looks "better", i.e., closer to the target, and by considering affine maps between compatible triangulations we can make this more precise.

The idea of compatible triangulations is interesting in the context of conformal mappings in the following way. As noted above, two n-gons need not have compatible triangulations unless we allow Steiner points to be added. If we allow Steiner points, then any two n-gons have a compatible triangulations. Can we take corresponding triangles to be similarities, or close to similarities? For triangles which touch original vertices of the polygons, this is clearly impossible, but the Riemann mapping theorem implies that all triangles except those in an ϵ neighborhood of the original vertices can be taken to almost similar, i.e., corresponding triangles can be mapped to each other $1 + \epsilon$ quasiconformal affine maps. Conversely, shrinking ϵ to zero and taking the limit of such piecewise affine maps gives a conformal map in the limit and proves the Riemann mapping theorem.

A couple of examples will serve to show that the metrics based on side length vectors and on quasiconformal mappings can be very different. In Figure 42 we have shown domains that are far apart in the side length sense but are quasiconformally close. Each is a square with a long narrow corridor attached, the second one being exactly half the size in both dimensions. Because each side of the corridor is half as long in the second domain, the normalized side length vectors differ by about $1/12$ in two coordinates, and hence the distance will be large. However, these two domains

are close in a quasiconforml sense. In each domain consider the half-annuli bounded by the dashed lines. We clam these can be mapped to each other by a quasiconformal map with small constant (at least if the inner and outer radii have large ratio). This is because the map

$$f_\alpha : z \to z|z|^{\alpha-1},$$

is quasiconformal on the plane with constant $K = \max(\alpha, \frac{1}{\alpha})$. This can be verified by a computation of its partial derivatives, but it is simpler to restrict to the upper half-plane, map the half-plane to an infinite horizontal strip by the (conformal) map $z \to \log z$ and note that our map is conjugated to $(x, y) \to (\alpha x, y)$ which is clearly K-quasiconformal. If we scale our first half-annuli so the outer radii is 1 and the inner is $r \ll 1$ then its image under f_α is a half-annuli with outer radius 1 and inner radius $R = r^\alpha = r/2$, if $\alpha = 1 - \frac{2}{\log r}$. We then extend out map to be conformal and linear on the remaining pieces of the domain, and obtain a K-quasiconformal map between the domain with K as close to 1 as we choose, depending only on the width of the corridor.

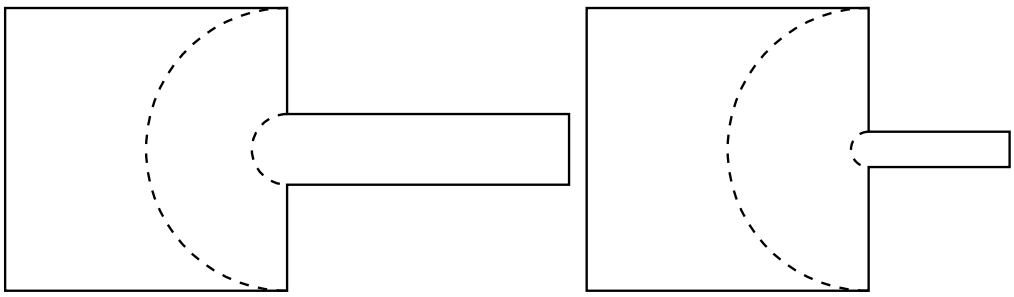


FIGURE 42. These two domain are close in the QC metric but far in the vector of side lengths metrics. The are far in the latter metric since the thin channels have significantly different lengths. They are close in the QC metric since the half-annulus regions between the dashed curves can be mapped to each other with small QC norm if the ratio between the inner and outer radius is large enough. Outside these regions we use linear conformal maps.

In Figure 43 we show two domains which are close in the vector sense but not in the QC sense. These domains are squares with two vertical slits removed, one attached to the top edge and the other to the bottom edge and each with length about 3/4's of the side length of the box and about ϵ apart. In the first domain

the slit attached the top is to the left of the other slit and in the second it is to the right. In terms of side lengths, this only requires changes of about size ϵ in the segments along the top and bottom of the square, so these domain are close in this sense. However, any homeomorphism of the interiors must have large distortion. For example, we draw a dashed vertical line in the left picture; its image under such a homeomorphism must look something like the dashed curve on the right, which requires a large QC constant.

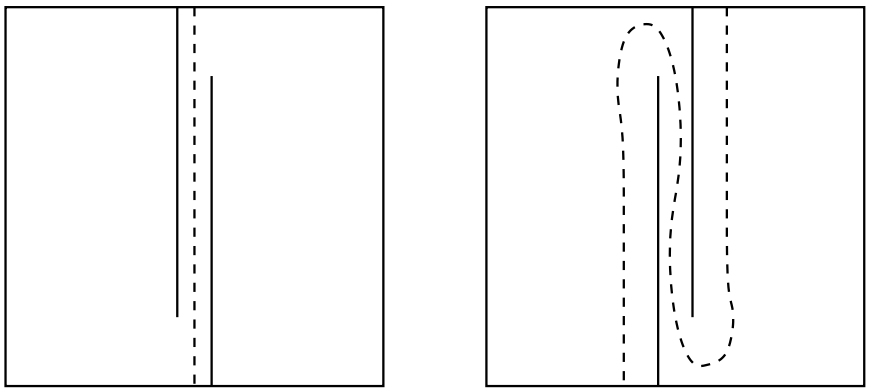


FIGURE 43. These are two domains which are close in the vector of side lengths sense, but not in the QC sense. Any homeomorphism of the interior which respects vertices must map the dashed vertical line on the left to something like the dashed curve on the right, which requires large distortion.

How can we bound the QC distance from below more explicitly? This can be done using conformal modulus, which will be discussed in a later chapter.

8. Schwarz-Christoffel iterations and Davis's method

Suppose Ω has polygonal boundary and $f : \mathbb{D} \to \Omega$ is conformal. The Schwarz-Christoffel formula (we abbreviate to "SC-formula" below) says

$$f(z) = A + C \int^z \prod_{k=1}^{n} (1 - \frac{w}{z_k})^{\alpha_k - 1} dw,$$

where $\alpha\pi = \{\alpha_1\pi, \ldots, \alpha_n\pi\}$, are the interior angles at the vertices $\mathbf{v} = \{v_1, \ldots, v_n\}$, and $\mathbf{z} = \{z_1, \ldots, z_n\} = f^{-1}(\mathbf{v})$ are the conformal preimages of the vertices (also know as the SC-parameters). For a fixed α, we can think of the formula as defining a map S from n-tuples in $\mathbb{T}$ to polygons (possibly self-overlapping). In fact, Möbius equivalent

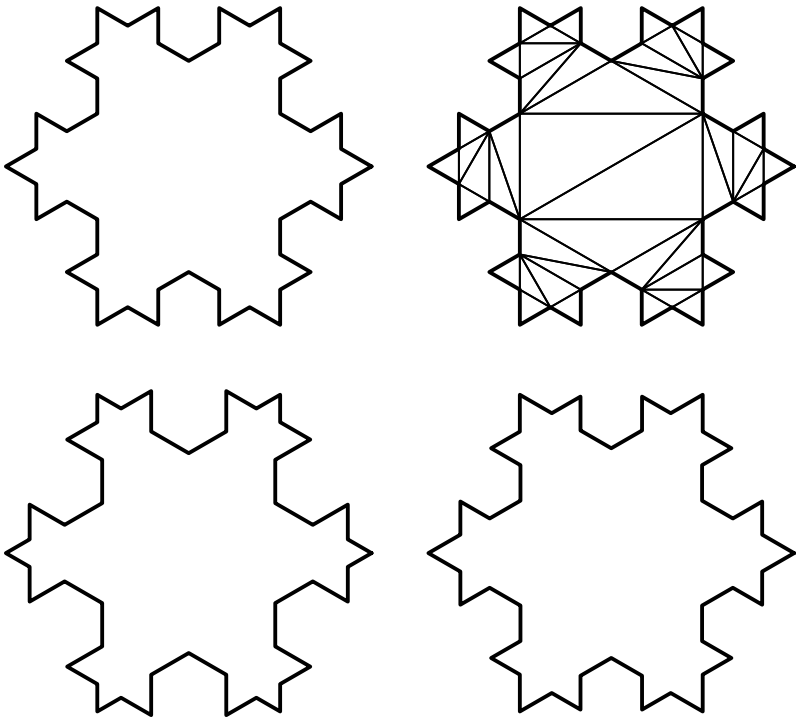


FIGURE 44. These polygons were drawn in Figure 36. The top show the target polygon and a triangluation. The bottom shows two attemps to calculate the SC-parameters using random walks and symmetry. The QC distance in the first case is 1.63525 (100 walks) and in the second is 1.11956 (1000 walks), justifying the idea that the second picture looks "better" than the first.

n-tuples give Euclidean similar polygons, so it is convenient to think of S as a map from $\mathbb{T}^n_*$ (n-tuples of distinct points on $\mathbb{T}$ modulo Möbius transformations) to P^n_* (complex n-tuples modulo similarities). We can identify $\mathbb{T}^n_* = \mathbb{R}^{n-3}$ as follows: fix a combinatorial triangulation of the n points, and for each pair of adjacent triangles let ρ_k be the cross ratio of the four vertices. This is a positive real number since the points lie on $\mathbb{T}$ (and if we take the correct ordering), so $\log \rho_k \in \mathbb{R}$. The original n-tuple (unique up to Möbius transformations) can easily be recovered from the $n-3$ values of $\log \rho_k$, so $\mathbb{T}^n_* = \mathbb{R}^{n-3}$.

Suppose we have a explicit way of guessing the SC-parameters for a given polygon, i.e., a map $G : P^n_* \to \mathbb{T}^n_* = \mathbb{R}^{n-g}$. Then $F = G \circ S$ gives a map $\mathbb{R}^{n-3} \to \mathbb{R}^{n-1}$. The

desired SC-parameters for P, $\mathbf{z}_*$, are a solution of $F(\mathbf{z}) = \mathbf{z}_0 = G(P)$ and hence are a fixed point of the iteration

$$\mathbf{z}_{k+1} = \mathbf{z}_k - A^{-1}(F(\mathbf{z}_k) - \mathbf{z}_0). \tag{6}$$

We call this an SC-iteration. If A is the derivative DF of F, this iteration is Newton's method for $n-3$ real variables. If we don't know DF explicitly, can take a discrete approximation using $n-3$ evaluations of F; we call this the "full iteration". If DF is close to the identity, then taking A to be the identity may also work and is much faster; this we call the "simple iteration". A compromise between these two extremes is to start A as the identity and to use Broyden updates at each step; this is called the "short-cut iteration" (A Broyden update multiplies A by a rank one matrix at each step, chosen to optimize the approximation to DF given the evaluations of F made so far. This method converges more slowly per iteration than the full iteration, but each iteration is faster to perform and it often beats full and simple iterations in practice).

One of the simplest such methods is due to Davis, which has the additional advantage of taking advantage of the geometry of the domain in a straightforward way. Suppose we are given points $\{z_1, \ldots, z_n\}$ on the unit circle. Compute an image polygon using the Schwarz-Christoffel formula with these parameters (and the known angles) and compare the side lengths of this polygon with the desired polygon. If a side is too short, the corresponding parameter values are moved apart in the next iteration and conversely. More precisely, if $\{z_1^k, \ldots, z_n^k\}$ is the current guess, and the image polygon has vertices $\{v_1^k, \ldots, v_n^k\}$ we define the next set of parameter guesses as

$$|z_k^{j+1} - z_{j-1}^{k+1}| = K|z_j^k - z_{j-1}^k| \frac{|v_j - v_{j-1}|}{|v_j^k - v_{j-1}^k|},$$

for $j = 0, \ldots, n$ where

$$K = 2\pi [\sum_j |z_j^k - z_{j-1}^k| \frac{|v_j - v_{j-1}|}{|v_j^k - v_{j-1}^k|}]^{-1},$$

is a normalizing constant (to make sure the new spacings add up to 2π) and $\mathbf{v} = \{v_0, \ldots, v_n\}$ are the vertices of the target polygon. An example of Davis' method is shown in Figure 45. Further details of the first ten steps of the iteration are given in Tables **??** and **??**.

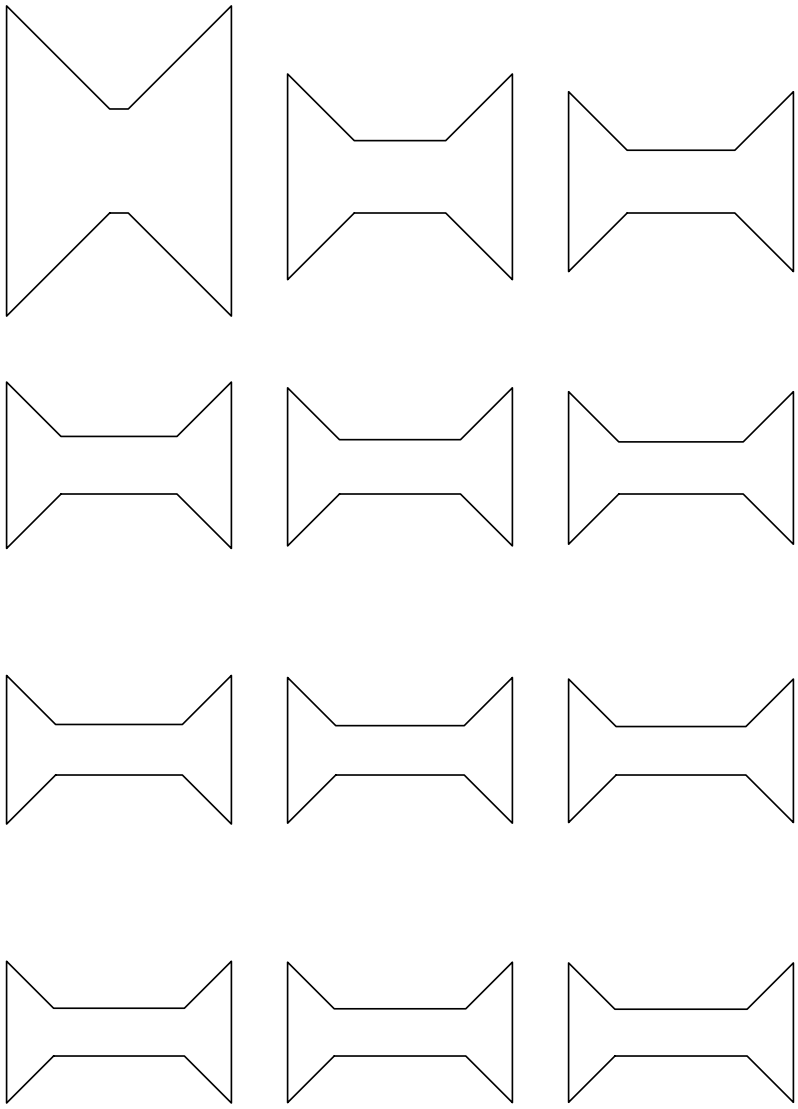


FIGURE 45. The first 12 iterations of Davis' method (including the first step where we assume equidistributed parameters). The upperbounds for the QC errors (for these iterations) obtained by triangulation are 16.7817, 2.37323, 1.74869, 1.4896, 1.34707, 1.25739, 1.19638, 1.15273, 1.1204, 1.09585, 1.07687, 1.06199

The method works in practice in many cases but is known to sometimes diverge even locally [**?**]. Davis' method assumes that increasing side length corresponds to increasing harmonic measure. However, in some examples, this is not case. See Figure 46. The edge on the far right makes angle $< \pi/2$ with both the adjacent edges. In this case, lengthening this side gives a polygon that strictly contains the first one, and the new right edge clearly has less harmonic measure. Thus we expect Davis' method to diverge in this case. However, when we do the experiment with this polygon and start iterating we get a kind of degeneration where all the z-parameters begin to cluster around one point of the unit circle. This corresponds to the origin being mapped to a point of the polygon which is tending towards the boundary. The behavior of the parameters under iteration is shown in Table 2.

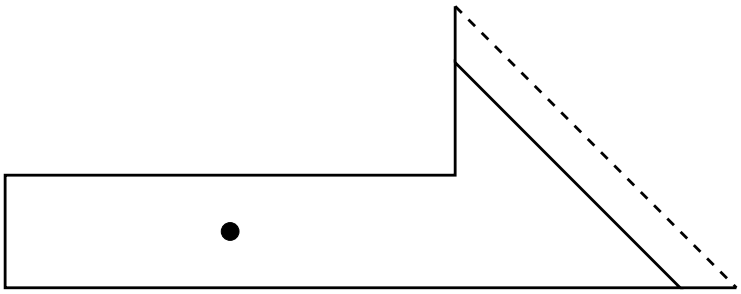


FIGURE 46. Lengthening the edge on the far right of the polygon decreases the harmonic measure of that edge from the given base point, and hence decreases the length of its conformal preimage on the circle (assuming the base point is mapped to the origin). Thus Davis' method will diverge from the correct answer given a starting point arbitrarily close to it. The example is taken from [**?**]. However, renormalizing seems to eliminate the divergence.

Davis' method is used in [**?**] by Banjai and Trefethen to give a $O(n)$ method for finding the prevertices that is practical for tens of thousands of vertices (the bound, however is an average case analysis, not a uniform estimate for all polygons).

The degeneration can be prevented by renormalizing the parameters each time we iterate. Choose two adjacent sides (say the first and second) and use a Möbius transformation of the disk to itself to map the corresponding three parameters to $1, i, -1$. Thus we are choosing a "center" for the polygon (i.e., the image of zero under the conformal map) from which these two sides each has harmonic measure

1.01811	0.412703	0.41663	4.06205	0.373695
0.721234	0.154016	0.181627	5.11414	0.112163
0.513933	0.0634117	0.0800658	5.58571	0.0400608
0.379569	0.0288556	0.0367562	5.82124	0.0167623
0.292544	0.0145189	0.0181267	5.95	0.00799366
0.234438	0.00801887	0.00971116	6.02679	0.00423033
0.193772	0.00479989	0.00563333	6.07655	0.00242752
0.163696	0.00306846	0.00350138	6.11144	0.00148175
0.140138	0.00206525	0.00230183	6.13773	0.000947121
0.120658	0.00144461	0.00157994	6.15888	0.000625846
0.103785	0.00103819	0.00111873	6.17682	0.000422947
0.0886404	0.000758795	0.000808338	6.19269	0.000289599
0.0747276	0.000558768	0.000590043	6.20711	0.000199199
0.0618022	0.000410843	0.00043094	6.2204	0.000136497
0.0497937	0.000298814	0.000311833	6.23269	0.0000923607
0.0387541	0.000212751	0.000221155	6.24394	0.0000610976
0.0288182	0.000146421	0.000151751	6.25403	0.0000390293
0.020169	0.0000958193	0.0000990789	6.2628	0.0000236892
0.0129987	0.0000582694	0.0000601445	6.27005	0.0000133538
0.00746055	0.0000318117	0.00003279	6.27565	0.000006755110

TABLE 2. The evolution of the SC parameters for Davis' method for the polygon in Figure 46. The iteration is started at equidistributed points, but seems to converge to a situation where all the parameters are clustered around a single point. This corresponds to the origin being mapped to a point in the image polygon which is closer and closer to the boundary. We can use Möbius transformations to renormalize the parameters by sending three of them to any three points we want. This is discussed below.

1/4. Such a point must remain in a compact region of the polygon (at least as long as the polygon itself stays in a compact set). This is because each of these two sides has harmonic measure 1/4 from this point and the complement has measure 1/2. There is a unique point of the domain at which this occurs and so the normalization prevents the kind of degeneration described above. I do not know if adding this normalization causes Davis' method to aways converge to the correct answer eventually.

The vertices of the polygon are

$$0, 2, 6, 4 + 2I, 4 + I, I$$

Note that we have placed a vertex of angle π on the bottom edge. We do this and normalize the edges $[0,2],[2,6]$ to have harmonic measure $1/3$ so that the center of the polygon is near the center of the long channel. This helps prevent the harmonic measure of any of the sides from getting two small. Fifty iterations of the normalized iteration gives harmonic measure (before renormalizing) of

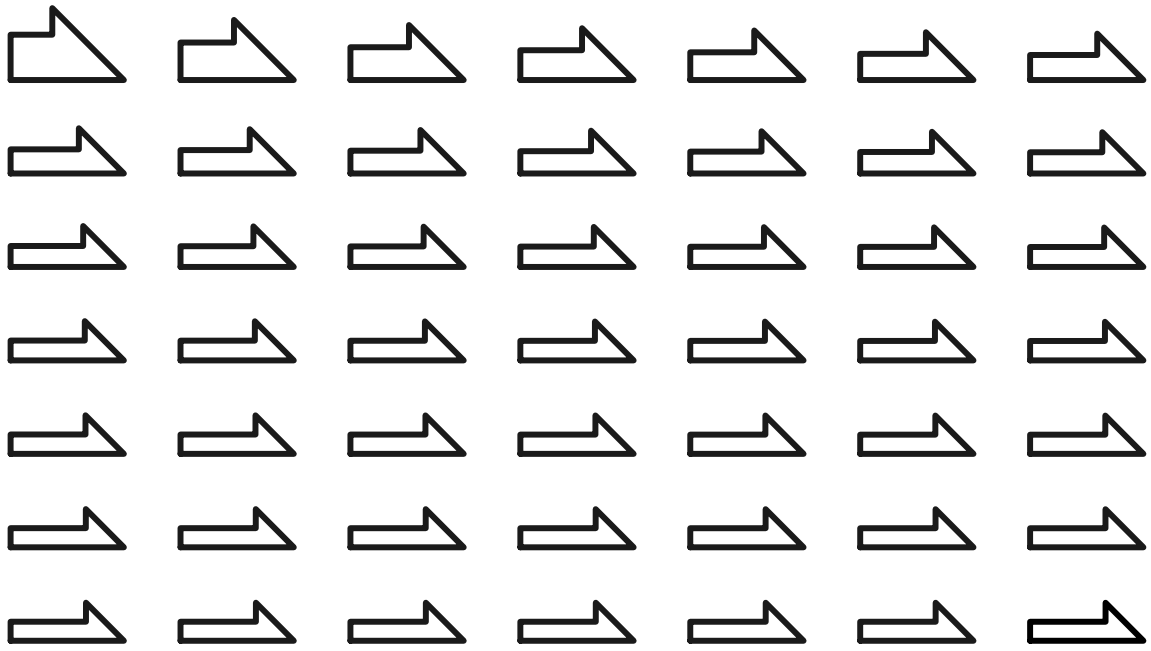
$$0.333411, 0.335319, 0.000186381, 0.000180184, 0.328768, 0.00213578$$

relative sidelengths of

$$0.134715, 0.267896, 0.190897, 0.0674928, 0.271171, 0.0678286$$

whereas the true relative sidelengths are

$$0.134876, 0.269752, 0.190744, 0.067438, 0.269752, 0.067438$$



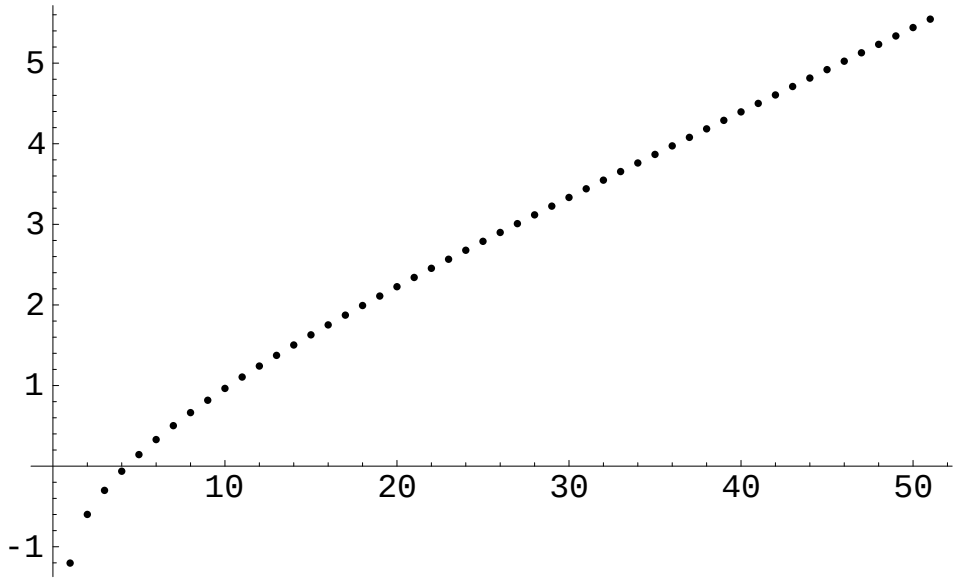


FIGURE 47. 49 iterations of Davis's method for the polygon in Figure 46 (with renormalizations) starting from equally spaced parameters. On the bottom is the graph of $-\log(K-1)$, which shows increasing accuracy.

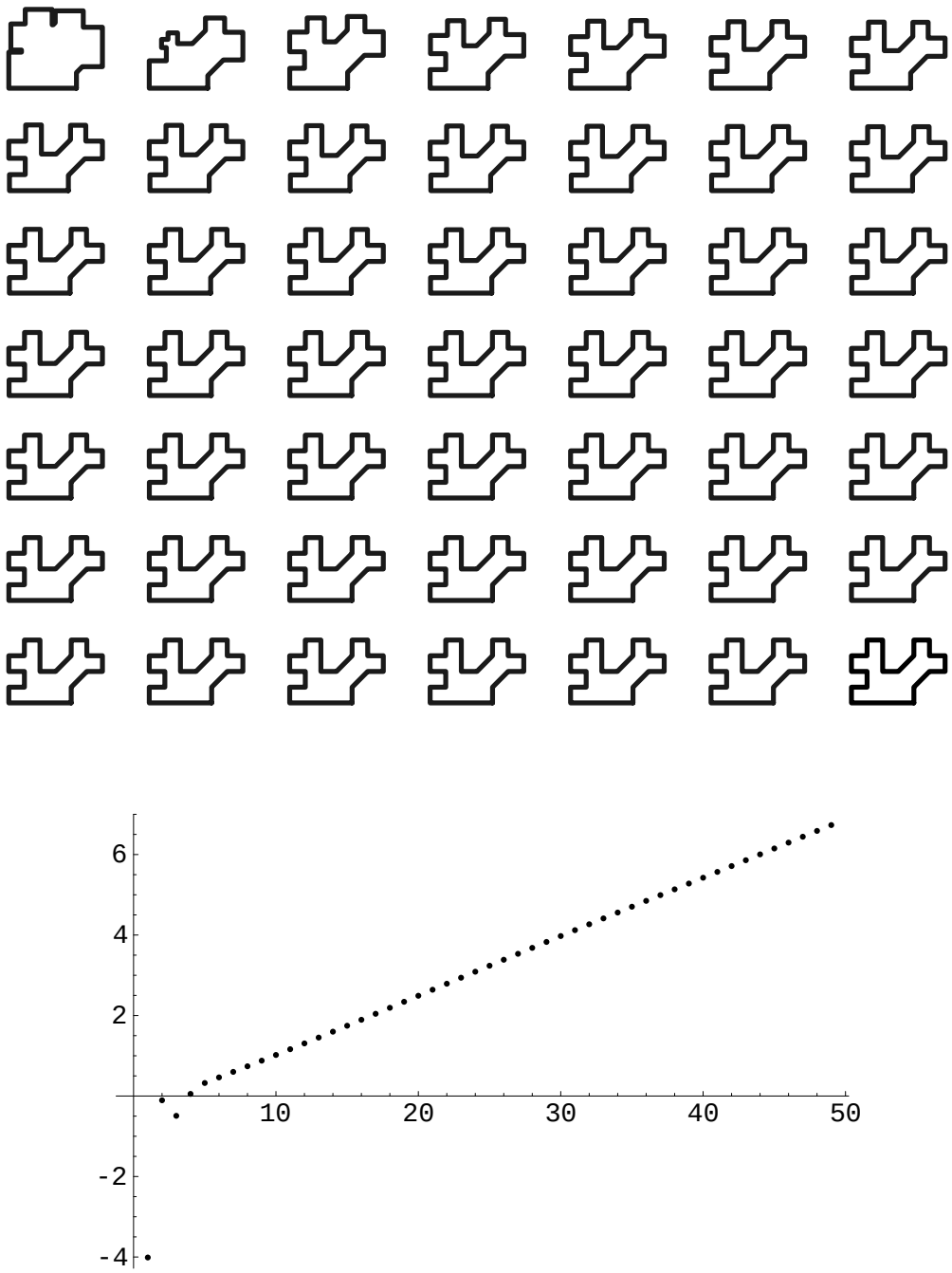


FIGURE 48. 49 iterations of Davis's method (with renormalizations) starting from equally spaced parameters. On the bottom is the graph of $-\log(K-1)$, which shows increasing accuracy.

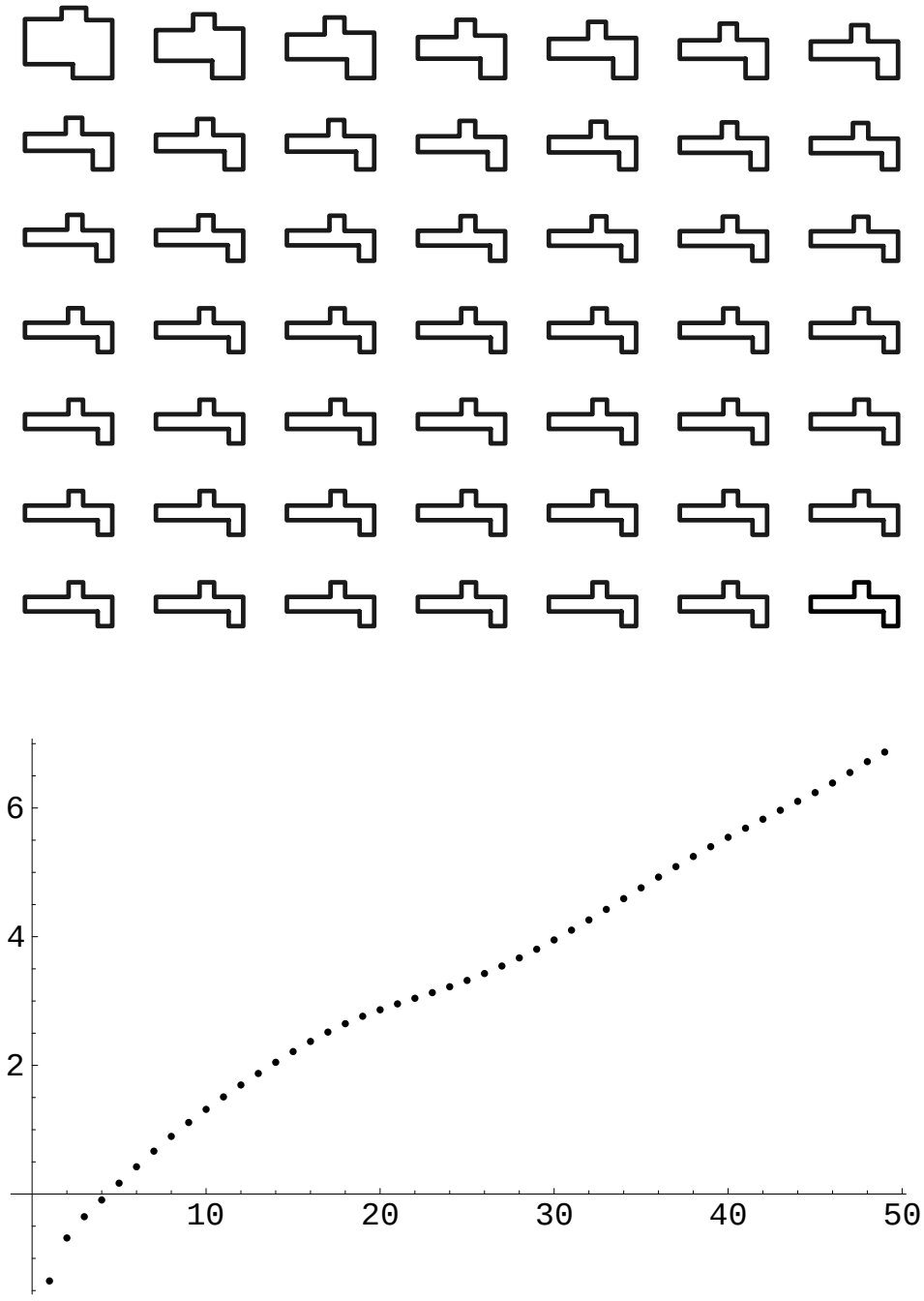


FIGURE 49. 49 iterations of Davis's method for another polygon.

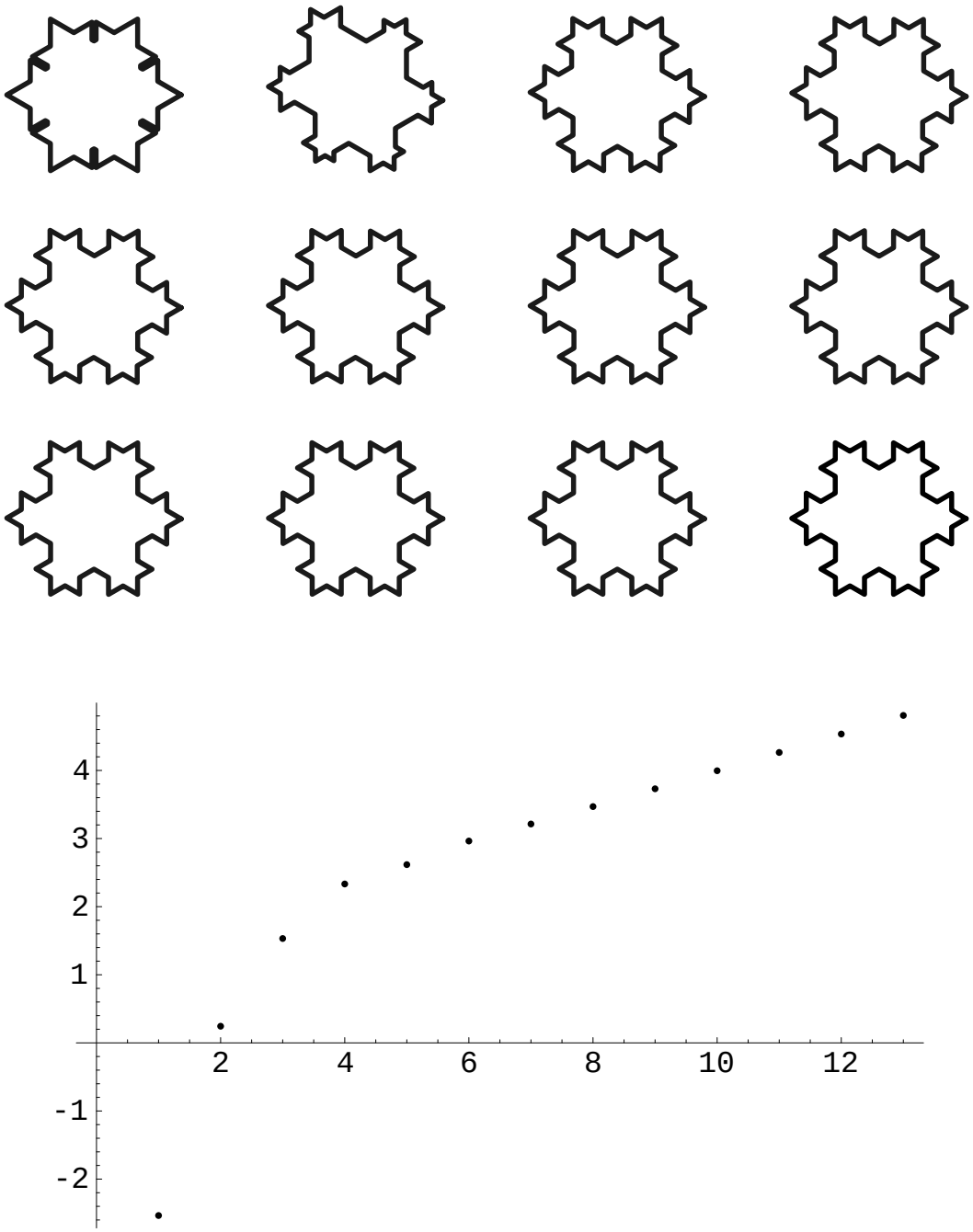


FIGURE 50. 12 iterations of Davis's method for the second generations von Koch snowflake.

CHAPTER 2

The Riemann mapping theorem

In this chapter we review basic results from one complex variable and potential theory and use them to give a proof of the Riemann mapping theorem (already quoted in the first chapter).

1. The hyperbolic metric

The hyperbolic metric on $\mathbb{D}$ is given by $d\rho_{\mathbb{D}} = 2|dz|/(1-|z|^2)$. This means that the hyperbolic length of a rectifiable curve γ in $\mathbb{D}$ is defined as

$$\ell_\rho(\gamma) = \int_\gamma \frac{2|dz|}{1-|z|^2},$$

and the hyperbolic distance between two points $z, w \in \mathbb{D}$ is the infimum of the lengths of paths connecting them (we shall see shortly that there is an explicit formula for this distance in terms of z and w).

A geodesic is a shortest possible path between points. Geodesics for the hyperbolic metric are circles orthogonal to the boundary. The orientation preserving isometries are exactly the Möbius transformations which preserve the disk, which all have the form $z \to e^{i\theta}(z-a)/(1-\bar{a}z)$, for some $\theta \in \mathbb{R}$ and $a \in \mathbb{D}$. The hyperbolic metric ρ_Ω on a simply connected domain Ω (or Riemann surface) is defined by transferring the metric on the disk to Ω by the Riemann map. We will sometimes write ρ for any hyperbolic metric when the domain is clear from context.

On the disk it is convenient to define the pseudo-hyperbolic metric

$$\rho(z,w) = |\frac{z-w}{1-\bar{w}z}|.$$

The hyperbolic metric between two points can then be expressed as

$$\psi(w,z) = \log \frac{1+\rho(w,z)}{1-\rho(w,z)}.$$

On the upper half-plane the corresponding function is

$$\rho(z,w) = |\frac{z-w}{w-\bar{z}}|,$$

and ψ is given as before. A hyperbolic ball in the disk is also a Euclidean ball, but the hyperbolic and Euclidean centers are different (unless they are both the origin).

The orientation preserving isometries of the hyperbolic disk are exactly the Möbius transformations that map the disk to itself. All of these have the form

$$e^{i\theta}\frac{z-a}{1-\bar{a}z},$$

where θ is real and $a \in \mathbb{D}$.

Recall the sine and cosine rules for hyperbolic geometry (e.g., see page 148 of Beardon's book [**?**]). Let T denote a hyperbolic triangle with angles α, β, γ and opposite side lengths denoted by a, b, c. See Figure 1. Then we have the Sine Rule,

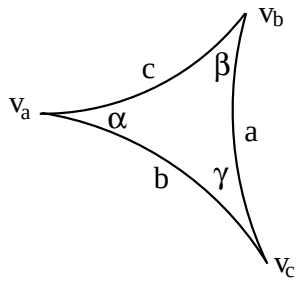


FIGURE 1. Definitions of a, b, c and α, β, γ

$$\frac{\sinh a}{\sin \alpha} = \frac{\sinh b}{\sin \beta} = \frac{\sinh c}{\sin \gamma} \tag{7}$$

the First Cosine Rule,

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos \gamma \tag{8}$$

and the Second Cosine Rule

$$\cosh c = \frac{\cos \alpha \cos \beta + \cos \gamma}{\sin \alpha \sin \beta} \tag{9}$$

LEMMA 19. *Möbius transformations of $\mathbb{D}$ to itself are isometries of the hyperbolic metric.*

PROOF. When f is a Möbius transformation of the disk we have

$$f(z) = \frac{z-a}{1-\bar{a}z},$$

$$f'(z) = \frac{1-|a|^2}{(1-\bar{a}z)^2}.$$

We define the hyperbolic gradient of f as

$$|\nabla_H f|(z) = \lim_{z \to w} \frac{\rho(f(z), f(w))}{\rho(z,w)}.$$

Thus

$$\begin{aligned}
|\nabla_H f(z)| &= \frac{1-|a|^2}{(1-\bar{a}z)^2} \frac{1-|z|^2}{1-|f(z)|^2} \\
&= \frac{1-|a|^2}{(1-\bar{a}z)^2} \frac{1-|z|^2}{1-|\frac{z-a}{1-\bar{a}z}|^2} \\
&= \frac{(1-|a|^2)(1-|z|^2)}{|1-\bar{a}z|^2 - |z-a|^2} \\
&= \frac{(1-|a|^2)(1-|z|^2)}{(1-\bar{a}z)(1-a\bar{z}) - (z-a)(\bar{z}-\bar{a})} \\
&= \frac{(1-|a|^2)(1-|z|^2)}{(1-\bar{a}z - a\bar{z} + |az|^2) - (|z|^2 - a\bar{z} - z\bar{a} + |a|^2)} \\
&= \frac{(1-|a|^2)(1-|z|^2)}{(1+|az|^2 - |z|^2 - |a|^2)} \\
&= 1.
\end{aligned}$$

Note that

$$\ell_\rho(f(\gamma)) \leq \int_\gamma |\nabla_H f|(z) \frac{|dz|}{1-|z|^2}.$$

Thus Möbius transformations multiply hyperbolic length by at most one. Since the inverse also has this property, we see that Möbius transformation preserve hyperbolic length. □

Simply connected, proper subdomains of the plane inherit a hyperbolic metric from the unit disk via the Riemann map. If $\varphi : \mathbb{D} \to \Omega$ is conformal and $w = \varphi(z)$ then $\rho_\Omega(w_1, w_2) = \rho_{\mathbb{D}}(z_1, z_2)$ defines the hyperbolic metric on Ω and is independent of the particular choice of φ. It is often convenient to estimate ρ_Ω in terms of the

more geometric "quasi-hyperbolic" metric on Ω which is defined as

$$\tilde{\rho}(w_1, w_2) = \inf \int_{w_1}^{w_2} \frac{|dw|}{\operatorname{dist}(w, \partial\Omega)},$$

where the infimum is over all arcs in Ω joining w_1 to w_2.

2. Schwarz's lemma

We start with the maximum prinicple:

LEMMA 20. *If f is holomorphic on the unit disk and $r < 1$, then*

$$\max_{|z|\leq r} |f(z)| \leq \max_{|z|=r} |f(z)|.$$

PROOF. By the Cauchy integral formula

$$|f(z)| = |\frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w - z} dw| \leq \int_0^{2\pi} |f(re^{i\theta})| \frac{d\theta}{2\pi r} \leq \max_{|z|=r} |f(z)|.$$

□

As a consequence, we seen that if f is holomorphic on $\mathbb{D}$ and

$$\limsup_{|z| \nearrow 1} |f(z)| \leq M,$$

then $|f(z)| \leq M$ on all of $\mathbb{D}$. We use the "limsup" since we don't know if f has continuous boundary values. If it does then we can just write

$$\sup_{\mathbb{D}} |f(z)| \leq \sup_{\mathbb{T}} |f(z)|.$$

Perhaps the most important fact relating the hyperbolic metric and holomorphic functions is that a holomorphic map $f : \mathbb{D} \to \mathbb{D}$ is a contraction of the hyperbolic metric, with equality iff if is a Möbius transformation. This fact is usually presented as:

LEMMA 21 (Schwarz's Lemma). *If $f : \mathbb{D} \to \mathbb{D}$ is holomorphic and $f(0) = 0$ then $|f'(0)| \leq 1$ with equality iff f is a rotation. Moreover, $|f(z)| \leq |z|$ for all $|z| < 1$, with equality for $z \neq 0$ iff f is a rotation.*

PROOF. Define $g(z) = f(z)/z$ for $z \neq 0$ and $g(0) = f'(0)$. This is a holomorphic function since if $f(z) = \sum a_n z^n$ then $a_0 = 0$ and so $g(z) = \sum a_n z^{n-1}$ has a convergent power series expansion. Since $\max_{|z|=r} |g(z)| \leq \frac{1}{r} \max_{|z|=r} |f| \leq \frac{1}{r}$. By the maximum principle $|g| \leq \frac{1}{r}$ on $\{|z| < r\}$. Taking $r \nearrow 1$ shows $|g| \leq 1$ on $\mathbb{D}$ and equality

anywhere implies g is constant. Thus $|f(z)| \leq |z|$ and $|f'(0)| = |g(0)| \leq 1$ and equality implies f is a rotation. □

In terms of the hyperbolic metric this says that

$$\psi(f(0), f(z)) = \psi(0, f(z)) \leq \psi(0, z),$$

which shows the hyperbolic distance from 0 to any point is non-increasing. For an arbitrary holomorphic self-map of the disk f and any point $w \in \mathbb{D}$ we can always choose Möbius transformations τ, σ so that $\tau(0) = w$ and $\sigma(f(w)) = 0$, so that $\sigma \circ f \circ \tau(0) = 0$. Since Möbius transformations are hyperbolic isometries, this shows

COROLLARY 22. *If $f : \mathbb{D} \to \mathbb{D}$ is a holomorphic then $\psi(f(w), f(z)) \leq \psi(w, z)$.*

There are numerous generalizations of Schwarz's lemma in the liturature. For example, replacing the hyperbolic metric by the pseudohyperbolic metric gives

$$\frac{|f(z) - f(w)|}{1 - |f(z)f(w)|} \leq \frac{|z - w|}{1 - |zw|}.$$

3. Square roots and logarithms

One of the more technical aspects of proving the Riemann mapping theorem is to make use of the hypothesis that the domain Ω is simply connected. Simply connected means that any closed curve γ in Ω can be continuously deformed to a point inside Ω. However, the property that we actually use is that if Ω is simply connected, then any non-vanishing, continuous complex valued function f on Ω has a continuous square root, i.e., there is a function g so that $g^2 = f$. The fact this holds for simply connected domains is result about covering spaces.

We will give the details below, but the basic idea is as follows. Let $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Then $z \to z^2$ is a covering map of $\mathbb{C}^*$ to itself. If $f : \Omega \to \mathbb{C}^*$ continuous and Ω is simply connected, then the theory of covering spaces says f has a "lift' $g : \Omega \to \mathbb{C}$ so that f equals g followed by the covering map $\mathbb{C} \to \mathbb{C}$, i.e., $f = g^2$.

In a similar way the map $z \to e^z$ is a covering map from $\mathbb{C}$ to $\mathbb{C}^*$ so for any map $f : \Omega \to \mathbb{C}^*$ there is a lifted map $g : \Omega \to \mathbb{C}$ so that $f = e^g$ (if Ω is simply connected). This g is a branch of $\log f$. Note that the imaginary part of g defines a continuous branch of $\arg(f)$ on Ω.

Now for the details. let $p : E \to B$ be continuous and surjective. An open set $U \subset B$ is evenly covered if the inverse image $p^{-1}(U)$ can be written as a disjoint union of sets V_α so that p restricted to each V_α is a homeomorphism onto U. If every point b of B has a neighborhood U that is evenly covered by p, then p is called a covering map. The reader should check that the maps described above are indeed covering maps.

A space X is simply connected if it is path connected and if its fundamental group is trivial, i.e., every closed loop in X can be homotoped to a point.

LEMMA 23 ([**?**], Lemma 8.4.1). *Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. Any path $f[0,1] \to B$ beginning at b_0 has a unique lift to a path $\tilde{f}$ in E beginning at e_0.*

LEMMA 24 (Exercise 8.4.12(a), [**?**]). *Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. Let $f : Y \to B$ be continuous with $f(y_0) = b_0$. If Y is locally path connected and simply connected then f can be lifted uniquely to a continuous map $\tilde{f} : (Y, y_0), \to (E, e_0)$.*

One of the uses we make of these topological facts is the following result that allows us to reduce the Riemann mapping theorem to the case of bounded domains.

LEMMA 25. *Any simply connected planar domain, except for the plane itself, can be conformally mapped to a bounded domain.*

PROOF. If the domain Ω is bounded, there is nothing to do. If Ω. omits a disk $D(x, r)$ then the map $z \to 1/(z - x)$ conformal maps Ω to a bounded domain. Otherwise, translate the domain so that 0 is on the boundary and consider a continuous branch of $\sqrt{z}$. The image is a 1-1, holomorphic image of Ω, but does not contain both a point and its negative. Since the image does contain some open ball, it also omits an open ball and hence can be mapped to a bounded domain by the previous case.

□

4. Morera's theorem and uniform convergence

THEOREM 26 (Morera's theorem). *If f is defined on a disk $D = D(0, r)$ and the integral $\int_T f(z)dz = 0$ for every triangle in D, then f is holomorphic in D.*

PROOF. Define $F(z) = \int_0^z f(\zeta)d\zeta$, where the integral is over the line segment from 0 to z. Consider the triangle with vertices $0, z, w$ with w close to z. Since the integral of f around this triangle is zero,

$$F(z) - F(w) = \int_0^z f(\zeta)d\zeta - \int_0^w f(\zeta)d\zeta = \int_w^z f(\zeta)d\zeta.$$

Since f is continuous, for any $\epsilon > 0$ we can choose $\delta >$ so that $|z - \zeta| \leq \delta$ implies $|f(z) - f(\zeta)| < \epsilon$ and hence

$$|F(z) - F(w) - f(z)(z - w)| \leq \epsilon|z - w|.$$

This means $F'(z)$ exists and equals f. Thus F is holomorphic and hence is $F' = f$, as desired. □

COROLLARY 27. *Suppose $\{f_n\}$ is a sequence of holomorphic functions on a domain Ω which converge uniformly on compact sets to a function f. Then f is also holomorphic.*

PROOF. Choose any disk $D \in \Omega$. For any triangle T in D, $\int_T f_n dz = 0$ since f_n is holomorphic. Since T is compact, we deduce $\int_T f dz = 0$ (uniform convergence implies the integrals converge) and hence f is holomorphic on D by Morera's theorem. Thus f is holomorphic on all of Ω. □

5. Equicontinuity and compactness

Suppose (Y, d) is a metric space and let $\mathcal{F}$ be a subset of $C(X, Y)$, the continuous functions from X to Y. If $x_0 \in X$, we say the family is equicontinuous at x_0 if for any $\epsilon > 0$ there is a neighborhood U of x_0 so that $d(f(x), f(y)) < \epsilon$ for every $x, y \in U$. If the family is equicontinuous at every point we simply say it is equicontinuous.

A family of functions $\mathcal{F}$ is called pointwise bounded if for every x $\{f(x) : f \in \mathcal{F}\}$ is a bounded set (different bounds for different x are allowed). See, e.g., Exercise 5, page 279 of Munkres' *Topolygy* [].

THEOREM 28 (Arzela's theorem). *If X is compact and $\{f_n\} \subset C(X, \mathbb{R}^d)$ is pointwise bounded and equicontinuous, then $\{f_n\}$ has a uniformly convergent subsequence.*

COROLLARY 29. *Suppose Ω is a planar domain and $\{f_n\}$ is a sequence of holomorphic functions mapping Ω into $\mathbb{D}$. Then there is subsequence that converges uniformly on compact subsets.*

PROOF. Pointwise boundedness is obvious. To prove equicontinuity, fix a point $z_0 \in \Omega$ and $D = D(z_0, r)$ where $r = \text{dist}(z_0, \partial\Omega)$ and let $g(z) = z_0 + rz$ map the unit disk to D. Then $h_n = f_n \circ g$ is a holomorphic map of the unit disk to itself, hence is a contraction of the hyperbolic metric, hence

$$|z - z_0| \leq r\delta \quad \Rightarrow \quad \psi(,0,z) \leq \delta \quad \Rightarrow \quad |h_n(z) - h_n(z_0)| \leq \psi(h_n(z), h_n(z_0)) \leq \delta,$$

which is equicontinuity at z_0. Now Arzela's theorem applies. □

6. The Poisson integral formula

If f is holomorphic on $\mathbb{D}$ with continuous boundary values then the Cauchy integral formula implies (taking $w = e^{i\theta}, dw = ie^{i\theta}d\theta$)

$$f(0) = \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w}dw = \int_0^{2\pi} f(e^{i\theta})e^{-i\theta}d\theta\frac{ie^{i\theta}d\theta}{2\pi} = \int_0^{2\pi} f(e^{i\theta})\frac{d\theta}{2\pi}.$$

If $\tau : \mathbb{D} \to \mathbb{D}$ is a Möbius transformation sending $0 \to z$ then $f \circ \tau$ is also holomorphic on $\mathbb{D}$ and

$$\begin{aligned} f(z) &= f \circ \tau(0) = \int_0^{2\pi} f \circ \tau(e^{i\theta})\frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} f(e^{i\theta})|\tau'(e^{i\theta})|\frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} f(e^{i\theta})\frac{1-|z|^2}{|z-w|^2}\frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} f(e^{i\theta})P_z(\theta)\frac{d\theta}{2\pi} \end{aligned}$$

where $P_z(\theta)$ is called the Poisson kernel. Since the kernel is real valued we also have

$$u(z) = \int_0^{2\pi} u(e^{i\theta})P_z(\theta)d\theta,$$

if u is the real part of a holomorphic finction (i.e., if u is harmonic).

The reader can check using the (Euclidean) law of cosines that

$$|z - w|^2 = 1 + r^2 - 2r\cos(\theta - \psi),$$

and thus that the Poission kernel is also given by the formula

$$P_z(w) = \frac{1 - r^2}{1 - 2r\cos(\theta - \psi) + r^2},$$

where $z = re^{i\psi} \in \mathbb{D}$ and $w = e^{i\theta} \in \mathbb{T}$.

Clearly minimum and maximum of $|z-w|^2$ are attained when $w = z/|z|$ (w is the radial projection of z) and $w = -z/|z|$ respectively. Since this is the denominator of P_z, the Poision kernel takes its maximum and minimum at these points and hence satisfies estimates

$$\frac{1-r}{1+r} = P_z(-\frac{z}{|z|} \leq P_z(w) \leq P_z(\frac{z}{|z|} = \frac{1+r}{1-r}.$$

If u is positive then replacing P_z by one of these bounds gives a bound for $u(z)$ in terms of its mean value, i.e.,

$$u(z) = \int P_z(e^{i\theta})u(e^{i\theta})\frac{d\theta}{2\pi} \leq \int \frac{1+r}{1-r}u(e^{i\theta})\frac{d\theta}{2\pi} = \frac{1+r}{1-r}u(0).$$

The analogous lower bound then gives

THEOREM 30 (Harnack's inequality). *If u is positive and harmonic on $\mathbb{D}$ then*

$$(\frac{1-|z|}{1+|z|})u(0) \leq u(z) \leq (\frac{1+|z|}{1-|z|})u(0).$$

In particular, if u is a positive harmonic function on $D(z, 2r)$ then

$$\max_{D(z,r)} u(z) \leq 9 \min_{D(x,r)} u(z).$$

If the "2" is replaced by some other $\lambda > 1$, we get a similar estimate with "9" replaced by a constant depending only on λ. Note that this says that if u is a positive harmonic function on $\mathbb{D}$, then $\log u$ is a Lipschtiz function from the disk (with its hyperbolic metric) to the reals (with the Euclidean metric) and the Lipschtiz constant does not depend on u. This is a striking example of equicontinuity.

Harnack's inequality holds on more general domains (with a larger constant). Suppose u is a positive harmonic function on a domain Ω and that K is compact connected set inside Ω. We can cover K by a finite number, N, of disks $\{D_j\}$, so that the double of each disk is in Ω. Thus for any positive harmonic function on Ω, the minimum and maximum values of u on any ball are within a factor 9 of each other. This implies the minimum and maximum values of u over K are within a factor of 9^N of each other. In other words:

LEMMA 31. *Suppose Ω is a domain and $K \subset \Omega$ is a compact set. Then there is a constant $C < \infty$ so that for any two points $z, w \in K$ and any positive harmonic function on Ω $u(z) \leq Cu(w)$. In particular, any sequence $\{u_n\}$ of positive harmonic*

functions on Ω has a subsequence which either tends to ∞ uniformly on every compact subset of Ω or is uniformly bounded on each compact subset of Ω.

If f is holomorphic on Ω and $0 < |f(z)| < 1$, Then $-\log|f(z)| = \Re(-log(f(z))$ is harmonic and positive, so Harnack's inequality applies. Thus any sequence of such functions the limit is either non-vanishing or identically zero. This is:

THEOREM 32 (Hurwitz's Theorem). *If $\{f_n\}$ is a sequence of non-vanshing holomorphic functions on a domain Ω, that converge uniformly on compact sets to a limiting function f, then f is either identially zero or nowhere zero.*

7. A proof of Riemann's theorem

We can now begin the proof of the Riemann mapping theorem. Fix a point $z_0 \in \Omega$ and let $\mathcal{F}_0$ be the class of 1-1 holomorphic functions from Ω into the unit disk so that $f(z_0) = 0$. This proof breaks up into three stages:

(1) Prove $\mathcal{F}_0 \neq \emptyset$.
(2) Prove there is an element $f \in \mathcal{F}_0$ which maximizes $|f'(z_0)|$.
(3) Prove this element maps Ω onto $\mathbb{D}$.

Proof of (1): If Ω is bounded the first step is trivial; we can take a linear map which shrinks Ω enough and moves it to the origin. Otherwise, apply Lemma 25 first.

Proof of (2): Let $M = \sup_{f\in\mathcal{F}_0} |f'(z_0)|$. Since $|f| \leq 1$, the Cauchy estimate (Corollary 8) implies $|f'(z_0)| \leq \text{dist}(z_0, \partial\Omega)$. Thus M is bounded (depending on z_0, but not on f).

Now choose a sequence so that $f_n'(z_0)$ approaches the supremum M and use Montel's theorem to show there is a holomorphic limit. Clearly this function maps into the unit disk and has the desired derivative at z_0, so we only have to show it is in $\mathcal{F}$, i.e., show it is 1-1. Choose any $w \in \Omega$ and note that $f_n(z) - f_n(w)$ is nowhere vanishing on $\Omega \setminus \{w\}$. By Hurwitz's theorem then $f(z) - f(w)$ is nowhere vanishing or identically 0 and the latter case does not happen since $|f'(z_0)| = M > 0$ ($M > 0$ since $\mathcal{F}$ is non-empty). Hence f never takes the same value twice, as desired.

Proof of (3): If $g \in \mathcal{F}$ omits a point of $\mathbb{D}$ from its image we will show how to construct another function in $\mathcal{F}$ with larger derivative at z_0. Thus the maximizing function must be onto.

Suppose $f \in \mathcal{F}$ omits the value w Let τ and σ be Möbius transformations of the disk to itself so that $\tau(w) = 0$ and $\sigma(\tau(f(z_0))) = 0$. Let $W = \tau(f(\Omega))$. Then W is a simply connected subdomain of the disk and omits 0 so there is a well defined branch of $z^{1/2}$ defined on Ω, call it S. Then $g = \sigma \circ S \circ \tau \circ f$ is holomorphic on Ω, 1-1, and maps z_0 to 0. Moreover, $\sigma \circ S \circ \tau$ fixes the origin and is the composition of two isometries and a strictly expanding map (all with respect to the hyperbolic metric), so its derivative at 0 is strictly greater than 1 in absolute value. Thus $|g'(z_0)| > |f'(z_0)|$, as claimed.

Finally we should observe that the conformal map we have constructed is essentially unique. Suppose f, g are two conformal maps of Ω to $\mathbb{D}$ which both send z_0 to 0. Then $f \circ g^{-1}$ is a conformal self-map of $\mathbb{D}$ fixing the origin and hence is a rotation by Schwarz's Lemma (Lemma 21). Thus f and g differ only by a rotation. If we require our map to satisfy $f'(z_0) > 0$, then it is uniquely determined.

Earlier we noted that Liouville's theorem implies that the plane cannot be conformally mapped to the disk, so the claim in Riemann's theorem is sharp.

8. Koebe's method

The proof of Riemann's theorem in the previous section seems non-constructive at first glance: we use compactness to say a function maximizing a certain derivative exists and argue by contradiction to show this map is 1-1, onto the disk. However, it does describe a simple algorithm for mapping a bounded simply connected domain Ω conformally to the disk with a given point z_0 mapping to the origin:

(1) Find a linear map $f : \Omega \to \Omega_0 \subset \mathbb{D}$ with z_0 mapping to 0.
(2) Assuming Ω_n has been defined, find point w on $\partial\Omega_n$ closest to 0.
(3) Choose Möbius transformations τ, σ of the disk to itself so that $\tau(w) = 0$ and $\sigma(\sqrt{\tau(f(z_0))}) = 0$.
(4) Let $\Omega_{n+1} = \sigma(\sqrt{\tau(\Omega_n)})$.
(5) Repeat steps until point w is within specified distance of unit circle.

In Step 3 of the proof of Riemann's theorem we merely stated that if f omitted a point of the disk then we could increase the derivative by composing with the map $\sigma \circ D \circ \tau$. However, it is easy to see that the multiplicative factor of this increase depends only on $|w|$, where w is the omitted point.

Consider the map $z \to z^2$. By Schwarz's lemma this is a strict contraction of the hyperbolic metric, although the hyperbolic derivative tends to 1 near the boundary as we can see from the explicit formula

$$\nabla_H(z^2) = \frac{2|z|}{1+|z|^2}.$$

We will actually use the fact the the inverse map, $\sqrt{z}$ is an expansion of the hyperbolic metric in the following sense. Suppose W is a simply connected subdomain of $\mathbb{D}$ which does not contain the origin. Then there is a well defined branch of $\sqrt{z}$ on W and since z^2 is a contraction for the hyperbolic metric, we have

$$\psi(\sqrt{z}, \sqrt{w}) > \psi(z, w),$$

for any pair of points $z, w \in W$. We can actually be a little more precise and say

$$|\nabla_H \sqrt{z}| \geq \frac{1+|z|}{2\sqrt{z}}.$$

If $\sqrt{z} = 1 - \epsilon$ this becomes

$$|\nabla_H \sqrt{z}| \geq \frac{1+(1-\epsilon)^2}{2(1-\epsilon)} = 1 + \frac{\epsilon^2}{2(1-\epsilon)} = 1 + O(|1-z|^2).$$

From this we can prove:

LEMMA 33. *Suppose $\Omega \subset \mathbb{D}$ is simply connected and omits the point $w \in \mathbb{D}$. Let τ, σ be Möbius self-maps of the disk $\tau(w) = 0$ and $\sigma(\tau(f(z_0))) = 0$ and let $S(z)$ be a branch of $\sqrt{z}$ root function on $\tau(\Omega)$. Then*

$$|(\sigma \circ S \circ \tau)'(0)| \geq \frac{1+|w|}{2\sqrt{|w|}}.$$

COROLLARY 34. *Suppose $d_n = \operatorname{dist}(\partial\Omega_n, 0)$ and $m = 4/(1-\sqrt{d_n})$. Then $d_{n+k} > \sqrt{d_n}$ for $k \geq m$. In particular, if $d_0 \geq 1/2$ then $1 - d_n = O(1/n)$.*

PROOF. Let $r = \sqrt{d_n}$. As long as $d_{n+k} < \sqrt{d_n} = r$ the derivative at 0 increases by a factor of $(1+r^2)/2r$ at each iteration. This is a contradiction if

$$(\frac{1+r^2}{2r})^k > r/d_n = 1/r,$$

or

$$k \geq \log\frac{1}{r} / \log\frac{1+r^2}{2r}.$$

A few simple estimates show

$$\log \frac{1}{r} / \log \frac{1+r^2}{2r} \leq 2(1-r)/\frac{1}{2}(\frac{1+r^2}{2r} - 1) \leq \frac{4r}{1-r} \leq 4/\sqrt{d_n},$$

so if k is larger than the right hand side, we get a contradiction. Thus $d_{n+k} > \sqrt{d_n}$ for $k \geq m$.

If $d_0 > 1/2$ we repeatedly take square roots of d_0 to get $s_n = d_0^{2^{-n}}$, these numbers approach 1 geometrically fast and the number of iterations where d_n is between s_k and s_{k+1} is at most $O(s_k^{-1})$, which grows exponentially in k. Thus the time to reach s_k is dominated by the time to cross between s_{k-1} and s_k. Thus $d_n > 1 - \epsilon$ after about $O(1/\epsilon)$ iterations. □

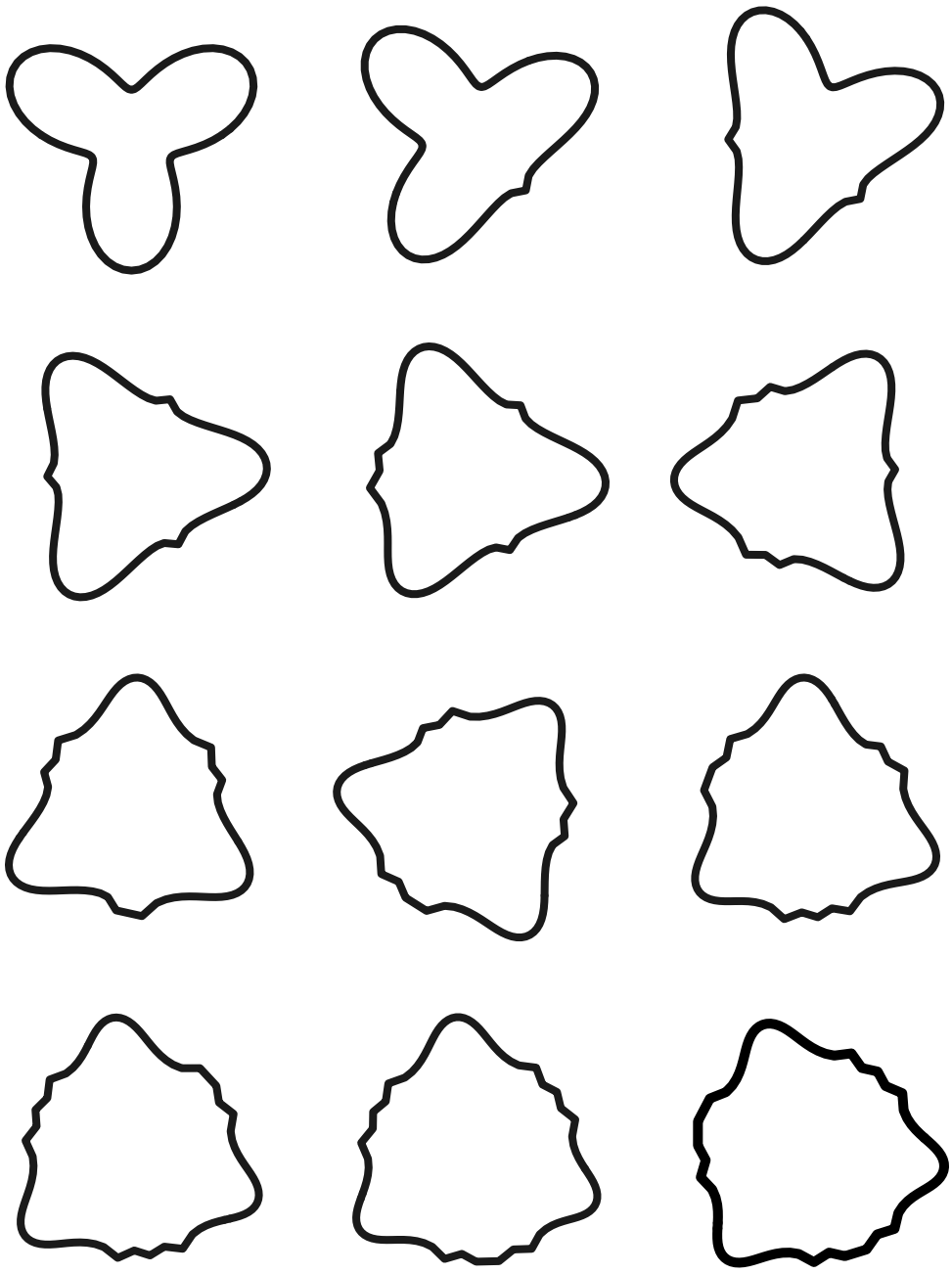


FIGURE 2. On the top left is a sudomain of the disk whose boundary is parameterized by $\gamma(t) = e^{it}\frac{1}{3}(3+\sin(t)))$. This is a polygon with 100 vertices defined by the points $t = k/100$, $k = 1, \ldots, 100$. The next 11 figures show the first 11 iterations of Koebe's method. The next figure show more iterations.

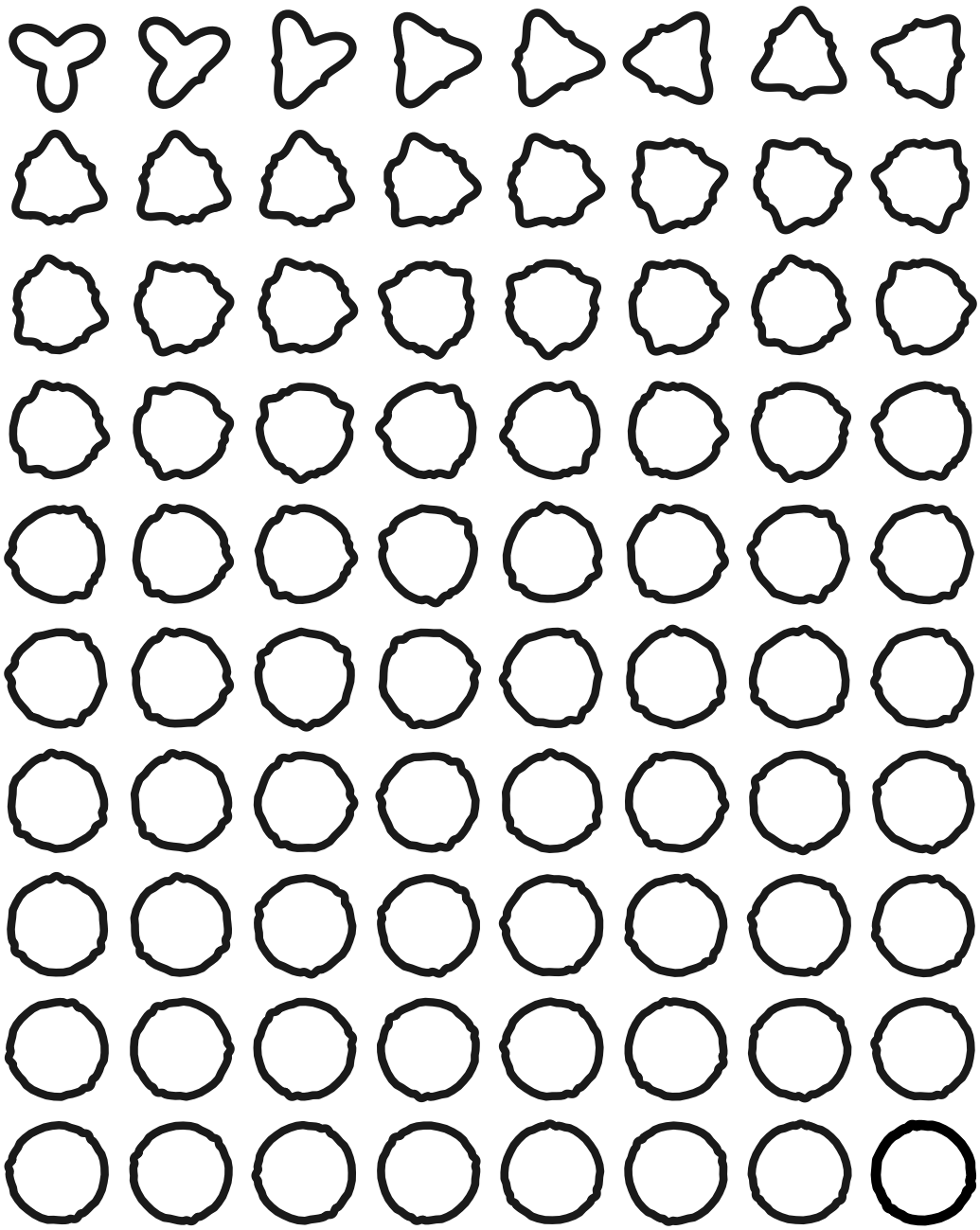


FIGURE 3. This shows the first 80 iterations of Koebe's method for the same domain as in Figure 2

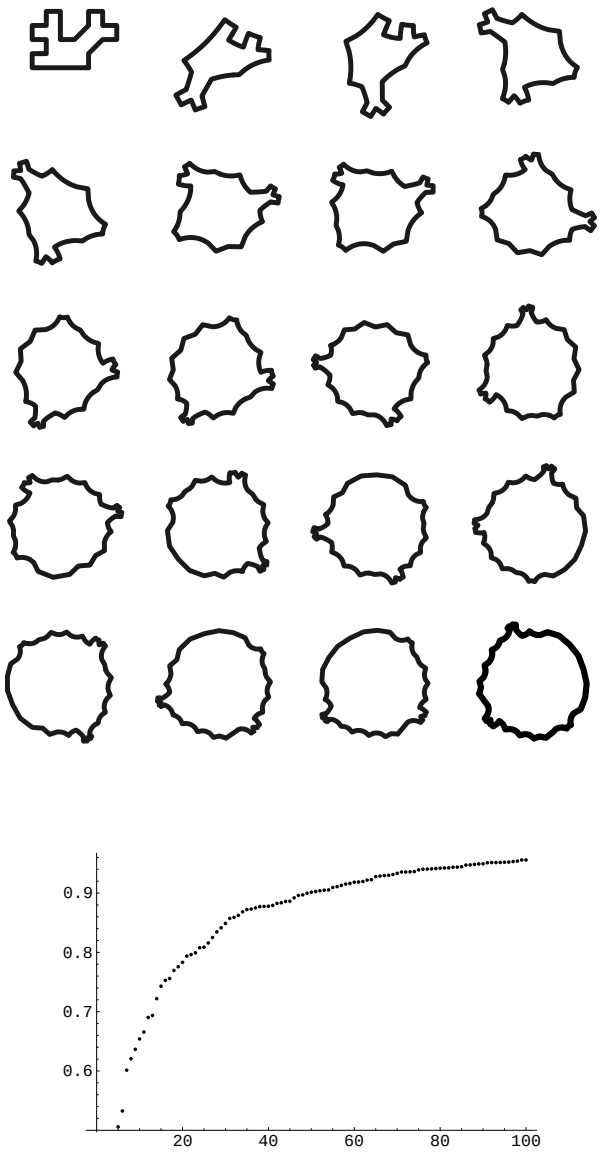


FIGURE 4. Koebe's method applied to a polygon. We have added 19 new, equally spaced vertices to the interior of each edge. On the bottom we have graphed the absolute value of the vertex closest to the origin at each iteration, up to 100 iterations.

9. Caratheodory's Theorem

Next we want to give conditions when the Riemann mapping has a continuous extension to the boundary. This occurs iff the boundary is locally connected and the hardest part of the proof is topological arguments that show the usual definition of local connectedness is equivalent to a more useful version. We will simply accept the following.

THEOREM 35. *Suppose Ω is a bounded simply connected plane domain. The following are equivalent.*

(1) *If U is a relatively open subset of $\partial\Omega$ and $z \in U$, then there is a connected, relatively open subset V so that $z \in V \subset U$.*
(2) *$\partial\Omega$ is a continuous image of a circle.*
(3) *for any $\delta > 0$ there is an $\epsilon > 0$ such that the following holds: if γ is a Jordan arc in Ω with length $\leq \epsilon$, then at least one component of $\Omega \setminus \gamma$ has diameter $\leq \delta$.*

Part (1) is the usual definition of local connectedness and part (3) is the version we shall use (i.e. we will show that (3) implies the Riemann map has a continuous boundary extension). In the course of the proof we will also use a few well known results from real analysis that we explicitly state here:

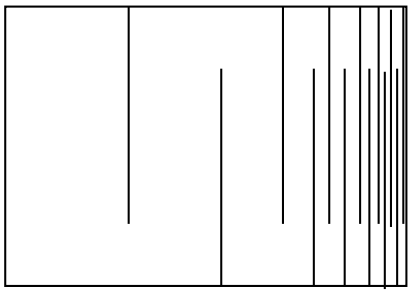


FIGURE 5. An example of a domain with a non-locally connected boundary. The Riemann map onto the interior of this domain fails to have a continuous boundary extension at one point. Examples can be constructed where it fails to have a continuous boundary extension at any point.

Fact 1: The Cauchy-Schwarz inequality

$$\int_I f(x)g(g)dx \leq (\int_I |f(x)|^2 dx)^{1/2} (\int_I |g(x)|^2 dx)^{1/2}.$$

Fact 2: If $\Omega_1 \supset \Omega_2 \supset \dots$ are nested open sets, $\text{area}(\Omega_1) < \infty$ and $\cap_n \Omega_n = \emptyset$, then $\text{area}(\Omega_n) \to 0$.

Fact 3: A continuous function on $\mathbb{D}$ has a continuous extension to the boundary iff it is uniformly continuous.

THEOREM 36 (Caratheodory). *Suppose $\Omega \subset \mathbb{C}$ is simply connected and condition (3) in Lemma 35 holds. Then the Riemann map $f : \mathbb{D} \to \Omega$ has a continuous extension to $f : \overline{\mathbb{D}} \to \overline{\Omega}$.*

PROOF. We assume $\partial\Omega$ satisfies (3) in Lemma **??** and will deduce that the Riemann map $f : \mathbb{D} \to \Omega$ is uniformly continuous, i.e., we have to show that given any ϵ, there is a $\delta > 0$ so that $z, w \in \mathbb{D}$, $|z - w| \leq \delta$ implies $|f(z) - f(w)| < \epsilon$.

Next we would like to assume $\text{area}(f(\{z : \frac{1}{2} < |z| < 1\})) < \infty$. If Ω is bounded, this is obvious since then $f(\mathbb{D}) = \Omega$ has finite area. If Ω is unbounded, consider $\tau(z) = 1/(z - f(0))$. This maps Ω to a domain containing ∞ and $\tau \circ f$ maps $D(0, \frac{1}{2})$ to a neighbohood of ∞. The remainder of the disk is mapped to a bounded region, hence has finite area. Rather than introduce a new symbol, we will let f denote the composed map and Ω the new image.

Fix $\eta > 0$ and choose δ_0 so small that $\text{area}(f(\{z : 1 - \delta_0 < |z| < 1\})) \leq \eta$. Suppose $\delta < \delta_0$ and fix $w \in \mathbb{T}$. Let D be the disk of radius δ around D. Note that $\text{area}(f(D \cap \mathbb{D})) \leq \text{area}(f(\{z : 1 - \delta_0 < |z| < 1\})) \leq \eta$.

Let γ_r be the circular arc in $\mathbb{D}$ centered at w of radius r. Then

$$\ell(f(\gamma_r)) \leq \int_{\gamma_r} |f'(w + re^{i\theta})| r d\theta \leq (r\pi \int_{\gamma_r} |f'(w + re^{i\theta})|^2 r d\theta)^{1/2}.$$

Now square and integrate with respect to $s \in (\delta/2, \delta)$,

$$\begin{aligned} \int_{\delta/2}^{\delta} \ell(f(\gamma_s))^2 ds &\leq \pi \int_{\delta/2}^{\delta} \int_{\gamma_s} |f'(re^{i\theta})|^2 r^2 dr d\theta \\ &\leq \delta \iint_{D \cap \mathbb{D}} |f'(re^{i\theta})|^2 r dr d\theta \\ &\leq \delta \text{area}(f(D \cap \mathbb{D})). \end{aligned}$$

Thus the average value of $\ell(f(\gamma_r))^2$ is at most 2η, so there is at least one value $r \in (\frac{\delta}{2}, \delta)$ with length less than this.

Since D does not contain 0, $W = f(D \cap \mathbb{D})$ does not contain z_0 and hence if $\delta < \delta_1$ is small enough, W must have diameter less than ϵ. Thus if $1 - |z| < \delta_1/2$ or $1 - |w| < \delta_1/2$ and $|z - w| < \delta_1/2$, then both points lie in a δ_1-ball around some boundary point. Thus $|f(w) - f(w)| \leq \epsilon$ by the argument above. If both $1 - |z| \geq \delta_1/2$ and $1 - |w| \geq \delta_1/2$ then both points lie in a compact subset of $\mathbb{D}$ and f is uniformly continuous on $D(0, 1 - \delta_1/2)$ by compactness. Thus there is a δ_2 so that for any two points in this disk $|z - w| \leq \delta_2$ implies $|f(w) - f(z)| \leq \epsilon$. Taking $\delta = \min(\delta_1, \delta_2)$ proves the result. □

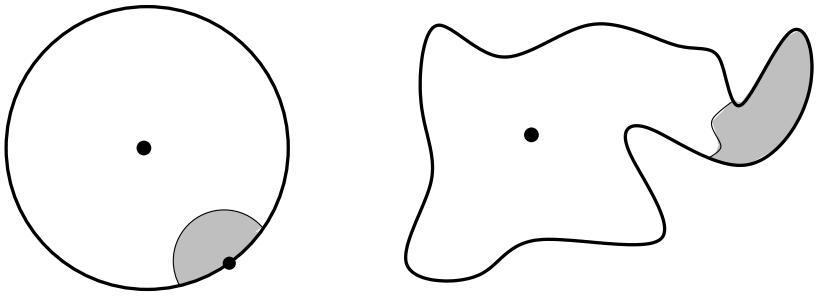


FIGURE 6. By assumption, if a circular crosscut of the disk maps to a curve of short length, it cuts off a subdomain with small diameter.

If the Riemann mapping $f : \mathbb{D} \to \Omega$ has a continuous extension, then $\partial\Omega$ is obviously a continuous image of a circle. Thus using Lemma **??**, we see that the Riemann map has a continuous extension to the whole boundary iff $\partial\Omega$ is locally connected. It is not to hard to deduce (3) of Lemma **??** directly; we will do this later when we have Beurling's lemma at our disposal (this says that the conformal image of a set of small diameter in Ω also has small diameter in $\mathbb{D}$). See Lemma **??**.

The proof also shows the Riemann map extends continuously to a single point $w \in \mathbb{T}$ if there is sequence of circular crosscuts γ_n centered at w so that $f(\gamma_n)$ divides Ω into two subdomains, and the one not containing $f(0)$ has diameter tending to zero. Such a sequence of crosscuts is closely related to the idea of compactifying a domain using prime ends, but we will not discuss thus further here.

10. Schwarz reflection

If $\partial\Omega$ is a polygon then Caratheodory's theorem implies the conformal map $f\mathbb{D} \to \Omega$ has a continuous extension to the boundary, but much more is true: f has a holomorphic extension at every boundary point except the preimages of the vertices. This follows from the well known:

THEOREM 37 (Schwarz Reflection principle). *Suppose $\Omega \subset \mathbb{H}$ and let Ω^* be the reflection of Ω across $\mathbb{R}$ and let Ω_0 be the interior of the closure of $\Omega \cup \Omega^*$. Suppose f is holomorphic on Ω, has a continuous extension to $E = \partial\Omega \cap \mathbb{R}$ and $f(E) \subset \mathbb{R}$. Extend f to Ω^* by setting*

$$f(z) = \overline{f}(\bar{z}).$$

Then the extended function is holomorphic on Ω_0.

PROOF. Supose $f = u + iv$ and $z = x + iy$. Then

$$g(z) = \overline{f(\bar{z})} = u(x, -y) - iv(x, -y) = a(x, y) + ib(x, y).$$

It is easy to check that a, b satisfy the Cauchy-Riemann equations ($a_x = u_x = v_y = b_y$, $a_y = -u_y = v_x = -b_x$) so g is holomorphic on Ω^*. Moreover, since f is real valued on E, we see that f has a continuous extension to Ω_0 which is holomorphic on $\Omega \cup \Omega^*$. We only have to show this extension is homomorphic on E.

Consider a small ball in Ω_0 centered at a point of E. Let T be a triangle in the ball. If such T does not intersect E, the integral is clearly 0. Similarly if T hits E at one vertex of along one edge, for then we can write the integral as a limit of integrals over curves that do not hit I and hence are zero.

Otherwise I divides T into two polygons whose interiors are contained in $\mathbb{H}$ and $\mathbb{L}$ respectively. The integrals over the boundaries of these two pieces add up to the integrals over T (since the sum differs from the integral over T by two integrals over I in opposite directions). Moreover each piece has integral zero, for the same reasons as above. □

The Schwarz reflection principle readily extends to reflections across boundary arcs which are circular arcs (simply map then to real axis by a Möbius transformation).

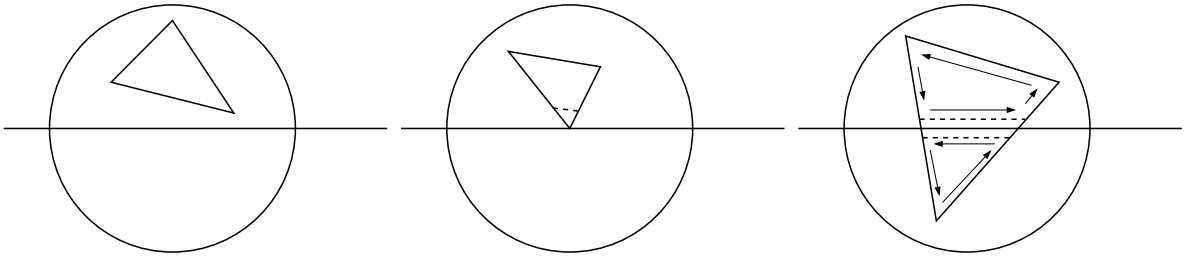


FIGURE 7. In the first case the integral is clearly zero and in the second case it is clearly a limit of zero integrals. In the third case, we can write the integral as the limit of the sum of two integrals inside the half-planes.

11. The maximum principle

The Poisson integral formula says that if u is harmonic on a neighborhood of a disk $D = D(x, r)$, then

$$u(z) = \int_0^{2\pi} u(x + re^{i\theta}) P_z(\theta) d\theta,$$

and hence

$$|u(z)| \leq \max_{\partial D} |u|.$$

Moreover, equality can occur only if u is constant on the boundary of the disk (and hence constant on the whole disk). This implies that a harmonic function on a domain which attains a maximum value must be constant. This is the maximum principle for harmonic functions.

If u is harmonic on a bounded domain Ω and has a continuous extension to the boundary (which we also call u), then it must attain a maximum on the closure of Ω (which is compact) and this value must be taken at a boundary point. If u does not have a continuous extension to the boundary, the following is still true.

LEMMA 38. *Suppose u is a harmonic function on a bounded domain Ω and that $\limsup_{z \to \partial\Omega} u(z) \leq M$. Then $u \leq M$ on Ω.*

PROOF. This is trivial if u is constant so suppose it is not and that $u - M$ takes a positive value somewhere. Then there is an $\epsilon > 0$ so that $\{z \in \Omega : u(z) > M + \epsilon\}$ is non-empty. It does not contain any sequence tending towards $\partial\Omega$, for this would contradict our limsup assumption. Thus this set has a compact closure in Ω

and u must attain a maximum value on this compact set. Thus u is constant, a contradiction. Thus $u - M \leq 0$ everywhere, as desired. □

The maximum principle implies that if two harmonic functions u, v both have continuous extensions to the boundary which agree everywhere, the $u - v \leq 0$ and $v - u \leq 0$, i.e., $u = v$. We will need the following generalization of this, which allows for finitely many discontinuities on the boundary.

LEMMA 39. *[Lindelöf's Maximum Principle] Suppose u, v are bounded harmonic functions on a bounded domain Ω which each have continuous boundary values everywhere on $\partial\Omega$, except for a finite set E. If the boundary values of u, v agree except on E, then $u = v$ on Ω.*

PROOF. Suppose $E = \{z_1, \ldots, z_n\}$ and let

$$g(z) = \log(\text{diam}(\Omega)) - \frac{1}{n}\sum_{k=1}^{n} \log|z - z_k|.$$

Then g is positive and harmonic on Ω and tends to $+\infty$ on the set E. Thus for any $\epsilon > 0$,

$$\limsup_{z \to \partial\Omega} \epsilon g(z) + v(z) - u(z) \geq 0,$$

so by the maximum principle for harmonic functions,

$$v(z) \geq u(z) - \epsilon g(z),$$

on Ω. Taking $\epsilon \to 0$ gives $v \geq u$. Reversing the roles of v and u gives $v = u$, as desired. □

This result can be extended to allow certian infinite exceptional sets (sets of zero logarithmic capacity).

12. Existence of Schwarz-Christoffel parameters

In the first chapter we introduced the Schwarz-Christoffel formula and showed that it always defines a holomorphic map of the disk whose boundary is a polygon with the given angles (but it may be self-intersecting). We left open the question of whether the conformal map onto every polygon can be represented in this way. We can now prove this.

We want to apply Lindelof's maximuum principle to the arguments of the Riemann map and the Schwarz-Christoffel map onto a polygonal region in order to show the two maps are the same. First we have to verify that these arguement functions are both bounded.

LEMMA 40. *Suppose G is a Schwarz-Christoffel map of the disk onto a polygonal domain. Then* $\arg(G')$ *is bounded.*

PROOF. By the Schwarz-Christoffel formula, $\arg(G'(w)) = \sum_n (\alpha_k - 1)[\arg(z_k - w) - \arg(z_k)]$, which is clearly a finite sum of bounded functions. □

LEMMA 41. *Suppose F is a Riemann map of the disk onto a polygonal domain. Then* $\arg(F')$ *is bounded.*

PROOF. By the Schwarz reflection principle, F has a holomorphic extension across each arc of $\partial\mathbb{D}$ between preimages of the vertices. Thus F' is bounded outside any neighborhood of the prevertices. Suppose w is the preimage of a vertex v at which the interior angle of the polygon is θ. Then $H = (F(z) - v)^{1/\theta}$ maps a neighborhood of w to a half-disk with an arc of $\partial\mathbb{D}$ mapping to the straight edge. Thus by Schwarz reflection, H has a holomorphic extension to a neighborhood of w and hence H' is bounded in a neighborhood of w. Thus $F' = \theta H^{\theta-1} H'$ and $\arg(F') = (\theta - 1)\arg(H) + \arg(H')$. We deduce that $\arg(F')$ is also bounded in a neighborhood of w (but note that $|F'|$ need not be bounded). Since $\arg(F')$ is bounded in sonme neighborhood of every point of $\partial\mathbb{D}$, compactness implies it is bounded on the whole circle. □

THEOREM 42. *The conformal map from the half-plane to a polygon satisfies the Schwarz-Christoffel formula for some choice of A and C. Similarly for the map from the unit disk to a polygon.*

PROOF. Let $F : \mathbb{D} \to \Omega$ be the conformal map of the disk to the interior of the polygon, let $\{z_k\}$ be the preimages of the vertices and let G be the locally 1-1 holomorphic map given by the Schwarz-Christoffel formula with these parameters (and the correct α_k's). We can choose C with $|C| = 1$ so that $\arg(CG') = \arg(F')$ on the interior of every parameter interval (but these arguments are not defined at the prevertices). By Lemmas **??** and 41 both these functions are bounded harmonic functions, so Lemma 39 implies $\arg(CG'(z)) = \arg(F'(z))$ on the interior of the disk.

Since these are the imaginary parts of $\log(CG')$ and $\log F'$ we deduce the real parts of $\log(CG')$ and $\log F'$ differ by a real additive constant and hence CG' an F differ by a real multiplicative constant. By putting this factor into C we may assume $CG' = F'$ thus CG and F differ by an additive constant, which we call A and we are done. □

13. Maps to a rectangle

Now that we have the Schwarz reflection theorem at our disposal, we can go back and verify formula (14) in Chapter 5.

LEMMA 43. *Suppose Ω is a $1 \times R$ rectangle, let $f_R : \Omega \to \mathbb{D}$ be conformal and let P be the cross ratio of the four images of the vertices of Ω. Then*

$$P = \exp(-\pi/R)\frac{1}{16}\prod_{n=1}^{\infty}\left(\frac{1+\exp(-2n\pi/R)}{1+\exp((-2n-1)\pi/R)}\right)^8.$$

PROOF. Since cross ratio is invariant under Möbius transformations we may assume f_R maps Ω to the upper half-plane, $\mathbb{H}$. To be even more specific, assume the vertices of Ω are $\{0, R, R+i, i\}$ and that these are mapped to $\{1, \infty, 0, P\}$ respectively. See Figure 8.

Applying Schwarz reflection to each of the sides, we can extend f_R to be conformal on each of the adjacent, similar rectangles, and mapping these rectangles to the lower half-plane. In fact, we can continue reflecting until the map is defined on the whole plane, as illustrated in the bottom of Figure 8. The gray squares are mapped to the upper half-plane and the white squares are mapped to the lower half-plane. When there is more than one way to reach a rectangle in the grid by reflections, it is easy to check that the alternate definitions of the extension agree. Moreover, since angles are doubled at the corners of the rectangles we see that the values $\{0, \infty, 1, P\}$ are taken on with multiplicity 2.

In other words, f_R is holomorphic function on the plane, except for poles of order 2 at the points $\mathcal{L}_1 = (2\mathbb{Z}+1)R + 2i\mathbb{Z}$ and it has zeros of order two at the points $\mathcal{L}_2 = 2R\mathbb{Z} + i(2\mathbb{Z}+1)$. Another function with this same property is

$$F_R(z) = \left(\prod_{n=-\infty}^{\infty}\left(\frac{1-\exp(\frac{\pi i}{R}(2ni + (R+i) - z))}{1-\exp(\frac{\pi i}{R}(2ni + R - z)}\right)^2.$$

It is easy to check that the infinite product converges in both directions and defines a function with periods $2R$ and $2i$. The numerators vanish exactly iff $z \in \mathcal{L}_1$ and

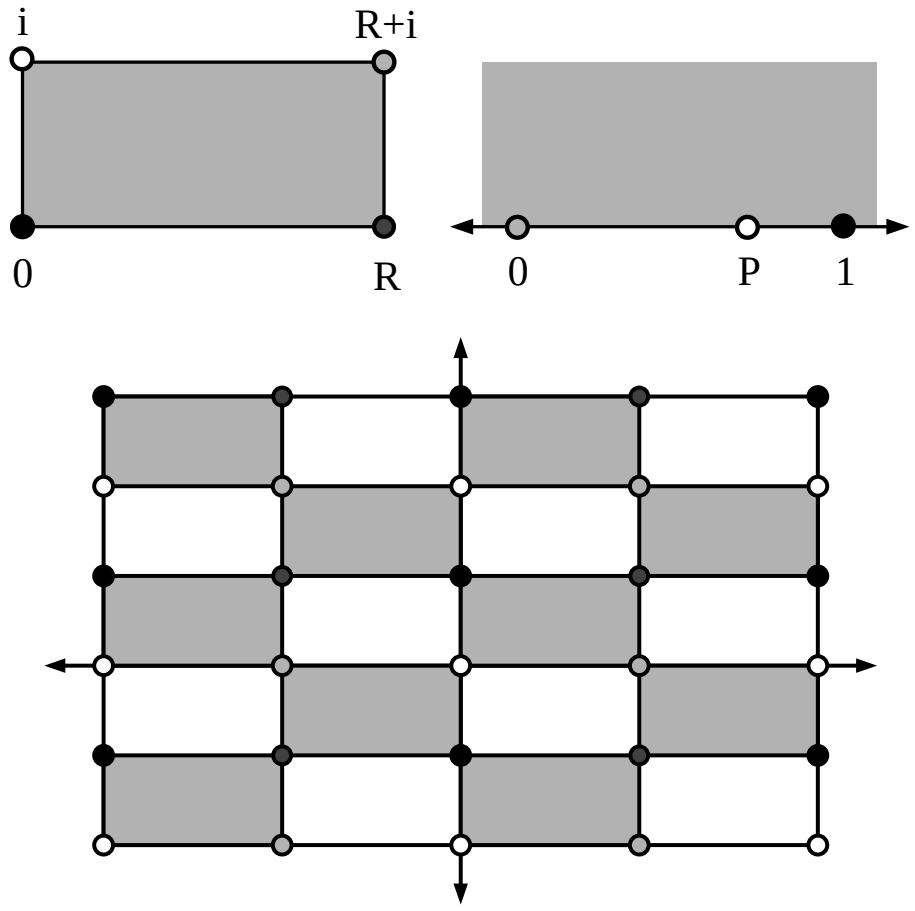


FIGURE 8. We assume f_R maps a rectangle to the upper half-plane with the vertics mapped as shown. After repeated reflection, this extends to a map of the plane to the Riemann sphere, with poles at the dark gray points, zeros at the light gray, value 1 at the black points and value $P \in (0,1)$ at the white points.

the denominators vanish exactly on $\mathcal{L}_2$. Thus f_R/F_R is holomorphic off $\mathcal{L}_2$. In fact, on $\mathcal{L}_2$ the poles cancel, so that f_R/F_R is actually holomorphic and periodic on the whole plane and hence bounded. By Liouville's theorem, f_R/F_R is constant. Since $f_R(0) = 1$, this means $f_R(z) = F_R(z)/F_R(0)$. In particular, $f_R(i) = F_R(i)/F_R(0)$, and both terms on the right hand side are explicitly computable.

First rewrite the product by reindexing some of the terms

$$F_R(z) = \prod_{n=-\infty}^{\infty} \frac{(1-\exp(\frac{\pi i}{R}(2ni+(R+i)-z)))(1-\exp(\frac{\pi i}{R}(2ni-(R+i)-z)))}{(1-\exp(\frac{\pi i}{R}(2ni+R-z)))(1-\exp(\frac{\pi i}{R}(2ni-R-z)))}.$$

Now separate the $n = 0$ terms and combine the n and $-n$ terms, and let $q = e^{-\pi R}$, $p = e^{-i\pi z/R}$. This gives

$$1 - \exp(\frac{\pi i}{R}((R+i) - z) = 1 - \exp(\pi i)\exp(\frac{-\pi}{R})\exp(\frac{-\pi i z}{R}) = 1 + qp,$$

and

$$1-\exp(\frac{\pi i}{R}(2ni+(R+i)-z) = 1-\exp(\frac{-2\pi n}{R})\exp(\pi i)\exp(\frac{-\pi}{R})\exp(\frac{-\pi i z}{R}) = 1+q^{2n+1}p,$$

Using these and similar formulas we get

$$\begin{aligned}
F_R(z) &= \frac{(1-\exp(\frac{\pi i}{R}((R+i)-z)))(1-\exp(\frac{\pi i}{R}(-(R+i)-z)))}{(1-\exp(\frac{\pi i}{R}(R-z)))(1-\exp(\frac{\pi i}{R}(-R-z)))} \\
&\quad \times \prod_{n=1}^{\infty}[\frac{(1-\exp(\frac{\pi i}{R}(2ni+(R+i)-z)))(1-\exp(\frac{\pi i}{R}(-2ni+(R+i)-z)))}{(1-\exp(\frac{\pi i}{R}(2ni+R-z)))(1-\exp(\frac{\pi i}{R}(-2ni+R-z)))} \\
&\quad \cdot \frac{(1-\exp(\frac{\pi i}{R}(2ni-(R+i)-z)))(1-\exp(\frac{\pi i}{R}(-2ni-(R+i)-z)))}{(1-\exp(\frac{\pi i}{R}(2ni-R-z)))(1-\exp(\frac{\pi i}{R}(-2ni-R-z)))}] \\
&= \frac{(1+qp)(1+q^{-1}p)}{(1+p)(1+p)} \prod_{n=1}^{\infty} \frac{(1+q^{2n+1}p)(1+q^{-2n+1}p)(1+q^{2n-1}p)(1+q^{-2n-1}p)}{(1+q^{2n}p)^2(1+q^{-2n}p)^2} \\
&= \frac{(1+qp)(1+q^{-1}p)}{(1+p)(1+p)} \prod_{n=1}^{\infty} \frac{(1+q^{2n+1}p)(p+q^{-2n+1})(1+q^{2n-1}p)(p+q^{-2n-1})q^{-4n}}{(1+q^{2n}p)^2(p+q^{-2n})^2q^{-4n}} \\
&= \frac{(1+qp)(1+q^{-1}p)}{(1+p)(1+p)} \prod_{n=1}^{\infty} \frac{(1+q^{2n+1}p)(p+q^{2n-1})(1+q^{2n-1}p)(p+q^{2n+1})}{(1+q^{2n}p)^2(p+q^{-2n})^2}
\end{aligned}$$

If we set $z = 0$ (i.e., $p = 1$), this becomes

$$\begin{aligned}
F_R(0) &= \frac{(1+q)(1+q^{-1})}{4} \prod_{n=1}^{\infty} \frac{(1+q^{2n+1})^2(1+q^{2n-1})^2}{(1+q^{2n})^4} \\
&= \frac{(1+q)(1+q^{-1})}{4(1+q)^2} \frac{(1+q^{2n-1})^4}{(1+q^{2n})^4} \\
&= \frac{1}{4q} \frac{(1+q^{2n-1})^4}{(1+q^{2n})^4}
\end{aligned}$$

If we set $z = i$, then $p = \exp(-\pi i z/R) = \exp(\pi/R) = 1/q$, and we get

$$\begin{aligned}
F_R(i) &= \frac{(1+1)(1+q^{-2})}{(1+q^{-1})^2} \prod_{n=1}^{\infty} \frac{(1+q^{2n})(q^{-1}+q^{2n-1})(1+q^{2n-2}p)(q^{-1}+q^{2n+1})}{(1+q^{2n-1}p)^2(q^{-1}+q^{2n})^2} \\
&= \frac{2(1+q^{-2})}{(1+q^{-1})^2} \prod_{n=1}^{\infty} \frac{(1+q^{2n})(1+q^{2n})(1+q^{2n-2})(1+q^{2n+2})q^{-2}}{(1+q^{2n-1})^2(1+q^{2n+1})^2 q^{-2}} \\
&= \frac{2(1+q^{-2})}{(1+q^{-1})^2} \prod_{n=1}^{\infty} \frac{(1+q^{2n})^2(1+q^{2n-2})(1+q^{2n+2})}{(1+q^{2n-1})^2(1+q^{2n+1})^2} \\
&= \frac{4(1+q)^2(1+q^{-2})}{(1+q^2)(1+q^{-1})^2} \prod_{n=1}^{\infty} \frac{(1+q^{2n})^4}{(1+q^{2n-1})^4} \\
&= 4 \prod_{n=1}^{\infty} \frac{(1+q^{2n})^4}{(1+q^{2n-1})^4}
\end{aligned}$$

Thus

$$P = f_R(i) = F_R(i)/F_R(0) = 16q \prod_{n=1}^{\infty} \left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^8 = 16e^{-\pi/R} \prod_{n=1}^{\infty} \left(\frac{1+e^{-2n\pi/R}}{1+e^{-(2n-1)\pi/R}}\right)^8$$

□

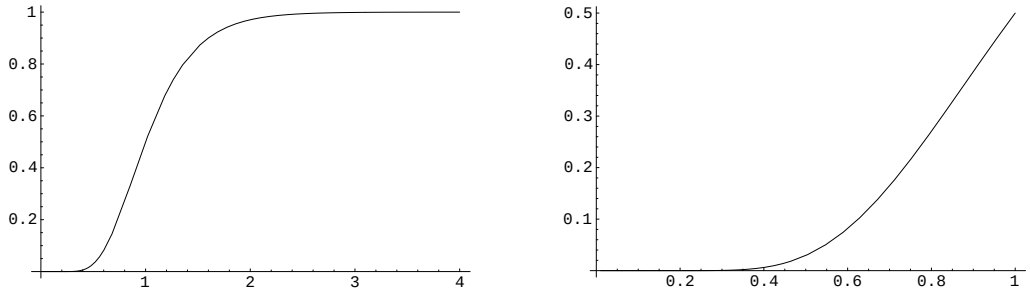


FIGURE 9. A plot of the function $R \to P$. As expected it tends exponentially fast to 0 and 1 as $R \to 0, R \to \infty$ respectively. Note $1\frac{1}{2}$ as required by symmetry.

Thus we have explicit (if somewhat involved) parameters for the Schwarz-Christoffel map onto any rectangle. A similar proof shows

$$Q(R) = \prod_{n=1}^{\infty} \left(\frac{1-q^{2n-1}}{1+q^{2n-1}}\right)^8.$$

CHAPTER 3

Representing conformal maps

We have seen that a conformal map from the disk onto a polygon is determined by the numbers $A, C, \{z_k\}, \{\alpha_k\}$, i.e., if we record these numbers then we "know" what the map is via the Schwarz-Christoffel formula. But what is the most convenient way to plug the parameters into the formula and compute an image point? We saw in the first Chapter than we can compute the power series expansion of the map at the origin, but that this converges slowly and that even for some simple polygons, very many terms are needed to give a good appoximation. In this chapter we will dicuss two ways to get around this problem: using multiple power series in disks that cover the unit disk or using numerical integration to directly evaluate the Schwarz-Christoffel formula.

1. The Carleson decomposition

Figures **??** and **??** indicate that using power series to evaluate conformal maps will be slow and inaccurate in general. This is because the accuracy of the power series decreases as we approach the radius of convergence (which is the just the distance from the center of the series to the closest singularity). However, if a power series has radius of convergence r and we only evaluate it within distance λr of the center (with $\lambda < 1$, then the accuracy is $O(\lambda^n)$. So we want to compute several power series expansions for $\int \prod (1 - \frac{w}{z_k})^{\alpha_k - 1}$ with different centers, so that for any $z \in \mathbb{D}$ we can find one of these expansions whose center is close enough to z to give a good approximation. The tool we use for this is the Carleson decomposition of the disk associated to the set of singularities $S = \{z_k\} \subset \mathbb{T}$.

Given an interval $I \subset \mathbb{T}$, the corresponding Carleson box Q is the region in the disk of the form $\{z = x + iy : z/|z| \in I, 0 < 1 - |z| < |I|\}$. The "top-half" of Q is $T(Q) = \{z \in Q : 1 - |z| > |I|/2\}$. This will be called a Whitney box, and its Euclidean diameter is comparable to its Euclidean distance from $\mathbb{T}$ (abusing notation we may

also call them Whitney "squares"). When I ranges over all dyadic intervals (i.e., all intervals of the form $[j2^{-n}, (j+1)2^{-n}]$), the corresponding Whitney boxes partition the disk into pieces with approximately unit hyperbolic size. Carleson squares are named after Lennart Carleson who used them in his solution of the corona problem and they are now ubiquitous in function theory [**?**], [**?**].

Dyadic Carleson squares form a tree under intersection of the interiors. Each square has a unique parent and two children. The parent of a dyadic Carleson square Q will be denoted Q^*. This obviously also induces a tree structure on Whitney boxes. We will say two dyadic Whitney boxes are neighbors if they are the same size and adjacent; each box therefore has a "left" and a "right" neighbor. One of these is a "sibling" in the sense that it shares a parent, while the other does not.

The decomposition will always begin with a root disk, which we take to be the disk of radius 1/2 around the origin. We will call this the unique type 1 piece. We break the unit circle into 16 intervals using equally spaced points starting at 1 (we could use other values than 16, but this will give decomposition pieces that are "roundish"). Let this list be denoted $\mathcal{L}$. For each interval, compute three numbers: the length, $|I|$, of the interval, the distance, $d(I, S)$, to the closest point of S (which is zero if the interval contains this point) and the distance $d_f(I, S)$ to the second closest point. The f stands for "feature" since this distance is sometimes called the feature distance in the computer science literature.

(1) If $d_f(I, S) \leq 4|I|$ then we say I is type 2 and we add the Whitney box with base I to the decomposition. The interval I is divided into two disjoint, equal length subintervals and each of these is added to the list $\mathcal{L}$ of intervals for testing. The interval I is removed from $\mathcal{L}$. For these intervals there are at least two points of S that are fairly close to I (when compared to $|I|$).
(2) If $d_f(I, S) > 4|I|$ and $d(I, s) > |I|$ then call I type 3 and let the Carleson square with base I be added to the decomposition. No new intervals are added to $\mathcal{L}$. These intervals are "far" from all points of S in the sense that tripling the interval misses all points of S.
(3) If $d_f(I, S) > 4|I|$ and $d(I, s) < |I|$ then call I type 3 and let the Carleson square with base I be added to the decomposition. No new intervals are

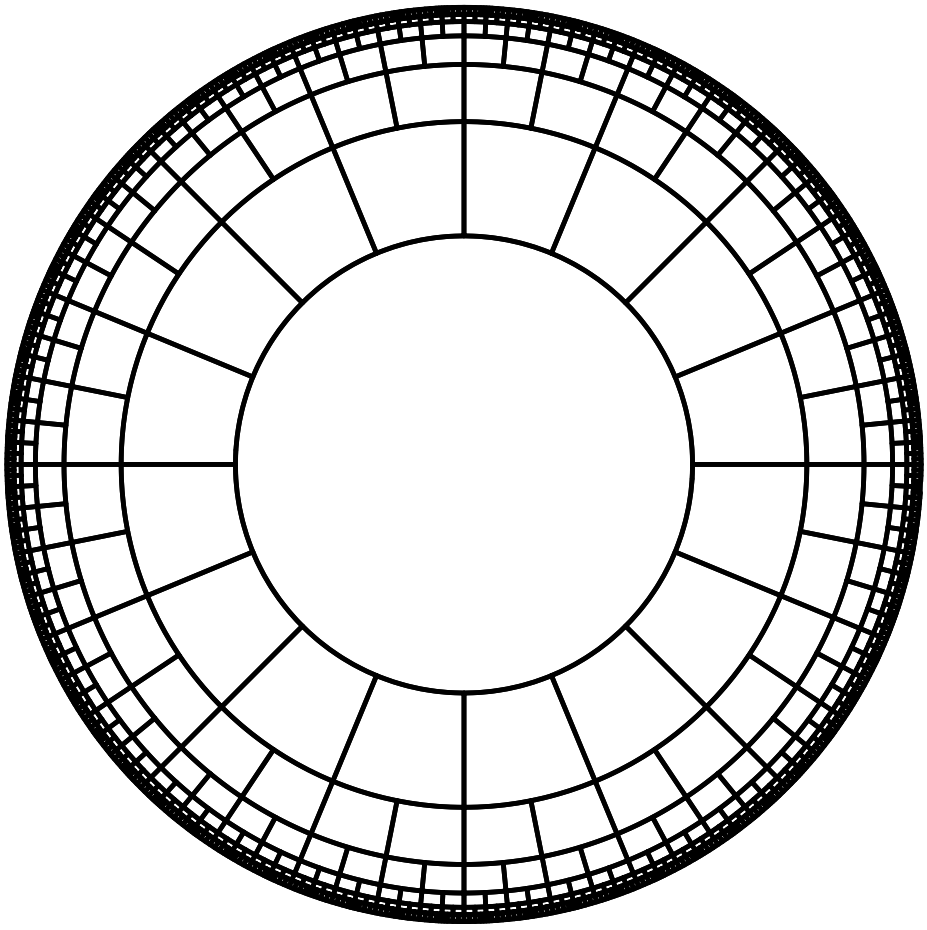


FIGURE 1. A decomposition of the disk into Whitney squares. There is a central disk which has 16 children, each of which has 2 children, and so on towards the boundary. The choice of 16 and 2 is arbitrary, but made here so that the Whitney boxes are "roundish" and so that each box is contained in a disk whose double is still in the unit disk. This means that any holomorphic function on the unit disk has a power series expansion around the center of each Whitney box which converges geometrically fast on the box.

added to $\mathcal{L}$. These intervals are close to some point of S (there is one contained in its triple) but far from all other points of S.

No interval created in this way can be shorter than $\frac{1}{10}$ the distance ϵ between the two closest points of S, and the created intervals are all disjoint, so the number of created intervals is at most $O(n \log \frac{1}{\epsilon})$. This upper bound is attained if we take $2n$

points with two points in each ϵ-neighborhood of a different nth root of unity. If the points of S are more evenly spaced then the number of decomposition pieces is more like $O(n)$.

Note that crowding of the z-parameters leads to more pieces in our decomposition. We shall see later that the extra effort needed to deal with crowding of the prevertices in S is roughly the same as N, the number of distinct pieces in the Carleson decomposition for S.

By adding a fifth type of decomposition piece, called an "arch", it is possible to guarantee that there are at most $O(n)$ pieces with a constant that is independent of n and the geometry of S. However, on the arches, the representation of the function is not with a power series, but with a Laurent series. While the arches provide a fast method if we assume infinite precision computations, if we stick to finite precision calculations then the extra space needed to deal with the Laurent series in the arches is about the same as the number of Whitney boxes that would be needed to fill in the arch and convert "the arched decomposition" into a "regular decomposition". Moreover, constructing the arched decomposition requires some more sophisticated ideas from computational geometry (namely the medial axis of the set S). For all these reasons, will leave the discussion of arches until later.

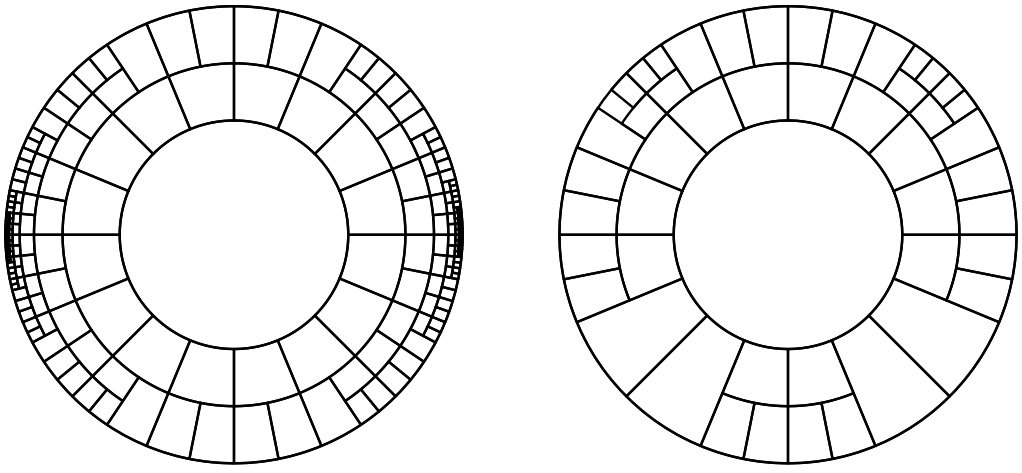


FIGURE 2. Two Carleson decompositions associated to different finite sets on the boundary. On the left is the set $\{0, 0.000628319, 3.14159, 3.14222\}$. On the right is the set $\{0, 1.5708, 3.14159\}$.

On each of the type 1, 2 and 3 pieces, we can compute a power series for $\int \prod_{k=1}^{n}(1-\frac{w}{z_k})^{\alpha_k-1}$. The center of the series is the "center" of the box. The radius of convergence is the distance to the nearest point of S (or more precisely the nearest point where $\alpha_k \neq 1$). By construction, each decomposition piece lies inside a compact subdisk of the disk of convergence and the ratio of these disk is bounded away from 1 uniformly. Thus the power series associated to a piece converges geometrically fast on that piece.

The power series for adjacent pieces may not agree, so we have to make them consistent. We take the power series for the root piece as is. We then compute the images of the 16 points on the boundary of the root piece which are the endpoints of the top edges of its children. Then for each child of the root we choose a, b so that $a + bf$ agrees at these two points with the values compute for the root. In general, if we have a series for a type 2 piece we compute the images of the two endpoints of its bottom edge and for the midpoint of the bottom edge. The maps for each of its children are then normalized by a linear map to agree with the parent map at the two endpoints of its top edge. Continuing is this way we can define a map from the union of type 1, 2 and 3 edges.

The type 3 boxes are clustered into groups, one corresponding to each of the components of $\mathbb{T} \setminus S$. By computing the images of two points on each of these components we can compute the line that the image boundary segment lies on. Doing this for the two components on either side of a point of S and computing their intersecting, we can find that vertex of the image polygon that corresponds to that parameter value. Thus we can compute all the vertices of the image polygon, using only the expansions on type 1, 2 and 3 pieces. For purposes of iteratively finding the parameters, this is all that is needed.

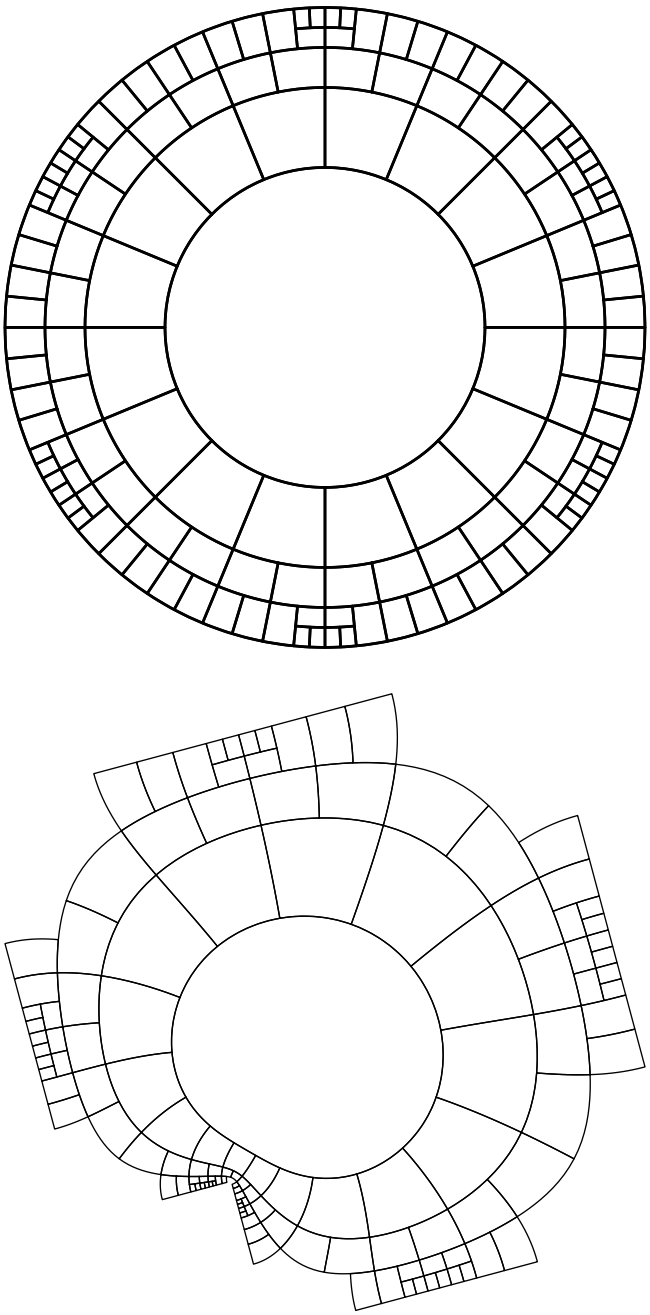


FIGURE 3. The top shows the Whitney-Carleson decomposition of the disk corresponding to six equidistributed points. The bottom shows the image of each box of type 1, 2 and 3 under a degree 10 polynomial, as derived in the text. The figure is rotated compared with the target because we have not yet normalized to find the A and C in the Schwarz-Christoffel formula. However, given the image above we can compute the lines containing the edges of the type 3 boxes and find the intersection points for adjacent edges. Once we know these vertices, we can rotate the figure to match the desired polygon. The result of this normalizing step is shown in Figure 5

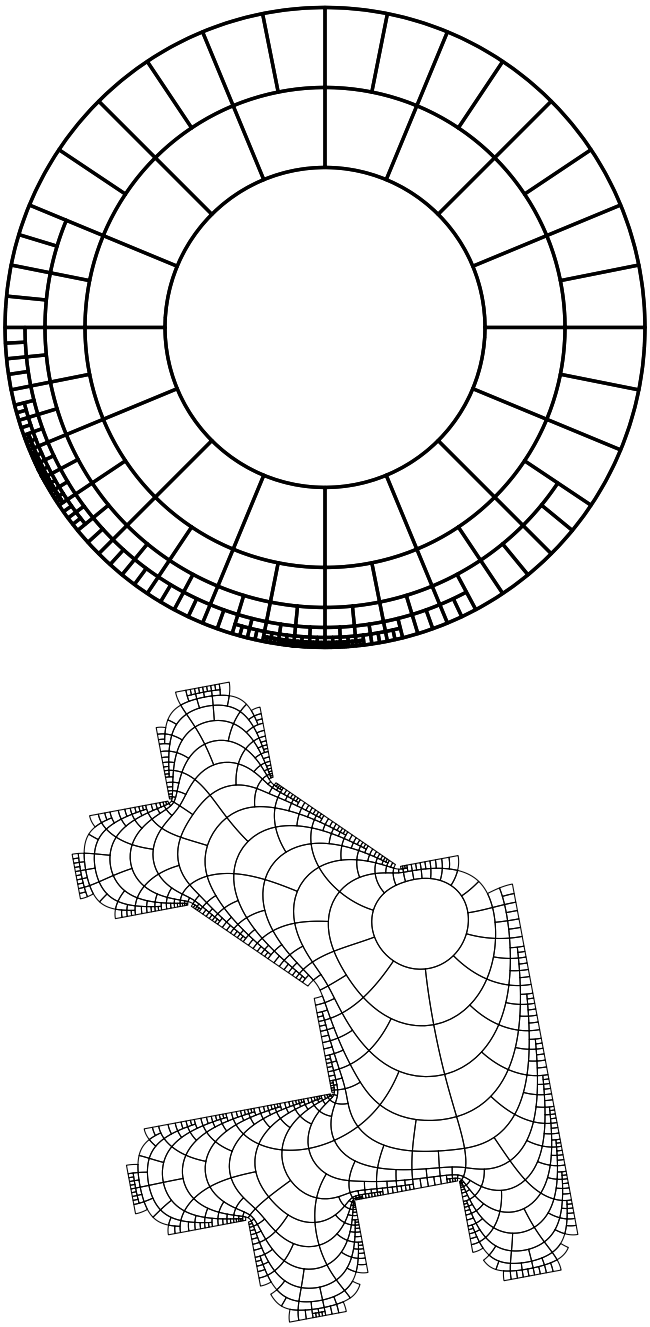


FIGURE 4. Same as for Figure 7 but for a more complicated 20-gon. By counting boxes in the image we can estimate the harmonic measure of any side. For example, the horizontal edge at the top left has measure about 2^{-15}

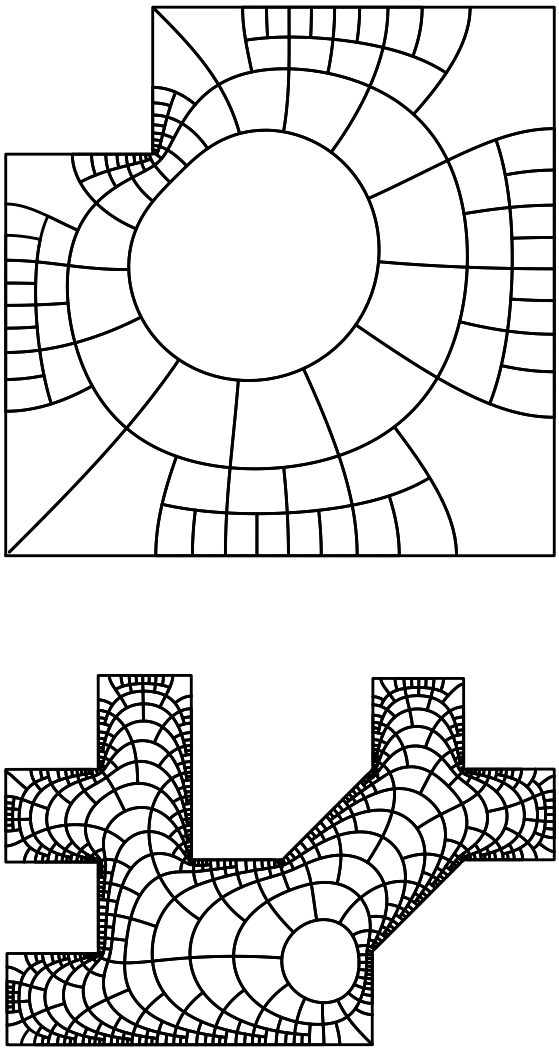


FIGURE 5. These are the normalized versions of Figures 7 and 8. We have computed the lines containing each edge and their intersection points and drawn the corresponding polygon (we have not cheated by using the given vertices except to normalize our figures to have the same first edge, thus setting A and C in the Schwarz-Christoffel formula). Note that the images of the type 4 boxes are not shown. We will see how to draw them in the next section. However, this final step is not needed if all we want is to compute the location of the vertices for use in a iterative solver.

2. An expansion around the singularities

We have already seen how to compute quick convergent power series expansions for the Schwarz-Christoffel map on the boxes of type 1, 2 and 3. This map need not have such an expansion on a type four box if that box contains a singularity, or it map not converge well, if the singularity is outside the box but nearby in an adjacent box. Suppose Q is a type 4 box, $z_k \in \mathbb{T}$ is the nearby parameter and v_k is the vertex of the image polygon corresponding to this parameter (recall that we can compute v_k using our computation on the type 3 boxes). Let θ_k be the interior angle of the polygon at this vertex. Suppose $f : \mathbb{D} \to \Omega$ is our conformal map. Let τ be a Möbius transformation that maps the upper half-plane to the unit disk with $z_k \to 0$, $-z_k \to \infty$ and $0 \to i$, i.e.,

$$\tau(z) = -i\frac{z - z_k}{z + z_k}, \qquad \tau^{-1}(z) = z_k\frac{1 + iw}{1 - iw}.$$

Choose $r_1 > r_2 > 0$ so that the disks $D(0, r_j)$, $j = 1, 2$ are mapped by τ^{-1} to disks $D_2 \subset D_1$, so that D_2 contains all the type 4 boxes associated to z_k and $\partial D_1 \cap \mathbb{D}$ is contained in the union of boxes of type 1, 2 and 3. Then

$$g(z) = \lambda_1(f(\tau^{-1}(r_1 z)) - v_k)^{\pi/\theta_k}, \tag{10}$$

maps $\mathbb{D} \cap \mathbb{H}$ to a neighborhood of 0 in some half-plane whose boundary contains 0 and by choosing a constant λ_1 with $|\lambda_1| = 1$ correctly, we may assume $g(z)$ maps the upper half-plane to itself. Thus by the Schwarz Refection Principle g has an holomorphic extension across the real line (at least in a neighborhood of 0 an so we can write

$$g(z) =,$$

(we can drop the zero term since $g(0) = 0$ by definition). Moreover,

$$f(z) = v_k + [\lambda_2 \sum_{k=1}^{\infty} a_k(\tau(z)/r_1)^k]^{\theta_k/\pi},$$

where λ_2 is chosen so the left hand side map into a cone of angle θ_k with the correct direction.

So this gives us a nice, compact representation for f, if we can compute power series for g. This can be done using a discrete Fourier transforms. Fix a positive integer N and consider the Nth roots of unity $R_N = \{w_k\} = \{e^{i2\pi k/N}\}$. Let V^N be

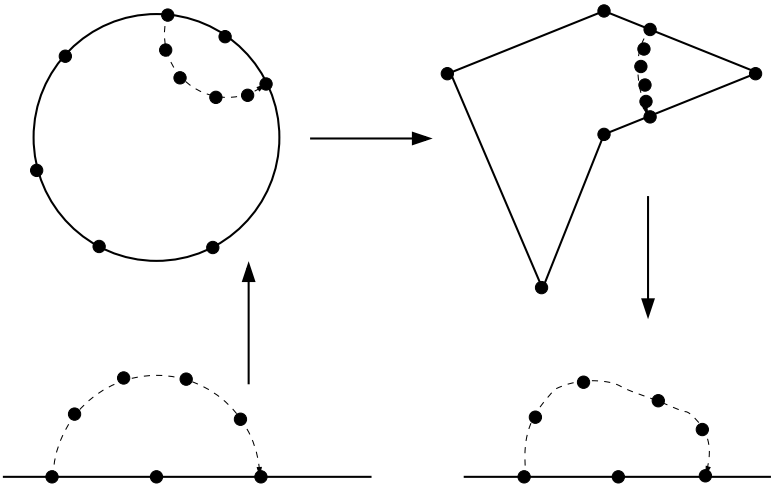


FIGURE 6. This illustrates how (10) is derived. We map the upper half-plane to the disk with 0 going to one of the Schwarz-Christoffel parameters (up arrow). Then we map to the polygon by f (across) and then open the vertex to 180 degrees by a power map (the down arrow). The resulting composition has a holomorphic extension to a neighborhood of 0 by the Schwarz reflection principle. We can compute the power series for this extension by sampling roots-of-unity on the half-circle as shown, mapping them forward as shown and defining the map on the corresponding roots-of-unity in the lower half-plane by reflection. Then apply an FFT to the data to get the coefficients of the power series.

the N-dimensional complex vector space of functions from R_N to $\mathbb{C}$. The functions $e_k(z) = z^k$, $k = 0, \ldots N-1$ form an orthonormal basis with respect to the inner product

$$\langle f, g \rangle = \frac{1}{N} \sum_{k=0}^{N-1} f(w_k)\overline{g(z_k)},$$

as can easily be seen by verifying

$$\|e_n\|^2 = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f(w_k^n)\overline{w_k^n} = \frac{1}{N} \sum_{k=0}^{N-1} 1 = 1,$$

$$\begin{aligned}\langle e_m, e_n\rangle &= \frac{1}{N}\sum_{k=0}^{N-1} w_k^m w_k^{-n} \\ &= \frac{1}{N}\sum_{k=0}^{N-1}(w_1^{m-n})^k \\ &= \frac{1}{N}\frac{1-(w_1^{m-n})^N}{1-w_1^{m-n}} = 0,\end{aligned}$$

where we have used the geometric formula $\sum_{k=0}^{N} z^k = \frac{1-z^N}{1-z}$ and the fact that $w_1^{m-n})^N (w_1^N)^{m-1} = 1$ since w_1 is an Nth root of unity. Therefore any function in V_N can be written

$$f(z) = \sum_{k=0}^{N-1} e_k\langle f, e_k\rangle = \sum_{k=0}^{N-1} a_k z^k,$$

for $z \in R_N$. Using the definition, each coefficient of f can be computed as a sum of N terms, so the whole expansion can be computed in time $O(N^2)$. However, the Fast Fourier Transform gives a $O(n \log n)$ algorithm for computing the same coefficients and will be discussed in Section 5.

The function g defined above maps the upper half-plane to itself and we can define g at the roots of unity that lie in the upper half-plane using (10) because we only need to evaluate f at points that lie in type 1, 2 or 3 boxes. We then define g at the roots of unity in the lower half-plane using $\overline{g(z)} = g(\bar{z})$. Then use the discrete Fourier transform to define a polynomial of degree $N-1$ g_0 which agrees with g at all the Nth roots of unity.

Unfortunately, this does not necessarily mean that g and g_0 are close anywhere else. If we start with $g(z) = z^N$ then the restriction to the Nth roots of unity gives the constant 1 functions and the resulting g_0 is also the constant 1. Fortunately, the maps we want to approximate are conformal and so we can do better than this.

Later on we will prove distortion estimates for conformal maps that say that if g is 1-1 and holomorphic on $D(0,2)$ then $|g'(z)| \le 3|g'(0)|$ on $D(0,1)$. Lets assume for the moment that $g'(0) = 1$. Let

$$g(z) = \sum_{k=0}^{\infty} b_k z_k,$$

be the power series for the function g and let

$$g_0(z) = \sum_{k=0}^{N-1} a_k z_k,$$

be the function derived using the discrete Fourier transform from the restriction of g to the Nth roots of unity. The coefficients $\{b_k\}$ can be obtained from the Cauchy integral formula as

$$b_k = \frac{1}{2\pi i} \int_{\mathbb{T}} \frac{g(z)}{z^{k+1}} dz = \int_0^{2\pi} g(e^{it}) e^{-ik} \frac{dt}{2\pi}.$$

Now break the unit circle into N arcs each of length $2\pi/N$ and centered at the Nth roots of unity. Then on each arc, the product $g(z)z^{-k}$ has second derivative bounded by $O(k^2)$ and therefore differs from a linear function by at most $O(k^2/N)$ on each arc. Therefore

$$|\int_{I_j} g(e^{it}) e^{-ik} \frac{dt}{2\pi} - g(w_j) w_j^k| = O(k^2/N^2).$$

Hence

$$|b_k - a_k| \leq O(k^2/N).$$

The coefficients $\{b_k\}$ satisfy $|b_k| = O(2^{-k})$, and therefore $|a_k| = O(\max(k^2/N, 2^{-k})$. Let $M = \lfloor \log_2 N \rfloor$ and

$$g_1(z) = \sum_{k=0}^{M} a_k z_k.$$

So if $r < 1$ and $|z| \leq r$, then

$$\begin{aligned} |g(z) - g_0(z)| &\leq \sum_{k=0}^{M} |a_k - b_k| r^k \sum_{k=M+1}^{\infty} |b_k| r^k + \sum_{k=M+1}^{N-1} |a_k| r^k \\ &= O(\frac{M^3}{N} + O(2^{-M}) + O(r^{-M}) \\ &= O(\frac{1}{N}(\log N + \frac{1}{1-r})). \end{aligned}$$

This tends to zero for any fixed $r < 1$ and $N \to \infty$. This the series we have constructed will uniformly approximate the conformal map.

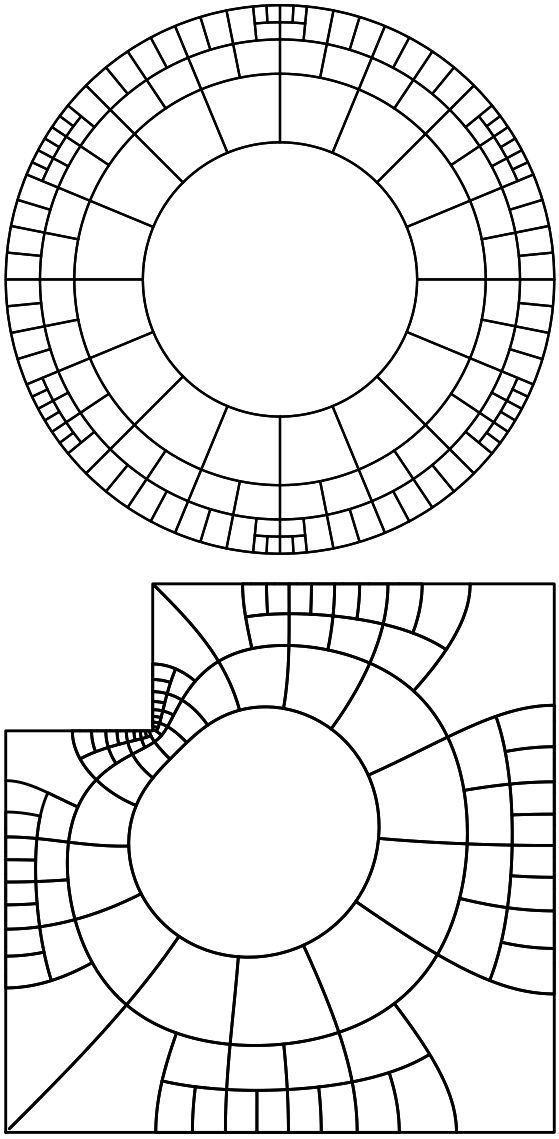


FIGURE 7. The top shows the Whitney-Carleson decomposition of the disk corresponding to six equidistributed points. The bottom shows the image of each box under a degree 10 polynomial, as derived in the text. The vertices of the polygon are found by computing two points on each edge using the type 3 boxes and then finding the intersection points of the corresponding lines.

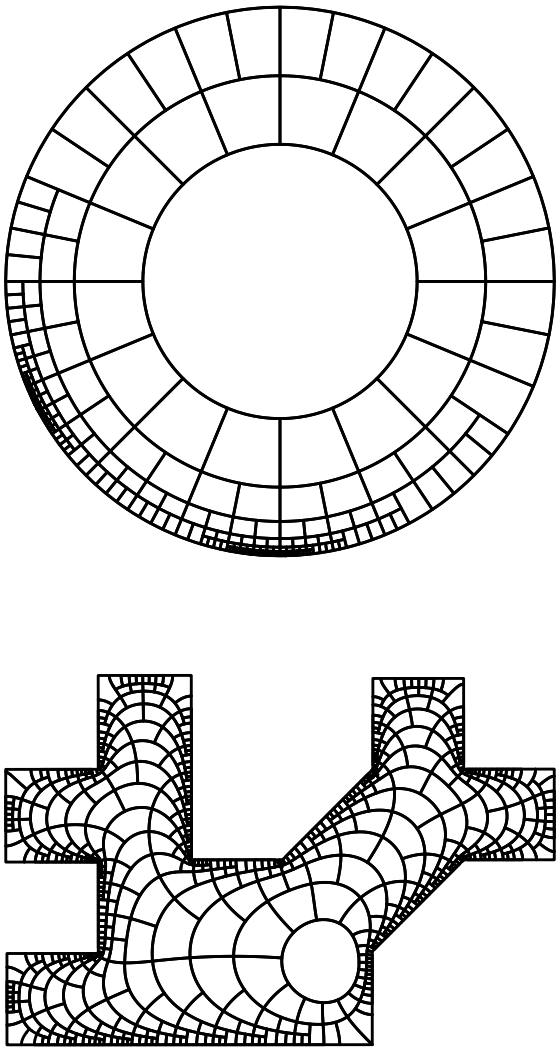


FIGURE 8. Same as for Figure 7 but for a more complicated 20-gon. By counting boxes in the image we can estimate the harmonic measure of any side. For example, the horizontal edge at the top left has measure about 2^{-15}

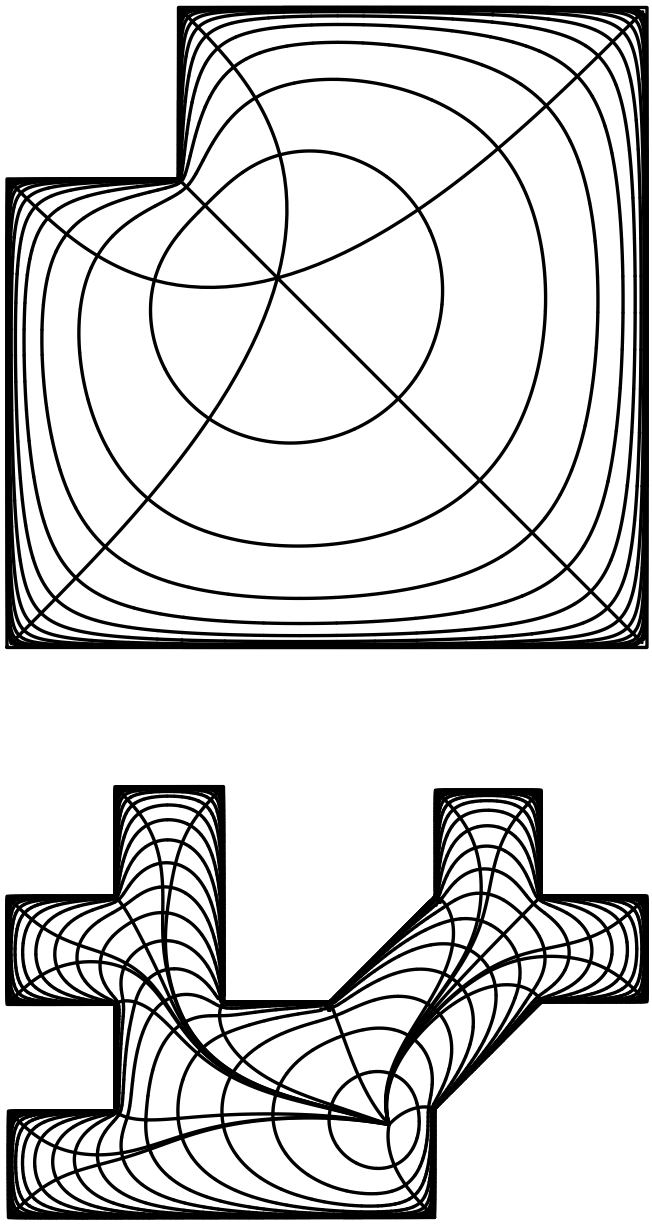


FIGURE 9. For the polygons in Figure 7 and 8 we have used the multiple power series representation to plot the images of the circles of radius $1 - 2^{-n}$ and the radial segments that end at the vertices. As before, polynomials of degree 10 are used for all approximations.

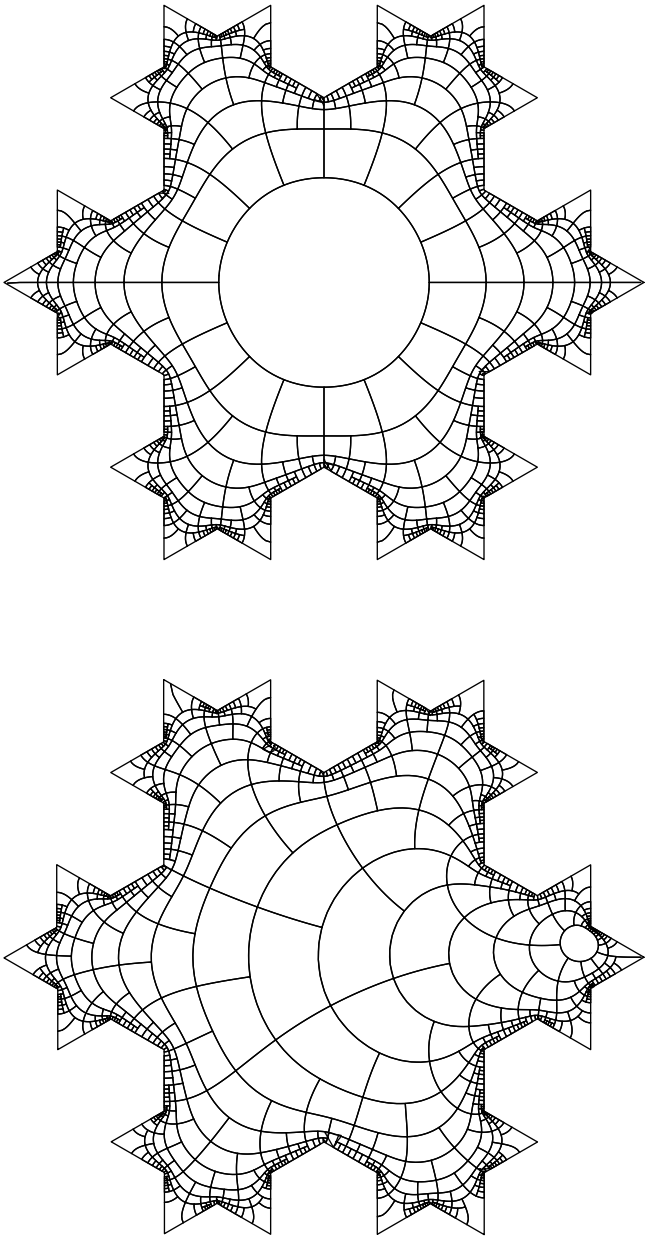


FIGURE 10. Images of the Whitney boxes for the second generation von Koch snowflake.

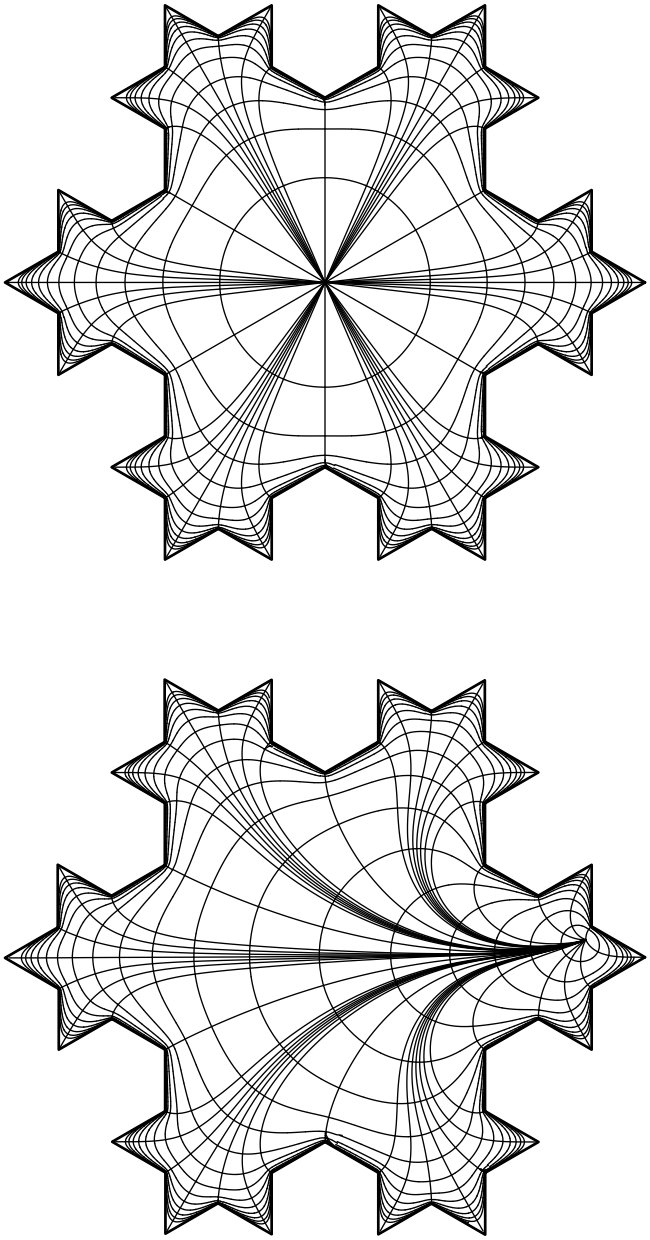


FIGURE 11. Level lines and radial images for the second generation von Koch snowflake. In the top picture the origin is mapped to the center of the snowflake and in the bottom the origin is mapped close to the boundary.

3. Gauss-Jacobi quadrature

The multiple power series described above have the advantage of covering the whole unit disk, so that to evaluate the conformal map at any point we merely have to decide which decomposition box contains the point and sum the corresponding power series. However, if we only want to evaluate the map at a few points, it may not be worth building the map everywhere. For example, in order to use certain iterative methods to estimate the unknown z-parameters, we only need to compute the vertices of the image polygons. In this case, it may be faster simply to numerically evaluate the integral in the Schwarz-Christoffel integral.

There are at least two choices: integrate f' on a ray from the origin to each parameter value (to find the position of the corresponding vertex relative to 0) or integrate $|f'|$ along the boundary arc between two parameters (to get the length of the corresponding polygonal side, which, with the known angles, is enough to determine the polygon). The boundary integral has the advantage of being real valued, whereas the interior integral in complex valued.

Suppose w is defined and integrable on $[a, b]$ and we want to evaluate

$$\int_a^b p(t)w(t)dt,$$

for $p \in \mathcal{P}_n$ (the polynomials of degree n). Think of $w = 1$ or $w(t) = (t-a)^\alpha$ as the main examples. If we are given any $n+1$ distinct points $\{x_k\}_0^n \subset [a, b]$ then p is determined by its values at these points, i.e., the map

$$p \to \{p(x_0), \dots p(x_n)\}$$

is an invertible map $\mathcal{P}_n \to \mathbb{R}^n$. Thus there must be real numbers w_k so that

$$\int_a^b p(t)w(t)dt = \sum_{k=0}^{n} w_k p(x_k), \tag{11}$$

holds for all $p \in \mathcal{P}_n$.

What are these weights more explicitly? Given the point set $\{x_k\}_0^n$ define the Lagrange polynomials

$$L_k(x) = \prod_{0 \le j \le n, j \ne k} \frac{x - x_j}{x_k - x_j}.$$

This is equal to 1 at x_k and equal to 0 at the other x_j's. We must have

$$p(x) = \sum_{k=0}^{n} p(x_k) L_k(x),$$

for any $p \in \mathcal{P}_n$, since both sides are degree n polynomials that agree at $n+1$ points. Thus

$$\int_a^b p(x)w(x)dx = \int_a^b \sum_{k=0}^{n} p(x_k)L_k(x)w(x)dx = \sum_{k=0}^{n} p(x_k)[\int_a^b L_k(x)w(x)dx].$$

Thus (11) holds with $w_k = \int_a^b L_k(x)w(x)dx$.

We can simplify this further by noting that

$$L_k(x) = \prod_{0 \le j \le n, j \ne k} \frac{x - x_j}{x_k - x_j} = \frac{p_n(x)}{(x - x_k)p_n'(x_k)},$$

since both sides are degree n polynomials that are 1 at x_k and 0 at x_j, $j \ne k$. Thus

$$w_k = \int_a^b \frac{p_n(x)}{(x - x_k)p_n'(x_k)} w(x)dx. \tag{12}$$

The discussion so far assumes that we are given the points $\{x_k\}$. If we are allowed to choose these points, then we have $n+1$ additional degrees of freedom, so we might hope to correctly evaluate integrals for even higher degree polynomials. In fact, we can choose $n+1$ points $\{x_k\}_0^n$ so that (11) holds for all polynomials of degree $\le 2n+1$.

The secret is to choose a polynomial p of degree $n+1$ which is orthogonal to every polynomial q of lesser degree, i.e., so that

$$\langle p, q \rangle_w = \int_a^b p(t)q(t)w(t)dt = 0,$$

for all $q \in \mathcal{P}_n$. Now let $\{x_k\}$ be the zeros of p and let $\{w_k\}$ be the weights which make (11) true for polynomials of degree $\le n$. If f is a polynomial of degree $\le 2n+1$, then long division of polynomials shows that we can write $f = a + bp$ where a, b are polynomials of degree $\le n$. Thus

$$\begin{aligned} \int_a^b f(t)w(t)dt &= \int_a^b a(t)w(t)dt + \int_a^b b(t)p(t)w(t)dt \\ &= \sum_k w_k a(x_k) + 0 \\ &= \sum_k w_k f(x_k), \end{aligned}$$

where the last line holds since $f = a$ on the zeros of p.

To see that it is not possible to increase the degree of f to $2n+2$, consider the function $\prod_{k=0}^{n}(t-x_k)^2$. It vanishes at the points $\{x_k\}$ so $\sum_k w_k f(x_k) = 0$, but $\int_a^b f(t)w(t)dt > 0$, at least if $w > 0$ since $f > 0$ except at $n+1$ points.

LEMMA 44. *For orthonormal polynomials we have*

$$w_k = \frac{k_n}{k_{n-1}} \frac{\langle p_{n-1}, p_{n-1}\rangle}{p_{n-1}(x_k)p_n'(x_k)},$$

where k_n is the leading coefficient of p_n (i.e., the coefficient of x^n).

To prove this we need two preliminary results. The first is:

LEMMA 45. *Let*

$$K_n(x,y) = \sum_{k=0}^{n} p_k(x)p_k(y).$$

Suppose $K(x,y)$ is a polynomial of degree n in both x and y. Then

$$\langle p(x), K(x,y)\rangle_{w(x)} = p(y),$$

holds for every polynomial p of degree n iff $K = K_n$.

PROOF. If p is polynomial of degree $\leq n$ then it has a n expansion in terms of the basis $p(x) = \sum a_m p_m(x)$, so

$$\begin{aligned}
\langle p(x), K_n(x)\rangle_w &= \langle \sum a_m p_m(x), \sum p_k(x)p_k(y)\rangle_w \\
&= \sum_{m,k} a_m p_k(y)\langle p_m(x), p_k(x)\rangle_w \\
&= \sum_k a_k p_k(y) \\
&= p(y),
\end{aligned}$$

so the equality holds when $K = K_n$. Conversely, some equality holds for K and all p. Fix w and choose $p(x) = K_n(x,w)$. Then

$$\langle K_n(x,w), K(x,y)\rangle_w = K_n(y,w).$$

But by our earlier calculation k_n has the reproducing property so

$$\langle K(x,w), K_n(x,y)\rangle_w = K(y,w).$$

Since the two left hand sides equal the same integral, we deduce $K(y,w) = K_n(y,w)$ for any y, w, which proves the lemma. □

The second preliminary result we need is:

THEOREM 46 (Christoffel-Darboux). *With notation as above,*

$$K_n(x,y) = \frac{k_n}{k_{n+1}} \frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x-y}$$

PROOF. Let $K(x,y)$ denote the right hand side above. The numerator is a polynomial in x of degree $\leq n+1$ and vanishes when $x=y$, so $K(x,y)$ is actually a polynomial in x of degree $\leq n$. Similary for y. Thus to show $K = K_n$ we only have to show it has the reproducing property of the previous lemma.

A bit of expanding and using $\langle p_n, p_{n+1}\rangle_w = 0$ shows

$$\begin{aligned}\langle p(x)m, K(x,y)\rangle_w &= \frac{k_n}{k_{n+1}}\langle (p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)), \frac{p(x)-p(y)}{x-y}\rangle_w \\ &+ \frac{k_n}{k_{n+1}} p(y)\langle p_{n+1}(x), \frac{p_n(y)-p_n(x)}{x-y}\rangle_w \\ &+ \frac{k_n}{k_{n+1}} p(y)\langle p_n(x), \frac{p_{n+1}(x)-p_{n+1}(y)}{x-y}\rangle_w\end{aligned}$$

Note that $(p(x)-p(y))/(x-y)$ has degree $\leq n-1$ as a polynomial in x and hence is orthogonal to p_n. Thus the first inner product is 0. Similarly for the second inner product. To compute the third inner product, write

$$\begin{aligned}\frac{k_n}{k_{n+1}} p(y) \frac{p_{n+1}(x)-p_{n+1}(y)}{x-y} &= k_n[\frac{y^{n+1}-x^{n+1}}{y-x} + \dots] \\ &= k_n x^n + \dots \\ &= p_n(x) + q(x,y),\end{aligned}$$

where q is a polynomial of degree $\leq n-1$ and hence orthogonal to p_n. Thus the third inner product equals $\langle p_n, p_n\rangle_w = k_{n+1}/k_n$, and hence

$$\langle p(x)m, K(x,y)\rangle_w = p(y).$$

By the previous lemma this implies $K = K_n$, as desired. □

LEMMA 44. We already know from (12) that

$$w_k = \frac{1}{p_n'(x_k)} \int_a^b \frac{p_n(x)}{x-x_k} w(x)dx$$

Since $p_n(x_k) = 0$,

$$\begin{aligned}\int_a^b \frac{p_n(x)}{x-x_k}w(x)dx &= \frac{1}{p_{n-1}(x_k)}\int \frac{p_n(x)p_{n-1}(x_k)}{x-x_k}w(x)dx \\ &= \frac{1}{p_{n-1}(x_k)}\int \frac{p_n(x)p_{n-1}(x_k)-p_{n-1}(x)p_n(x_k)}{x-x_k}w(x)dx \\ &= \frac{1}{p_{n-1}(x_k)}\frac{k_n}{k_{n-1}}\int K_n(x,x_k)w(x)dx \\ &= \frac{k_n}{k_{n-1}p_{n-1}(x_k)}.\end{aligned}$$

□

For more general functions, the difference between out discrete estimate and the actual integral can be bounded as follows:

$$E_n(f) = \int_a^b f(t)w(t)dt - \sum_k w_k f(x_k) = \frac{f^{(2n)}(\zeta)}{(2n)!k_n^2},$$

where ζ is some point in (a,b) and k_n is the coefficient of the power t^n in $p(t)$. For Schwarz-Christoffel integrals, the most relevant case is when w is a Jacobi weight

$$w(t) = (1-x)^\alpha(1+x)^\beta,$$

when this estimate is known to be (see []),

$$E_n(f) = f^{(2n)}(\zeta)\frac{2^{2n+\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)\Gamma(n+\alpha+\beta+1)n!}{\Gamma(2n+\alpha+\beta+1)\Gamma(2n+\alpha+\beta+2)(2n)!}.$$

If $\alpha = \beta = 0$ then $w = 1$ and p is a Legendre polynomial. Then the error bound simplifies to

$$E_n(f) = f^{(2n)}(\zeta)\frac{2^{2n+1}(n!)^4}{(2n+1)((2n)!)^3}.$$

Consider a simple case like $f(t) = e^t$. Then all the derivatives of f are bounded on $[-1,1]$ and using Stirling's formula

$$n! \sim n^n e^{-n}\sqrt{2\pi n},$$

we see that

$$E_n(e^t) = O(n^{-2n}).$$

On the other hand, the nth order Taylor series for e^t only approximates it to within $1/n! \gg n^{-n}$ on $[-1,1]$. Thus the numerical integration using n points should give about twice as many correct digits as term-by-term integration of the nth order power

series. Of course, we can just double the number of terms in the power series to obtain the same accuracy.

Even for fairly small n, both error estimates will be less than machine precision, so a practical comparison would have to estimate the work to produce the approximations. Very roughly, if $f(t) = \prod_{j=1}^{N}(1-\frac{t}{z_j})^{\alpha_k-1}$ is the sort of function that arises in the Schwarz-Christoffel formula, then naively, it takes N multiplications to evaluate f at one point and so takes $O(nN)$ operations to compute the Gauss-Jacobi approximation to the integral. To compute the power series expansion, we have to multiply N individual power series each of length n. Naively multiplication of two of these takes $O(n^2)$, but the fast Fourier transform allows this to be done in $O(n \log n)$. Thus about $o(Nn \log n)$ work is needed to compute the integral via power series, which is only slightly worse. This does not seem to be a decisive win for either method (especially since we probably only need $n \leq 20$ to attain machine precision), so any practical comparison boils down to estimating the size of the multiplicative constants implicit in the big-O estimates. This depends on the particular implementations, and we will not address it further.

So efficient numerical integration is possible if we can

(1) find $p_{n+1} \in \mathcal{P}_{n+1}$ so that $p_{n+1} \perp \mathcal{P}_n$ and $\|p_{n+1}\|_w = \langle p_{n+1} p_{n+1} \rangle_w = 1$,
(2) find the zeros $\{x_k\}$ of p_{n+1}
(3) find the weights $\{w_k\}$.

The first step is the main difficulty. Once we have the polynomial p, we can use Newton's method to find the roots of p_{n+1} and the weights are given by

$$w_k = -\frac{k_{n+1}}{k_n} \frac{1}{p_{n+1}(x_k) p_n'(x_k)}.$$

Suppose $\{p_k\}_0^n$ are orthonormal polynomials of degree k and the coefficient of x^k in p_k is c_k. We can find a polynomial (orthogonal to $\mathcal{P}_n$, but not necessarily of unit norm) p_{n+1} by taking any $(n+1)$st degree polynomial p and subtracting aways its orthogonal projection onto each of the 1-dimensional subspaces corresponding to these vectors, i.e.,

$$p_{n+1}(x) = p(x) - \sum_{k=0}^{n} p_k(x) \langle p, p_k \rangle_w.$$

Since we get to choose p, we take $p = xp_n$, so that

$$\begin{aligned} p_{n+1}(x) &= xp_n(x) - \sum_{k=0}^{n} p_k(x)\langle xp_n, p_k\rangle_w \\ &= xp_n(x) - p_n(x)\langle xp_n, p_n\rangle_w - p_{n-1}(x)\langle xp_n, p_{n-1}\rangle_w - \sum_{k=0}^{n-2} p_k(x)\langle p_n, xp_k\rangle_w \\ &= xp_n(x) - p_n(x)\langle xp_n, p_n\rangle_w - p_{n-1}(x)\langle xp_n, p_{n-1}\rangle_w \\ &= p_n(x)(x - \langle xp_n, p_n\rangle_w) - p_{n-1}(x)\langle xp_n, p_{n-1}\rangle_w \\ &= p_n(x)(x - a_n) - p_{n-1}(x)b_n \end{aligned}$$

We have used the facts that $\langle xf, g\rangle_w = \langle f, xg\rangle_w$ and that p_n is perpendicular to xp_k if $k < n-1$. The polynomial constructed is not necessarily of unit norm, but we can fix this by replacing p_{n+1} by

$$\frac{p_{n+1}}{\|p_{n+1}\|_w}.$$

To implement the method we have to be able to compute the recursion coefficients

$$\begin{aligned} a_n &= \langle xp_n, p_n\rangle_w \\ b_n &= \langle xp_n, p_{n-1}\rangle_w \\ c_n &= \|p_{n+1}\|_w = \|p_n(x - a_n) - p_{n-1}b_n\|_w. \end{aligned}$$

Recall that each of these inner products is an integral of the form

$$\int_a^b f(t)w(t)dt.$$

We already know p_n (by induction) so we could find its roots and use these to exactly evaluate such integrals for polynomials of degree $\leq 2n-1$. However, the inner products above involve polynomials of degree up to $2n+1$, and using the roots of p_n will definitely give a wrong answer for $\int_a^b tp_n^2(t)w(t)dt$. Therefore these coefficients should be computed by other means.

For evaluating Schwarz-Christoffel integrals, we will only need to consider the case of Jacobi weights with a singularity at one endpoint (or possibly neither endpoint), i.e., weights of the form $w(x) = (x-a)^\alpha$ on the interval $[a, b]$. However, we can compute an integral of the form

$$\int_a^b (\sum_{k=0}^{n} a_k x^k)(x-a)^\alpha dx,$$

using the following observation. A polynomial $p(x) = \sum_{k=0}^{n} a_k x^k$ has a Taylor expansion around any point, including the point a. This Taylor expansion, must also be a polynomial of degree n. Thus we can write

$$\sum_{k=0}^{n} b_k (x-a)^k = p(x) = \sum_{k=0}^{n} a_k x^k.$$

Then

$$\begin{aligned}
\int_a^b (\sum_{k=0}^{n} a_k x^k)(x-a)^\alpha dx &= \int_a^b \sum_{k=0}^{n} b_k (x-a)^k (x-a)^\alpha \\
&= \sum_{k=0}^{n} b_k \int_a^b (x-a)^{k+\alpha} dx \\
&= \sum_{k=0}^{n} b_k \frac{(b-a)^{k+\alpha+1}}{k+\alpha+1}.
\end{aligned}$$

So now we have to compute the $\{b_k\}$ from the $\{a_k\}$. Note that

$$\sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} a_k (x-a+a)^k = \sum_{k=0}^{n} a_k [\sum_{j=0}^{k} (x-a)^j a^{k-j} \binom{k}{j}],$$

so be get $\mathbf{b} = (b_0, \ldots, b_n)$ we just have to apply the matrix

$$M = (m_{jk}) = a^{k-j} \cdot \binom{k}{j},$$

to the vector $\mathbf{a} = (a_0, \ldots, a_n)$. This can be done naively in $O(n^2)$, but Chapter 3, Section 5 shows how to apply this $n \times n$ matrix to a vector in time only $O(n \log n)$.

If $w(t) = 1$, the Gauss-Jacobi polynomials specialize to the Legendre polynomials. Here are the first ten Legendre polynomials for the interval $[a, b] = [-1, 1]$, generated by the recursion $p_0 = 1\text{m}$ $p_{n+1} = ((2n+1)xp_n - np_{n-1})/(n+1)$:

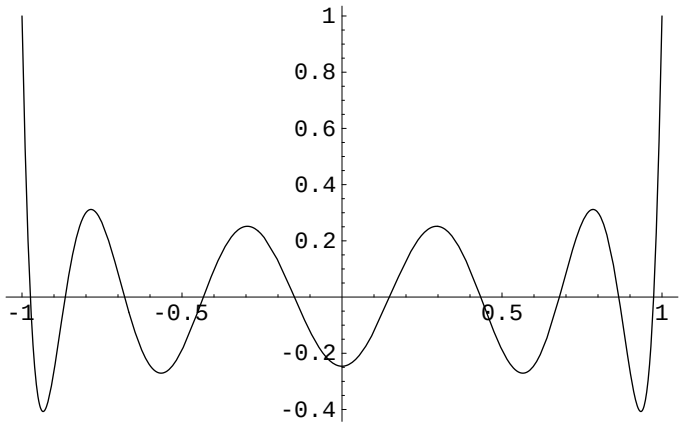
$$
\begin{aligned}
P_1(x) &= x, \\
P_2(x) &= -\left(\frac{1}{2}\right) + \frac{3\,x^2}{2} \\
P_3(x) &= \frac{-3\,x}{2} + \frac{5\,x^3}{2} \\
P_4(x) &= \frac{3}{8} - \frac{15\,x^2}{4} + \frac{35\,x^4}{8} \\
P_5(x) &= \frac{15\,x}{8} - \frac{35\,x^3}{4} + \frac{63\,x^5}{8} \\
P_6(x) &= -\left(\frac{5}{16}\right) + \frac{105\,x^2}{16} - \frac{315\,x^4}{16} + \frac{231\,x^6}{16} \\
P_7(x) &= \frac{-35\,x}{16} + \frac{315\,x^3}{16} - \frac{693\,x^5}{16} + \frac{429\,x^7}{16} \\
P_8(x) &= \frac{35}{128} - \frac{315\,x^2}{32} + \frac{3465\,x^4}{64} - \frac{3003\,x^6}{32} + \frac{6435\,x^8}{128} \\
P_9(x) &= \frac{315\,x}{128} - \frac{1155\,x^3}{32} + \frac{9009\,x^5}{64} - \frac{6435\,x^7}{32} + \frac{12155\,x^9}{128} \\
P_{10}(x) &= -\left(\frac{63}{256}\right) + \frac{3465\,x^2}{256} - \frac{15015\,x^4}{128} + \frac{45045\,x^6}{128} - \frac{109395\,x^8}{256} + \frac{46189\,x^{10}}{256}
\end{aligned}
$$

4	1.999984228457721944767532072144696487557194483115
5	2.000000110284471879766230094981509385528232424409
6	1.999999994772707155704064026794080767157032702 62
7	2.000000000001790471398897952280272539688258955 16
8	1.999999999999995360426618966771985355557337015 82
9	2.000000000000000009413606407259771939641429494 2
10	1.999999999999999999984637929765349118496057595 3
11	2.000000000000000000000020601345717327893623870 9
12	1.999999999999999999999999768928657480891021 33
13	2.000000000000000000000000000021998288119455387
14	1.999999999999999999999999999999820014656592 10
15	2.000000000000000000000000000000001279196248
16	1.999999999999999999999999999999999999202885
17	2.000000000000000000000000000000000000000439
18	2.000000000000000000000000000000000000000000

TABLE 1. Approximating $\frac{\pi}{2}\int_{-1}^{1}\cos(\frac{\pi}{2}t)dt$ using the roots of the nth Legendre polynomial.

4	1.0985703536493604213694507148231753197893152746 43
5	1.0986092418124719605204127411395244504262000897 26
6	1.0986120681169406437641500147652327985037897430 85
7	1.0986122727383456082370423301475424478891767017 7
8	1.0986122875191782529458652680660122250343151602 4
9	1.0986122885853231560758415201023000760438081814
10	1.0986122886621485872861135030048483168226650251
11	1.0986122886676806754603956463038864607321813301
12	1.0986122886680788273422876748043133557118525407
13	1.0986122886681074716527849715264721291023344 49
14	1.0986122886681095317892196695313251920927242 42
15	1.0986122886681096799212936635863512799859204 4
16	1.0986122886681096905705207364007642199981394 3
17	1.0986122886681096913359733724191428226330298 5
18	1.0986122886681096913909859076206421219228913
19	1.0986122886681096913949391857585049768764517
20	1.0986122886681096913952232475480128000949082
21	1.0986122886681096913952436571211548679735 54
22	1.0986122886681096913952451234301287602618 56
23	1.0986122886681096913952452287696849626202 0
24	1.0986122886681096913952452363368843107935 1
25	1.0986122886681096913952452368804590989262 2
26	1.0986122886681096913952452369195041669804
27	1.0986122886681096913952452369223086820056
28	1.0986122886681096913952452369225101173798
29	1.0986122886681096913952452369225245851 48
30	1.0986122886681096913952452369225256242 45

TABLE 2. Approximating $\log(3) = \int_{-1}^{1} \frac{1}{2+t} dt$ using the roots of the nth Legendre polynomial. *Mathematica* gives the first 50 digits of $\log 3$ as 1.0986122886681096913952452369225257046474905578227



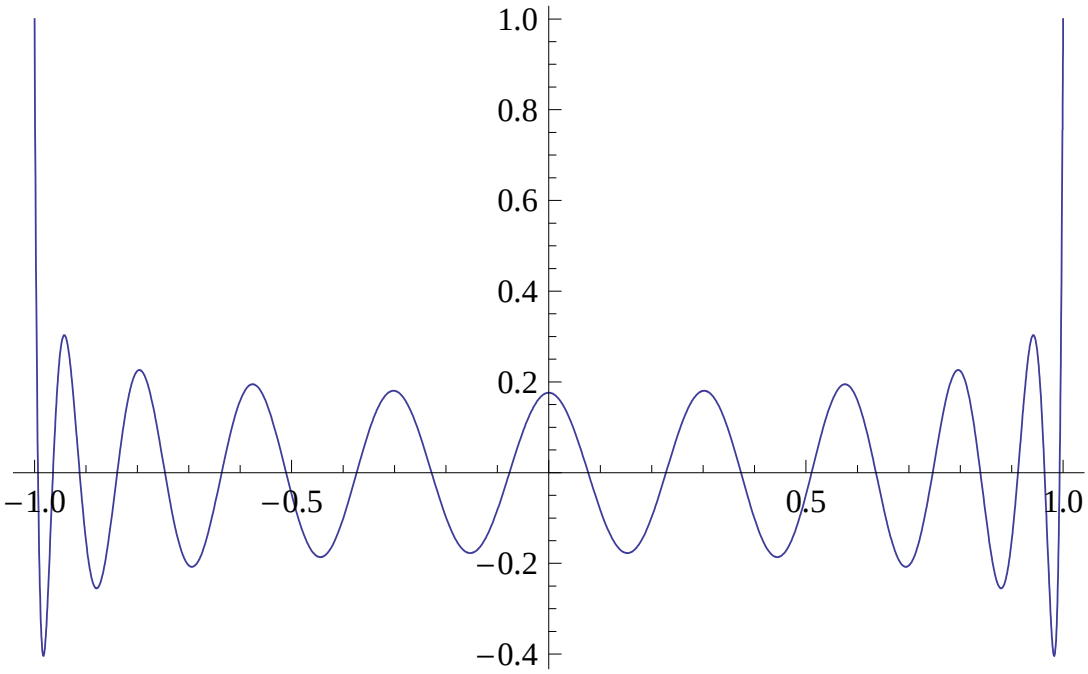


FIGURE 12. Some example of Legendre polynomials. The roots of P_n are the optimal n points to sample to compute an integral of the form $\int_{-1}^{1} f(t)dt$ in the sense that they will give the correct answer if f if a polynomial of degree at most $2n+1$. Shown are $n = 10, 20$.

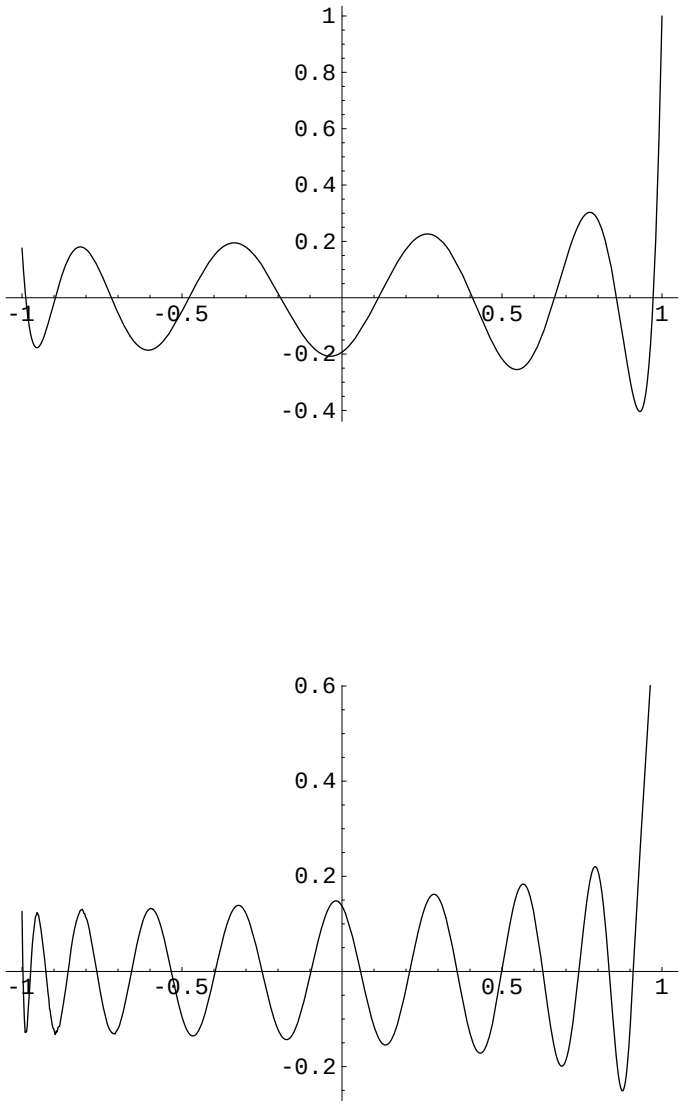


FIGURE 13. Some examples of Jacobi polynomials. This is P_{10} and P_{20} for the weight $(1+x)^{-1/2}$ (so there is only a singularity at the left end of the interval).

4. The fast Fourier transform

As we saw earlier in this chapter, efficient computation of the Schwarz-Christoffel formula requires efficient manipulation of power series (either to compute the power series of the map locally on each piece of our Carleson decomposition or to find the orthogonal polynomials used in the numerical integration). Most of these manipulations can be easily interpreted as applying a matrix to a vector and in this section we will review known results about applying structured matrices to vectors quickly. This summary is taken mostly from [**?**] and [**?**].

An n term power series centered at 0 has the form

$$p(z) = \sum_{k=0}^{n-1} a_n z^k,$$

and is polynomial of order $n-1$. Thus p is also determined by its values at any n distinct points $\{w_k\}_{k=1}^n$ and as we saw earlier can be recovered from these values using the Lagrange polynomials as

$$p(z) = \sum_{k=1}^{n} p(w_k) L_k(z),$$

where

$$L_k(z) = \prod_{1 \le j \le n, j \ne k} \frac{z - w_j}{w_k - w_j}.$$

Given the values $\{p(w_k)\}$, this formulation takes about $O(n^2)$ additions and multiplications to evaluate the coefficients of p. Similarly, given the coefficients of p it takes about $O(n^2)$ steps to evaluate it at the n points $\{w_k\}$. Can we do these conversions between the two representations of p faster. The answer is yes, at least if we assume the $\{w_k\}$ are special points, namely the nth roots of unity.

In this case, the problem of evaluating a polynomial or recovering its coefficients from its values is the same as computing a discrete Fourier transform. As before, fix a positive integer N and consider the Nth roots of unity $R_N = \{w_k\} = \{e^{i2\pi k/N}\}$. Let V^N be the N-dimensional complex vector space of functions from R_N to $\mathbb{C}$. The functions $e_k(z) = z^k$, $k = 0, \ldots N-1$ form an orthonormal basis with respect to the inner product

$$\langle f, g \rangle = \frac{1}{N} \sum_{k=0}^{N-1} f(w_k) \overline{g(z_k)}.$$

Therefore any function in V_N can be written

$$f(z) = \sum_{k=0}^{N-1} e_k \langle f, e_k \rangle = \sum_{k=0}^{N-1} a_k z^k,$$

for $z \in R_N$. Let $\mathrm{FFT}(N)$ denote the number of complex additions and multiplications to compute all N of the numbers $\{a_k\}$ given the N numbers $\{f(w_k\}$. Suppose that $N = 2M$. We claim that

$$\mathrm{FFT}(N) \leq 2\mathrm{FFT}(M) + O(N).$$

To prove this suppose $\{w_j\}_0^{N-1}$ are the Nth roots of unity and $\{v_j\}_0^{M-1}$ are the Mth roots of unity. Then write down the definition of the coefficients and split the sum into the sums over the even and odd terms:

$$\begin{aligned}
a_k &= \frac{1}{N} \sum_{j=0}^{N-1} f(w_j) w_j^k \\
&= \frac{1}{2M} \sum_{j=0}^{2M-1} f(w_j) w_j^k \\
&= \frac{1}{2M} \sum_{j=0}^{M-1} f(w_(2j) w_{2j}^k + \frac{1}{2M} \sum_{j=0}^{M-1} f(w_{2j+1}) w_{2j+1}^k \\
&= \frac{1}{2M} \sum_{j=0}^{M-1} f(v_(j) v_j^k + \frac{1}{2M} \sum_{j=0}^{M-1} f(v_j w_1) v_j^k w_1^k \\
&= \frac{1}{2}(b_k + w_i^k c_k),
\end{aligned}$$

where $\{b_k\}$ is the discrete Fourier transform on R_M of the function $g(z) = f(z)$ and $\{c_k\}$ is the discrete Fourier transform of the function $h(z) = f(w_1 z)$. Both $\{b_k\}$ and $\{c_k\}$ can be computed with $\mathrm{FFT}(M)$ operations and it clearly takes only $O(N)$ operations to combine them to get $\{a_k\}$. Thus the claim is proved. If $N = 2^n$ is a power of 2, then

$$\begin{aligned}
\mathrm{FFT}(N) &= 2\mathrm{FFT}(N/2) + O(N) \\
&= 4\mathrm{FFT}(N/4) + 2O(N/2) + O(N) \\
\ldots &= 2^n \mathrm{FFT}(2) + O(N) + O(N) + \cdots + O(N) \\
&= O(N \log N).
\end{aligned}$$

There are $O(N \log N)$ FFT algorithms for every N, but we shall shall only present the one above, for powers of 2.

Computing a discrete Fourier transform is the same as applying the Fourier matrix to a vector, where the Fourier matrix is given by

$$F_n = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \omega^{(n-1)(n-2)} \end{pmatrix}$$

where ω is an nth root of unity. The fast Fourier transform (FFT) applies F_n to an n-vector in time $O(n \log n)$ [**?**]. F_n is unitary and its conjugate transpose, F^*, can also be applied in $O(n \log n)$ time. The discrete Fourier transform (DFT) takes a n-long sequence of complex numbers $\{a_k\}_0^{n-1}$ and a n-root of unity ω and returns the values of the polynomial $p(z) = a_0 + a_1 x + \dots a_{n-1} z^{n-1}$ at the points $z = \{1, \omega, \omega^2, \dots, \omega^{n-1}\}$. Composing DFT with itself returns the original sequence times n. It turns our that being able to apply this matrix to a vector quickly is the key to many other fast computations.

5. Fast power series mainulations

Suppose $f(z) = \sum_{k=0}^{n} a_k z^k$ and $g(z) = \sum_{k=0}^{n} b_k z^k$. How fast can we multiply, divide or compose these series? Let $M(n)$ denote the number of field operations it takes to multiply two power series of length n. The usual process of convolving the coefficients shows $M(n) = O(n^2)$. A divide and conquer method of Karatsuba and Ofman [**?**] improves this to $O(n^\alpha)$ with $\alpha = \log 3 / \log 2$, but the fastest known method uses the Fast Fourier Transform [**?**], which shows $M(n) = O(n \log n)$ (two power series of length n can be multiplied by taking the DFT of each, multiplying the results term-by-term, taking the DFT of the result and finally dividing by n).

Other operations on power series are generally estimated in terms of $M(n)$. For example, inversion (finding the reciprocal power series, $1/f$, given the series for f) is $O(M(n))$. Like several other operations on power series, this is most easily proven using Newton's method (applied to series rather than numbers). For example, $1/f$ is the solution of the equation $\frac{1}{g} - f = 0$. If g is an approximate solution with $n > 0$

terms correct, then

$$g = g - \frac{\frac{1}{g} - f}{-1/g^2} = g - \frac{fg - 1}{z^n} g z^n,$$

has $2n$ correct terms. The right side requires two multiplications and so the work to compute inversions is $O(M(n)) + O(M(n/2)) + \cdots + O(1) = O(M(n))$.

Given inversion, one can divide power series (multiply f by $1/g$) compute $\log f$ (integrate f'/f term-by-term) or $\exp(f)$ (solve $\log g = f$ by Newton's method) all in time $O(M(n))$.

Composition of power series is a little harder. Brent and Kung showed that given power series f, g of order n and $g_0 = 0$, the composition $f \circ g$ can be computed in time $\mathrm{Comp}(n) = O(\sqrt{n \log n} M(n))$. Using FFT multiplication, this gives $O(n^{3/2} \log^{3/2} n \log\log n)$. They also showed that reversion (i.e., given f find g so $f \circ g(z) = z$) can be solved using Newton's method with the iteration

$$g \to g - \frac{f \circ g}{f' \circ g},$$

which doubles the number of correct terms in g with every step. Thus $\mathrm{Rev}(n) = O(\mathrm{Comp}(n)) = O(\sqrt{n \log n} M(n))$.

Fortunately, there are some special cases when composition is faster. For example, if we want to post-compose with a linear fractional transformation $\sigma(z) = (az + b)/(cz + d)$, this is the same as adding and dividing series, so is only $O(M(n))$.

Pre-composing by σ is more complicated. A function f is called algebraic if it satisfies

$$P_d(z) f(z)^d + \cdots + P_0(z) = 0,$$

for some polynomials $P_0, \cdots P_d$. Clearly every rational function is algebraic with $d = 1$. The power series of such functions satisfy linear recursions and n terms of the series can be computed in $O(n)$. Moreover, pre-composition by algebraic functions is fast; if f has p terms, g has q terms and is algebraic of degree d then the first n terms of $f \circ g$ can be computed in time $O(qd^2 \frac{p(q-v)}{n} M(n + pv) \log n)$ where v is the valuation of P_d (the largest power of z which divides $P_d(z)$) and q is the maximum of the degrees of P_i, plus 1. For linear fractional transformations $v = 0$ and $q = 2$ so the time to pre-compose by such a map is $O(M(n) \log n) = O(n \log^2 n)$. (There is an extensive generalization of the algebraic case to fast manipulations of holonomic

functions, as developed by van der Hoeven [**?**], although we do not need to use it here.)

This is too slow for our purposes. Fortunately, the only times we will have to pre-compose with a Möbius transformation correspond to various manipulations of power and Laurent series in the fast multipole method and all of these can be accomplished in $O(n \log n)$ by fast application of Toeplitz, Hankel and Pascal matrices as shown by Tang in [**?**]. A matrix is called circulant if each column is a down-shift of the previous one, is called Toeplitz if it is constant on diagonals and called Hankel if it is constant on antidiagonals. The general forms of these three types are:

$$C(x) = \begin{pmatrix} x_1 & x_n & x_{n-1} & \dots & x_2 \\ x_2 & x_1 & x_n & \dots & x_3 \\ x_3 & x_2 & x_1 & \dots & x_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n-1} & x_{n-2} & \dots & x_1 \end{pmatrix},$$

$$T_(x) = \begin{pmatrix} x_0 & x_1 & x_2 & \dots & x_{n-1} \\ x_{-1} & x_0 & x_1 & \dots & x_{n-2} \\ x_{-2} & x_n & x_0 & \dots & x_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{-n+1} & x_{-n+2} & x_{-n+3} & \dots & x_0 \end{pmatrix},$$

$$H(x) = \begin{pmatrix} x_{-n+1} & x_{-n+2} & x_{-n+3} & \dots & x_0 \\ x_{-n+2} & x_{-n+3} & x_{-n+4} & \dots & x_1 \\ x_{-n+3} & x_{-n+4} & x_{-n+5} & \dots & x_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0 & x_1 & x_2 & \dots & x_{n-1} \end{pmatrix}$$

A circulant matrix can be applied to a vector using three applications of FFT, i.e., because $C_n(x)$ applied to a vector y is the same as $\mathrm{IFFT}(\mathrm{FFT}(x) \cdot \mathrm{FFT}(y))$. A Toeplitz matrix can be embedded in a circulant matrix of the form

$$C_{2n} = \begin{pmatrix} T_n & S_n \\ S_n & T_n \end{pmatrix}$$

where

$$S_n = \begin{pmatrix} 0 & x_{-n+1} & x_{-n+2} & \dots & x_{-1} \\ x_{n-1} & 0 & x_{-n+1} & \dots & x_{-2} \\ x_{n-2} & & 0 & \dots & x_{-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \dots & 0 \end{pmatrix}$$

To apply T to an n-vector y, append n zeros to y to get a $2n$-vector, apply C_n and take the first n coordinates of the result. This takes $O(n \log n)$ time. If H is a Hankel matrix then $R \cdot H$ is a Toeplitz matrix where R is the permutation matrix which is 1's on the main anti-diagonal and 0 elsewhere, i.e., it reverses the order of the coordinates of a vector. Thus $H = R \cdot (R \cdot H)$, is a Toeplitz matrix followed by a permutation and can clearly be applied in time $O(n \log n)$ as well.

The Pascal matrix is lower triangular with its (j,k)th entry being the binomial coefficient $C_i^j = \binom{i}{j}$.

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{n-1}^0 & C_{n-1}^1 & C_{n-1}^2 & \dots & C_{n-1}^{n-1} \end{pmatrix}$$

This matrix can be written as $P = \mathrm{diag}(v_1) \cdot T \cdot \mathrm{diag}(v_2)$ where

$$v_1 = (1, 1, 2!, 3!, \dots, (n-1)!),$$

$v_2 = \frac{1}{v_1}$ (term-wise) and T is the Toeplitz matrix

$$T = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \frac{1}{2!} & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(n-1)!} & \frac{1}{(n-2)!} & \frac{1}{(n-3)!} & \dots & 1 \end{pmatrix}$$

The diagonal matrices can be applied in $O(n)$ and the Toeplitz in $O(n \log n)$ and hence so can P. Similarly for the transpose of P.

Now for the applications to fast multipole translation operators. There are three types of conversions to consider. First, local to local translation

$$\sum_{k=0}^{n-1} a_k (z-a)^k \quad \to \quad \sum_{k=0}^{n-1} b_k (z-b)^k,$$

then multipole to local

$$\sum_{k=0}^{n} a_n (z-a)^{-k} \quad \to \quad \sum_{k=0}^{n} b_n (z-b)^k,$$

and finally, multipole to multipole,

$$\sum_{k=0}^{n} a_n(z-a)^{-k} \quad \to \quad \sum_{k=0}^{n} b_n(z-b)^{-k}.$$

Let $c = b - a$ and consider the local-to-local translation. We have

$$\sum_{k=0}^{n-1} a_k(w-c)^k = \sum_{k=0}^{n-1} a_k \sum_{j=0}^{k} w^j(-c)^{k-j}\binom{k}{j},$$

so the matrix corresponding to local translation has kth column

$$((-c)^k, (-c)^{k-1}\binom{k}{1}, \ldots, (-c)^0\binom{k}{k}, 0, \ldots, 0)^t$$

or

$$LL = \begin{pmatrix} 1 & -c & c^2 & \ldots & (-c)^{n-1} \\ 0 & 1 & -2c & \ldots & (-c)^{n-2}C_{n-1}^1 \\ 0 & 0 & 1 & \ldots & (-c)^{n-3}C_{n-1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

This matrix is equal to

$$\mathrm{diag}(1, -\mathrm{z}, \ldots, (-\mathrm{z})^{\mathrm{n}-1}) \cdot \mathrm{P}' \cdot \mathrm{diag}(1, -\mathrm{z}^{-1}, \ldots (-\mathrm{z})^{-\mathrm{n}+1}),$$

where P' is the transpose of P. The diagonal matrices can be applied in $O(n)$ time and P' can be applied in $O(n \log n)$. Thus local-to-local translations can be done this fast.

Similarly, the multipole-to-multipole and multipole-to-local transformations correspond to applying the matrices

$$MM = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \binom{1}{1}c & 1 & 0 & \ldots & 0 \\ \binom{2}{2}c^2 & \binom{2}{1}c & 1 & \ldots & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{n-1}{n-1}c^{n-1} & \binom{n-1}{n-2}c^{n-2} & \binom{n-1}{n-3}c^{n-3} & \ldots & 1 \end{pmatrix}$$

$$ML = \begin{pmatrix} -c^{-1} & c^{-2} & c^{-3} & \ldots & c^{-n+1} \\ -c^{-2} & 2c^{-3} & -3c^{-4} & \ldots & \\ -c^{-3} & 3c^{-4} & -6c^{-5} & \ldots & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -c^{-n+1} & (n-1)c^{-n} & -\binom{n}{2}c^{-n-1} & \ldots & (-1^{n-1}\binom{2p-2,p-1}{c})^{-2n-1} \end{pmatrix}.$$

We can rewrite these matrices as

$$MM = \mathrm{diag}(1, \mathrm{z}, \ldots, \mathrm{z}^{\mathrm{n}-1}) \cdot \mathrm{P} \cdot \mathrm{diag}(1, \mathrm{z}^{-1}, \ldots \mathrm{z}^{-\mathrm{n}+1}),$$

$$ML = \mathrm{diag}(1, \mathrm{z}^{-1}, \ldots, \mathrm{z}^{1-\mathrm{n}}) \cdot \mathrm{P} \cdot \mathrm{P}' \cdot \mathrm{diag}(-\mathrm{z}^{-1}, \mathrm{z}^{-2} \ldots (-\mathrm{z})^{-\mathrm{n}}),$$

where P' is the transpose of P. As with local translations, these are compositions of diagonal matrices (which can be applied in $O(n)$) and matrices that can be applied in $O(n \log n)$ time.

We will also use structured matrices to compute expansions around ∞ of functions of the form $\int \frac{d\mu(z)}{(z-w)^k}$, $k = 1, 2, 3$. We will only consider the Cauchy transform ($k = 1$) since the others can be obtained by term-by-term differentiation of that one. Suppose $f(z) = \sum_{k=0}^{n} a_k z^k$ is a power series for an analytic function, bounded by one and defined on $\mathbb{D}$ and $\varphi(x, y)$ is a polynomial in x and y of uniformly bounded degree. Then the Cauchy transform

$$F(w) = \int_S \frac{f(z)\varphi(x, y)dxdy}{z - w},$$

is analytic in w outside $S = [-\frac{1}{2}, \frac{1}{2}]^2$, so has an expansion $F(w) = \sum_{k=1}^{\infty} b_n w^{-n}$. Given $\{a_k\}_0^p$, thinking of φ as fixed, we want to compute $\{b_k\}_1^n$. For each monomial of the form $z^k x^a y^b$ we can precompute the expansion using explicit formulas ($O(n)$ for each of $O(n)$ monomials) and then we simply apply the resulting matrix to the vector $\{a_k\}$. Naively, we can do this in time $O(n^2)$.

Actually we can compute the expansion in only $O(n \log n)$. Let $d\mu = x^a y^b dxdy$ restricted to $Q = [0, 1]^2$. We want to compute the expansion at ∞ of

$$\begin{aligned} F(w) = \iint \frac{z^n}{w - z} d\mu(z) &= \iint z^n \frac{1}{w}(1 + \frac{z}{w} + (\frac{z}{w})^2 + \ldots)d\mu(z) \\ &= \sum_{k=0}^{\infty} w^{-k-1} \iint z^{n+k} d\mu(z) \\ &= \sum_{k=1}^{\infty} a_{k,n} w^{-k}, \end{aligned}$$

where

$$a_{k,n} = c(n + k + 1, a, b) = \iint_Q (x + iy)^{n+k-1} x^a y^b dxdy.$$

Since $a_{k,n}$ only depends on $k + n$, A is a Hankel matrix. As noted above, a $n \times n$ Hankel matrix can be applied to a n-vector using FFT in time $O(n \log n)$.

The individual coefficients have explicit formulas involving Euler's Beta function. Evaluations for a few small values of a, b (as given by *Mathematica* are)

$$c(n,0,0) = \frac{i - i^{n+1} + 2(1+i)^n}{2+3n+n^2},$$

$$c(n,1,0) = \frac{2i - i^n + in + 2(1+I)^n((2-i)+n)}{(1+n)(2+n)(3+n)},$$

$$c(n,2,0) = \frac{i(6+2i^n+5n+n^2) + 2(1+i)^n((4-4i)+n((5-2i)+n))}{(1+n)(2+n)(3+n)(4+n)}$$

$$c(n,1,1) = -\frac{1+i^n-2(1+i)^n(2+n)}{(1+n)(2+n)(4+n)},$$

$$c(n,2,1) = \frac{-3(2i)i^n - n + 2(1+i)^n(1+n)((4-i)+n)}{(1+n)(2+n)(3+n)(5+n)},$$

$$c(n,0,1) = \frac{-1 - i^{n+1}(2+n) + 2(1+i)^n((2+i)+n)}{(1+n)(2+n)(3+n)}.$$

Thus n-term Laurent expansions for Beurling transforms of the appropriate degree n polynomials can can be computed in time $O(n \log n)$.

CHAPTER 4

Some geometric function theory

In this chapter and the next we introduce ideas that are fundamental to geometric function theory: modulus, capacity, distortion estimates and quasiconformal maps.

1. Conformal modulus

A conformal invariant is a number which is invariant under conformal mappings. We are often in the situation where we wish to know the value of some conformal invariant (e.g., that harmonic measure of the edge of a polygon) and are able to estimate some other conformal invariant (e.g., the modulus of some path family in the polygon). Using a known relation between the invariants, we can turn an esitmate for one into an estimate for the other.

Probably the most important example of a conformal invariant is the (conformal) modulus.

Suppose Γ is a family of locally rectifiable paths in a planar domain Ω and ρ is a non-negative Borel function on Ω. We say ρ is admissible for Γ if

$$\ell(\Gamma) = \inf_{\gamma\in\Gamma} \int_\gamma \rho ds \geq 1,$$

and define the modulus of Γ as

$$\mathrm{Mod}(\Gamma) = \inf_{\Omega} \int_M \rho^2 dxdy,$$

where the infimum is over all admissible ρ for Γ. This is a well known conformal invariant whose basic properties are discussed in many sources such as Ahlfors' book [**?**]. It reciprocal is called the extremal length of the path family. Modulus and extremal length satisfy several properties that are helpful in estimating these quantities.

LEMMA 47 (Conformal invariance). *If $\mathcal{F}$ is a family of curves in a domain Ω and f is a one-to-one analytic mapping from Ω to Ω' then $M(\mathcal{F}) = M(f(\mathcal{F}))$.*

PROOF. This is just the change of variables formulas

$$\int_{f(\gamma)} \rho \circ f ds = \int_\gamma \rho ds,$$

$$\int_{f(\Omega)} (\rho \circ f)^2 dxdy = \int_\Omega \rho dxdy.$$

These imply that if $\rho \in \mathcal{A}(\mathcal{F})$ then $\rho \circ f^{-1} \in \mathcal{A}(f(\mathcal{F}))$, and thus $M(f(\mathcal{F})) \leq M(\mathcal{F})$. We get the other direction by considering f^{-1}. □

LEMMA 48 (Monotonicity). *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are collections such that every $\gamma \in \mathcal{F}_1$ contains some curve in $\mathcal{F}_2$ then $M(\mathcal{F}_1) \leq M(\mathcal{F}_2)$ and $\lambda(\mathcal{F}_1) \geq \lambda(\mathcal{F}_2)$.*

The proof is immediate since $\mathcal{A}(\mathcal{F}_1) \supset \mathcal{A}(\mathcal{F}_2)$.

LEMMA 49 (Grötsch Principle). *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are families of curves in disjoint domains then $M(\mathcal{F}_1 \cup \mathcal{F}_2) = M(\mathcal{F}_1) + M(\mathcal{F}_2)$.*

LEMMA 50. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are families of curves in disjoint domains and every curve of $\mathcal{F}$ contains both a curve from $\mathcal{F}_1$ and $\mathcal{F}_2$, then $\lambda(\mathcal{F}) \geq \lambda(\mathcal{F}_1) + \lambda(\mathcal{F}_2)$.*

PROOF. If $\rho_1 \in \mathcal{A}(\mathcal{F}_i)$ for $i = 1, 2$, then $\rho = t\rho_1 + (1-t)\rho_2$ is admissible for $\mathcal{F}$. Since the domains are disjoint we may assume $\rho_1\rho_2 = 0$ everywhere so taking

$$t = t^2 M(\mathcal{F}_1) + (1-t^2)M(\mathcal{F}_2),$$

gives

$$m(\mathcal{F}) \leq t^2 M(\mathcal{F}_1) + (1-t^2)M(\mathcal{F}_2) = (M(\mathcal{F}_1)^{-1} + M(\mathcal{F}_2)^{-1})^{-1},$$

as required. □

The fundamental example is to compute the modulus of the path family connecting opposite sides of a $a \times b$ rectangle; this serves as the model of almost all modulus estimates. So suppose $R = [0, b] \times [0, a]$ is a a long and b high rectangle and Γ consists of all rectifiable curves in R with one endpoint on each of the sides of length a. Then each such curve has length at least b, so if we let ρ be the constant $1/b$ function on R we have

$$\int_\gamma \rho ds \geq 1,$$

for all $\gamma \in \Gamma$. Thus this metric is admissible and so

$$\text{Mod}(\Gamma) \leq \iint_T \rho^2 dxdy = \frac{1}{b^2} ab = \frac{a}{b}.$$

To prove a lower bound, we use the well known Cauchy-Schwarz inequality:

$$(\int fgdx) \leq (\int f^2dx)(\int g^2dx).$$

To apply this, suppose ρ is an admissible metric on R for γ. Every horizontal segment in R connecting the two sides of length a is in Γ, so since γ is admissible,

$$\int_0^b \rho(x,y)dx \geq 1,$$

and so by Cauchy-Schwarz

$$1 \leq \int_0^b (1 \cdot \rho(x,y))dx \leq \int_0^b 1^2dx \cdot \int_0^b \rho^2(x,y)dx.$$

Now integrate with respect to y to get

$$\int_0^a 1dy \leq b \int_0^a \int_0^b \rho^2(x,y)dxdy,$$

or

$$\frac{a}{b} \leq \iint_R \rho^2 dxdy,$$

which implies $\text{Mod}(\Gamma) \geq \frac{b}{a}$. Thus we must have equality.

Another useful computation is the modulus of the family of path connecting the inner and out boundaries of the annulus $A = \{z : r < |z| < R\}$. An argument similar to the one above shows that the modulus of this family is $\frac{1}{2\pi} \log \frac{R}{r}$.

2. Modulus and cross ratio; rectangles revisited

A generalized quadrilateral Q is a Jordan domain in the plane with four specified boundary points x_1, x_2, x_3, x_4 (in counterclockwise order). We define the modulus of Q, $M_Q(x_1, x_2, x_3, x_4)$ (or just M_Q or $M(Q)$ if the points are clear from context), as the modulus of the path family in Q which connects the arc (x_1, x_2) to the arc (x_3, x_4). This is also the unique positive real number M such that Q can be conformally mapped to a $1 \times M$ rectangle with the arcs (x_1, x_2), (x_3, x_4) mapping to the opposite sides of length 1. In this paper, we will be particularly concerned with the case when $Q = \mathbb{D}$ and we are given four points in counterclockwise order on the unit circle.

Given a generalized quadrilateral Q with four boundary points x_1, x_2, x_3, x_4, the quadrilateral Q' with vertices x_2, x_3, x_4, x_1 is called the reciprocal of Q and it is easy to see that $\text{Mod}(Q') = 1/\text{Mod}(Q)$. One way to compute the modulus is to define

the harmonic function on Q which is 0 on (x_1, x_2) and 1 on (x_3, x_4) and has normal derivative zero on the other two arcs. Then

$$\mathrm{Mod}(Q) = \iint_Q |\nabla u|^2 dx dy.$$

(This is clear on a rectangle and it is easy to check that both sides are conformal invariants.)

If the four points lie on $\mathbb{T}$, then since cr and $M_{\mathbb{D}}$ are both invariant under Möbius transformations of the disk to itself, each must be a function of the other in this case. The usual way to represent this function (e.g., as in Ahlfors' book [**?**]) is to map the disk to the upper half plane, $\mathbb{H}$, sending the points a, b, c to $0, 1, \infty$ respectively and d to $-P = \mathrm{cr}(a, b, c, d) \in (-\infty, 0)$. Then $M_{\mathbb{D}}(a, b, c, d)$ is the same as the modulus of the path family in $\mathbb{H}$ connecting $(-\infty, -P)$ to $(0, 1)$. By symmetry, this is twice the modulus of the path family of closed curves in the plane which separate $[-P, 0]$ from $[1, \infty]$. We will denote this modulus by $M(P)$. The transformation $z \to (z-1)/(z+P)$ sends $0, 1, \infty, -P$ to $-\frac{1}{P}, 0, 1, \infty$, so by Möbius invariance of modulus and the fact that conjugate quadrilaterals have reciprocal moduli, we see that

$$M_{\mathbb{H}}(\frac{1}{P}) = \frac{1}{M_{\mathbb{H}}(P)}, \tag{13}$$

and hence $M_{\mathbb{H}}(1) = 1$ and $M(1) = 1/2$.

We saw earlier that the the relation is given by

$$P + 1 = \exp(2\pi M)\frac{1}{16}\prod_{n=1}^{\infty}(\frac{1 + \exp((1-2n)2\pi M)}{1 + \exp((-2n)2\pi M)})^8. \tag{14}$$

For $M > 0$ the infinite product converges and for M large (say $M \geq 1$) we have

$$\prod_{n=1}^{\infty}(\frac{1 + \exp((1-2n)2\pi M)}{1 + \exp((-2n)2\pi M)})^8 = 1 + 8e^{-2\pi M} + O(e^{-4\pi M}).$$

Thus for $P \geq 1$, (equivalently $M \geq 1$), we have

$$\log(P + 1) = 2\pi M - \log 16 + 8e^{-2\pi M} + O(e^{-4\pi M}),$$

which implies

$$P \simeq \exp(2\pi M).$$

For $0 < P \leq 1$, (equivalently $0 < M \leq 1$), we can use (13) to deduce

$$\log(\frac{1}{P} + 1) = \frac{\pi}{2M} - \log 16 + 8e^{-\pi/(2M)} + O(e^{-\pi/M}),$$

which implies

$$P \simeq \exp(-\frac{\pi}{2M}).$$

In other words,

$$M \simeq \frac{1}{2\pi} \log P, \qquad P \gg 1,$$
$$M \simeq \frac{\pi}{2|\log P|}, \qquad P \ll 1,$$

Thus for $\mathbf{x} = \{x_1, x_2, x_3, x_4\} \subset \mathbb{T}$, since $\mathrm{Mod}_{\mathbb{D}} = 2M$,

$$M_{\mathbb{D}}(\mathbf{x}) \simeq \frac{1}{\pi} \log |\mathrm{cr}(\mathbf{x})|, \qquad |\mathrm{cr}(\mathbf{x})| \gg 1,$$
$$M_{\mathbb{D}}(\mathbf{x}) \simeq \frac{\pi}{|\log |\mathrm{cr}(\mathbf{x})||}, \qquad |\mathrm{cr}(\mathbf{x})| \ll 1,$$

The conformal modulus of the path family connecting the length 1 sides in a $1 \times R$ rectangle is R. By conformal invariance this must also be the modulus of the unit disk with the four prevertices. If we assume these prevertices are symmetric with respect to both the real and imaginary axes, then their arguments can be easily computed from their cross ratio. However, the cross ratio is a function of R as given by (14), so we can solve the SC parameter problem for rectangles with this function.

Another elegant connection between modulus and cross ratios is given in [**?**] (Bagby shows that conformal modulus for a ring domain is given by minimizing an integral involving logarithms of cross ratios).

3. Pfluger's theorem and Beurling's estimate

The usefulness of extremal length is its ability to estimate a conformal invariant in terms of geometry (length and area). Our main application of this idea is the following special case of a theorem of Pfluger:

THEOREM 51. *Suppose $K \subset \mathbb{D}$ is compact with a smooth boundary and contains 0 in its interior and and $E \subset \partial\mathbb{D}$ is compact and let $\mathcal{F}$ be the family of curves in $\mathbb{D} \setminus K$ separating 0 from E. Then there is a C (depending only on K) such that $\mathcal{H}^1(E) \leq C \exp(-\pi M(\mathcal{F}))$.*

This is not sharp since the right hand side can be positive for sets of zero length. The sharp version uses logarithmic capacity in place of $\mathcal{H}^1$ measure on the left hand side. See Exercise **??** for the sharp version of Pfluger's theorem.

PROOF. We assume E has positive length, since there is nothing to prove otherwise. Furthermore, we may assume E is actually a finite union of closed intervals. Let μ be Lebesgue measure restricted to E, normalized to have mass 1. Define the potential function

$$U_\mu(z) = \log \frac{1}{|z|} * \mu,$$

and set

$$v(z) = U_\mu(z) + U_\mu(\frac{1}{z}).$$

Then v is symmetric with respect to the unit circle, has negative logarithmic poles at 0 and ∞ and bounded above by $2|\log \mathcal{H}^1(E)| + O(1)$ (to prove this note that since $-\log|z|$ is decreasing with $|z|$, the integral defining U is maximized when E is an interval and z is its midpoint). Since v is symmetric with respect to $\mathbb{T}$ we have $\partial v/\partial n = 0$ on $\partial\mathbb{D} \setminus E$. Now suppose $\gamma \in \mathcal{F}$ and let Ω be the component of $\mathbb{D} - \gamma$ containing 0. Since v is harmonic in $\mathbb{D}$ except for a logarithmic pole at 0 we can apply Green's theorem to get

$$\int_\gamma |\nabla v| ds \;\; \geq \;\; -\int_\gamma \frac{\partial v}{\partial n} ds = \lim_{\epsilon \to 0} \int_{|z|=\epsilon} \frac{\partial v}{\partial n} ds = 2\pi.$$

Thus $|\nabla v|/2\pi \in \mathcal{A}(\mathcal{F})$, so

$$M(\mathcal{F}) \leq \int_{\mathbb{D}\setminus K} (\frac{|\nabla v|}{2\pi})^2 dxdy.$$

Note that $|\nabla v|^2 = \frac{1}{2}\Delta(v^2)$, and use Green's theorem

$$\begin{aligned} \iint_{\mathbb{D}\setminus K} |\nabla v|^2 dxdy &= \frac{1}{2} \iint_{\mathbb{D}\setminus K} \Delta(v^2) dxdy \\ &= -\frac{1}{2} \int_{\partial\mathbb{D}} \frac{\partial v^2}{\partial n} ds - \frac{1}{2} \int_{\partial K} \frac{\partial v^2}{\partial n} ds \\ &\leq -(\max_{\mathbb{T}} v) \int_{\partial\mathbb{D}} \frac{\partial v}{\partial n} ds + O(1) \end{aligned}$$

Using Green's theorem gives

$$-\int_{\partial\mathbb{D}} \frac{\partial v}{\partial n} ds = 2\pi$$

Hence,

$$M(\mathcal{F}) \leq \frac{1}{(2\pi)^2} |2\log \mathcal{H}^1(E)| 2\pi + O(1),$$

as desired. □

COROLLARY 52. *Suppose Ω is a Jordan domain, $z_0 \in \Omega$ with $\operatorname{dist}(z_0, \partial\Omega) \geq 1$ and $E \subset \partial\Omega$. Let $\mathcal{F}$ be the family of curves in Ω which separates $D(z_0, 1/2)$ from E. Then $\omega(z_0, E, \Omega) \leq C \exp(-\pi M(\mathcal{F}))$.*

PROOF. Since both harmonic measure and modulus are conformally invariant we need only verify this when $\Omega = \mathbb{D}$. But this is the previous theorem. □

If $E \subset \partial\Omega$ is an arc then the inequality is actually a similarity.

COROLLARY 53 (Ahlfors distortion theorem). *Suppose Ω is a Jordan domain, $z_0 \in \Omega$ with $\operatorname{dist}(z_0, \partial\Omega) \geq 1$ and $x \in \partial\Omega$. For each $0 < t < 1$ let $\theta(t)$ be the length of $\Omega \cap \{|w - x| = t\}$. Then there is an absolute $C < \infty$, so that*

$$\omega(z_0, D(x, r), \Omega) \leq C \exp(-\pi \int_r^1 \frac{dt}{\theta(t)}).$$

PROOF. Let K be the disk of radius $1/2$ around z_0 and let $\mathcal{F}$ be the family of curves in Ω which separate $D(x, r) \cap \partial\Omega$ from K. Let $\mathcal{F}_1 \subset \mathcal{F}$ be the collection of curves of the form

$$L_t = \Omega \cap \{|w - x| = t\}.$$

if ρ is admissible for $\mathcal{F}$ then it is admissible for $\mathcal{F}_1$ and hence

$$1 \leq \int_{L_t} \rho ds \leq (\int_{L_t} \rho^2 ds)\theta(t),$$

so

$$\int_r^1 \int_{L_t} \rho^2 ds dt \geq \int_r^1 \frac{dt}{\theta(t)}.$$

This proves

$$M(\mathcal{F}) \geq \int_r^1 \frac{dt}{\theta(t)},$$

which proves the result by the previous corollary. □

For an alternate version of this using line segments instead of circular arcs, see Exercise **??**.

COROLLARY 54 (Beurling's estimate). *There is a $C < \infty$ so that if Ω is simply connected, $z \in \Omega$ and $d = \operatorname{dist}(z, \partial\Omega)$ then for any $0 < r < 1$ and any $x \in \partial\Omega$,*

$$\omega(z, D(x, rd), \Omega) \leq C r^{1/2}$$

PROOF. Apply Corollary 53 at x and use $\theta(t) \le 2\pi t$ to get

$$\exp(-\pi \int_{rd}^{d} \frac{dt}{\theta(t)t}) \le C \exp(-\frac{1}{2}\log r) \le C\sqrt{r}.$$

□

LEMMA 55. *If Ω is simply connected then*

$$\omega(z_0, D(x,r), \Omega) \le C[\frac{r}{\operatorname{dist}(z, \partial\Omega)}]^{1/2}.$$

PROOF. Use extremal length. □

COROLLARY 56. *If Ω is simply connected and $w \in \partial\Omega$, then*

$$\int_{D(w,r)\cap\partial\Omega} |\log \frac{1}{|z-w|} d\omega_{z_0}(z) \le C[\frac{r}{\operatorname{dist}(z, \partial\Omega)}]^{1/2}.$$

PROOF. Cut the disk into concentric annuli $a_n = \{z : 2^{-n}r \le |z-w| \le 2^{-n+1}r\}$. By Beurling estimate the singleton $\{w\}$ has zero harmonic measure, so $\int_D = \sum \int_{A_n}$. However, the integral over A_n is bounded by $n2^{-n/2} \log \frac{1}{r} \log 2$ which sums to $O(|\log r/\operatorname{dist}(z_0, \partial\Omega)|)$. □

4. Logarithmic capacity

Suppose μ is a positive Borel measure on $\mathbb{R}^2$ and define its energy integral by

$$I(\mu) = \iint \log \frac{2}{|z-w|} d\mu(z) d\mu(w).$$

We put the "2" in the numerator so that the integrand is non-negative when $z, w \in \mathbb{T}$ (in this paper we will only consider the capacity of subsets of $\mathbb{T}$). If $E \subset \mathbb{R}^2$ is Borel, let $\operatorname{Prob}(E)$ be the set of positive Borel measures with $\mu(E) = \|\mu\| = 1$ and define its logarithmic capacity as

$$\operatorname{cap}(E) = \frac{1}{\inf\{I(\mu) : \mu \in \operatorname{Prob}(E)\}}.$$

For subsets of the circle, cap is non-negative, monotone and is countably subadditive (Lemma 4, page 24, [**?**]; this is where we need the "2" in the definition of the energy integral). If $\operatorname{cap}(E) > 0$, there is a unique measure which minimizes the energy integral, which is called the equilibrium measure (it is also equal to the harmonic

measure of $S^2 \setminus E$ with respect to infinity). An alternate version of logarithmic capacity is

$$\widetilde{\text{cap}}(E) = \sup\{\exp(-I(\mu)) : \mu \in \text{Prob}(E)\}.$$

The exponential in the definition is a technical convenience and gives it nice scaling, i.e., $\widetilde{\text{cap}}(tE) = t \cdot \widetilde{\text{cap}}(E)$. The two versions of logarithmic capacity are related by the equations

$$\widetilde{\text{cap}} = \exp(-1/\text{cap}), \quad \text{cap} = (\log \widetilde{\text{cap}}^{-1})^{-1}.$$

Note that if $E \subset \mathbb{T}$ then $\widetilde{\text{cap}}(E) = 0$ iff $\text{cap}(E) = 0$. Thus we may speak of sets of positive or zero capacity without specifying which definition we mean and we will use both versions throughout the paper.

Logarithmic capacity is closely related to the usual Robin constant γ_E defined by

$$\gamma_E = \inf\{I(\mu) - \log 2 : \mu \in \text{Prob}(E)\} = \frac{1}{\text{cap}(E)} - \log 2.$$

The $\log 2$ enters because we put a "2" in our energy integral, whereas the usual definition does not.

If $f : \mathbb{D} \to \Omega$ is conformal and $E \subset \partial\Omega$ then we will call $\text{cap}(f^{-1}(E))$ the capacity of E with respect to Ω (the value depends on the choice of f, but whether or not it is zero is independent of f).

The connection between extremal length and logarithmic capacity is given by the following result, Pfluger's Theorem, e.g., Theorem 9.17 of [**?**],

LEMMA 57. *Suppose $E \subset \mathbb{T}$ is compact, $K \subset \mathbb{D}$ is compact and connected and $\mathcal{P}$ is the path family in $\mathbb{D}$ connecting K to E. Then $\widetilde{\text{cap}}(E) \simeq \exp(-\pi\lambda(\mathcal{P}))$, with constants that depend only on K.*

One particular consequence is the following.

COROLLARY 58. *If f is a conformal map on $\mathbb{D}$ and takes the boundary value 0 at every point of $E \subset \mathbb{T}$, then $\text{cap}(E) = 0$*

PROOF. Suppose $K \subset \mathbb{D}$ is compact and choose r so small that $D(0,r) \cap f(K) = \emptyset$. Then the extremal length of the path family connecting K to E in $\mathbb{D}$ is greater than for the family crossing the annulus $\{z : \epsilon < |z| < r\}$ in Ω. Taking $\epsilon \to 0$ and using the estimate for annuli discussed above proves the result. □

Using Lemma 62, which we will prove later, one can show that it suffices to assume f has radial limit 0 on E in Lemma 58.

Suppose $\partial\Omega$ is bounded in $\mathbb{R}^2$ and $f : \mathbb{D} \to \Omega$ is conformal. For $0 < r < 1$, let

$$a_f(r) = \text{area}(\Omega \setminus f(D(0,r))).$$

Since $\partial\Omega$ is compact it is easy to see that this tends to zero as $r \to 1$.

LEMMA 59. *There is a $C < \infty$ so that the following holds. Suppose $f : \mathbb{D} \to \Omega$ and $\frac{1}{2} \le r < 1$. Let $E = \{x \in \mathbb{T} : |f(sx) - f(rx)| \ge \delta \text{ for some } r < s < 1\}$. Then the extremal length of the path family $\mathcal{P}$ connecting $D(0,r)$ to E is bounded below by $\delta^2/Ca(r)$.*

PROOF. Suppose $z, w \in \Omega$, suppose γ is the hyperbolic geodesic connecting z and w and suppose $\tilde{\gamma}$ is any path in Ω connecting these points. By the Gehring-Hayman inequality [**?**], there is a universal $C < \infty$ such that $\ell(\gamma) \le C\ell(\tilde{\gamma})$ (here $\ell(\gamma)$ denotes the length of γ). In other words, up to a constant, the hyperbolic geodesic has the shortest Euclidean length amongst all curves in Ω connecting the two points.

Now suppose we apply this with $z = f(sx)$ and $w \in f(D(0,r))$. Then the length of any curve from w to z is at least $1/C$ times the length of the hyperbolic geodesic γ between them. But this geodesic has a segment γ_0 that lies within a uniformly bounded distance of the geodesic γ_1 from $f(rx)$ to z. By the Koebe distortion theorem γ_0 and γ_1 have comparable Euclidean lengths, and clearly the length of γ_1 is at least δ. Thus the length of any path from $f(D(0,r))$ to $f(sx)$ is at least δ/C. Now let $\rho = C/\delta$ in $\Omega \setminus f(D(0,r))$ and 0 elsewhere. Then ρ is admissible for $f(\mathcal{P})$ and $\iint \rho^2 dxdy$ is bounded by $C^2 a(r)/\delta^2$. Thus $\lambda(\mathcal{P}) \ge \frac{\delta^2}{C^2 a(r)}$. □

If f has radial limits on $E \subset \mathbb{T}$ then the previous lemma is still valid for $s = 1$. For subsets of the circle it is known that

$$\widetilde{\text{cap}}(E) \ge |E|/C, \tag{15}$$

(e.g., XI.2.E in [**?**]). Combining this with Lemmas and 57 and 59 gives

COROLLARY 60. *If $f : \mathbb{D} \to \Omega$ is a conformal map onto a bounded domain then for any $\delta > 0$,*

$$|\{x \in \mathbb{T} : |f(x) - f(rx)| \ge \delta\}| \to 0,$$

as $r \to 1$. Thus a conformal map of $\mathbb{D}$ with bounded image has radial limits Lebesgue almost everywhere.

LEMMA 61. *Suppose $f : \mathbb{D} \to \Omega$ is conformal and for $R \geq 1$,*

$$E = \{x \in \mathbb{T} : |f(x)| \geq R \operatorname{dist}(f(0), \partial\Omega)\}.$$

Then $\operatorname{cap}(E) \leq CR^{-1/2}$ (with C independent of Ω).

PROOF. Assume $f(0) = 0$ and $\operatorname{dist}(0, \partial\Omega) = 1$ and let $\rho(z) = |z|^{-1}/\log R$ for $z \in \Omega \cap \{1 < |z| < R\}$. Then ρ is admissible for the path family connecting $D(0, 1/2)$ to $\partial\Omega \setminus D(0, R)$ and $\iint \rho^2 dxdy \leq 2\pi/\log R$. By the Koebe distortion theorem $f^{-1}(D(0, 1/2))$ is contained in a compact subset of $\mathbb{D}$, independent of Ω. The result follows by Lemma 57. □

Lemma 61 also follows from a stronger result of Balogh and Bonk in [**?**].

Given a compact set $E \subset \mathbb{T}$ we will now define the associated "sawtooth" region W_E and a 2-quasiconformal map between W_E and $\mathbb{D}$ which keeps E fixed pointwise. Suppose $\{I_n\}$ are the connected components of $\mathbb{T} \setminus E$ and for each n let $\gamma_n(\theta)$ be the circular arc in $\mathbb{D}$ with the same endpoints as I_n and which makes angle θ with I_n (so $\gamma_n(0) = I_n$ and $\gamma_n(\pi/2)$ is the hyperbolic geodesic with the same endpoints as I_n). Let $C_n(\theta)$ be the region bounded by I_n and $\gamma_n(\theta)$, and let $W_E(\theta) = \mathbb{D} \setminus \cup_n C_n(\theta)$.

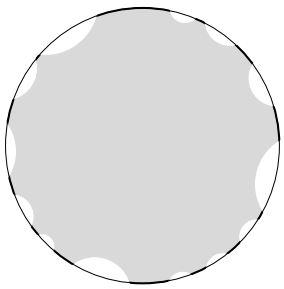


FIGURE 1. The sawtooth domain W_E

For the rest of the paper we will let $W_E = W_E(\pi/8)$ (and let $W_E^* \subset \mathbb{D}^*$ be its reflection across $\mathbb{T}$). We can map $\mathbb{D}$ to W_E by a 2-quasiconformal map f as follows. First let f be the identity on $W_E(\pi/2)$. Then map $U_n = C_n(\pi/2) \setminus C_n(\pi/4)$ (which is

a crescent of angle $\pi/4$) to $V_n = C_n(\pi/2) \setminus C_n(3\pi/8)$ (which is a crescent of angle $\pi/8$) as follows: map U_n to the cone $\{z : 0 < \arg(z) < \pi/4\}$ by a Möbius transformation, then to $\{z : 0 < \arg(z) < \pi/8\}$ by halving the angle and then to V_n by another Möbius transformation. Finally, map $C_n(\pi/4)$ to $C_n(3\pi/8) \setminus C_n(\pi/8)$ by a Möbius transformation. See Figure 2. It is easy to check these maps can be chosen to match up along the common boundaries and hence define a 2-quasiconformal map.

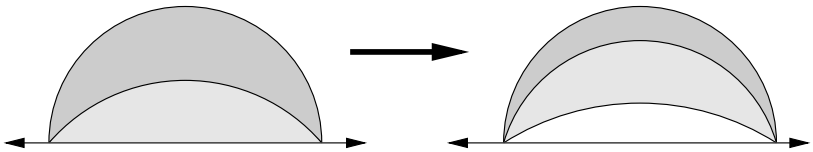


FIGURE 2. Mapping the disk to W_E

If $f : \mathbb{D} \to \Omega$ and $0 < r < 1$, then define

$$d_f(r) = \sup\{|f(z) - f(w)| : |z| = |w| = r \text{ and } |z - w| \le 1 - r\}.$$

If $\partial\Omega$ is bounded in the plane, then it is easy to see this goes to zero as $r \nearrow 1$, since otherwise any neighborhood of $\partial\Omega$ would contain infinitely many disjoint disks of a fixed, positive size.

LEMMA 62. *Suppose $f : \mathbb{D} \to \Omega \subset S^2$ is conformal. Then for any $\epsilon > 0$ there is a compact set $E \subset \mathbb{T}$ with $\operatorname{cap}(\mathbb{T} \setminus E) < \epsilon$ such that f is continuous on $\overline{W_E}$.*

PROOF. By applying a square root and a Möbius transformation, we may assume that $\partial\Omega$ is bounded in the plane. Given $r < 1$ let

$$E(\epsilon, r) = \{x \in \mathbb{T} : |f(sx) - f(tx)| > \epsilon \text{ for some } r < s < t < 1\}$$

and note that by Lemmas 57 and 59

$$\widetilde{\operatorname{cap}}(E(\epsilon, r)) \le \exp(-\pi\epsilon^2/Ca(r)).$$

Moreover, this set is open since f is continuous at the points sx and tx. So if we take $\epsilon_n = 2^{-n}$, and use the relationship between cap and $\widetilde{\operatorname{cap}}$ we can choose r_n so close to 1 that $\operatorname{cap}(E_n) \equiv \operatorname{cap}(E(\epsilon_n, r_n)) \le \epsilon 2^{-n}$. If we define $E = \mathbb{T} \setminus \cup_{n>1} E_n$, then E is closed and $\mathbb{T} \setminus E$ has capacity $\le \epsilon$ by subadditivity .

To show f is continuous at every $x \in \overline{W_E}$, we want to show that $|x - y|$ small implies $|f(x) - f(y)|$ is small. We only have to consider points $x \in \partial W_E \cap \mathbb{T}$. First

suppose $y \in \partial W_E \cap \mathbb{T}$. Choose the maximal n so that $s = |x - y| \le 1 - r_n$. Then $x, y \notin E_n$, so

$$|f(x) - f(y)| \le |f(x) - f(sx)| + |f(sx) - f(sy)| + |f(sy) - f(y)|.$$

The first and last terms on the right are $\le \epsilon_{n-1}$ by the definition of E. The middle term is at most $d_f(1 - s)$ (which tends to 0 as $s \to 0$). Thus $|f(x) - f(y)|$ is small if $|x - y|$ is.

Now suppose $x \in \partial W_E \cap \mathbb{T}$, $y \in \partial W_E \setminus \mathbb{T}$. From the definition of W_E it is easy to see there is a point $w \in \partial W_E \cap \mathbb{T}$ such that $|w - y| \le 2(1 - |y|) \le 2|x - y|$. For the point w we know by the argument above that $|f(x) - f(w)|$ is small. On the other hand, if $t = 1 - |y|$, then

$$|f(y) - f(w)| \le |f(y) - f(tw)| + |f(tw) - f(w)|.$$

The first term is bounded by $Cd_f(1 - t)$ and the second is small since $w \notin E_n$. Thus $|f(x) - f(y)|$ is small depending only on $|x - y|$. Hence f is continuous on $\overline{W_E}$. $\square$

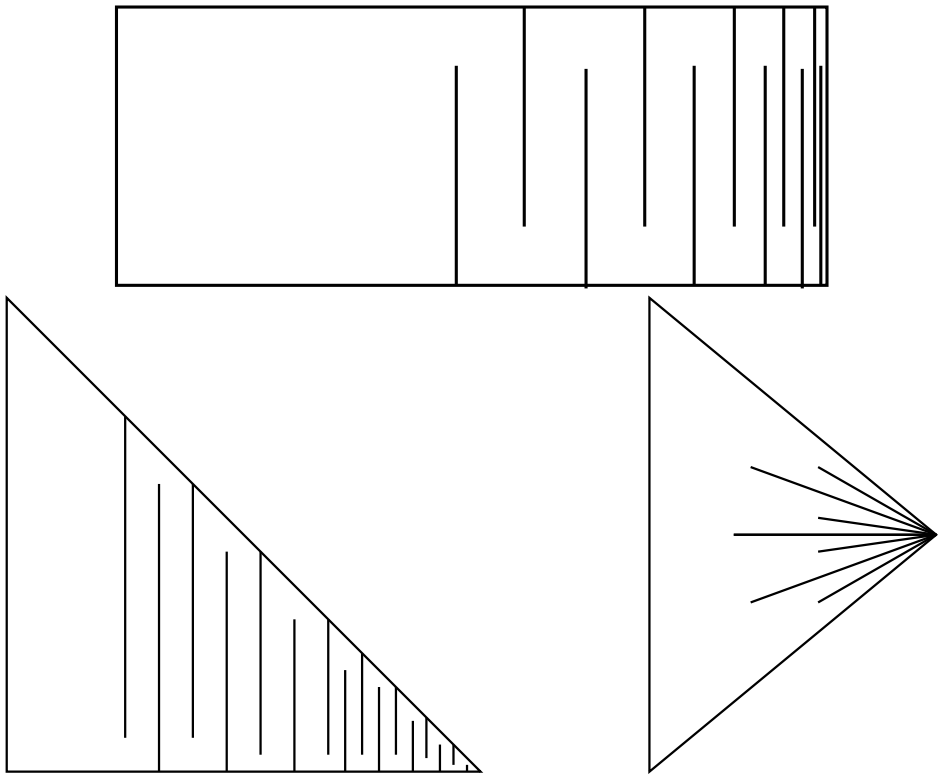


FIGURE 3. On top is a non-locally connected boundary. There is one ray on which the conformal map does not have a limit. The domain on the lower left has a continuous extension everywhere on the boundary, but has a radial image of infinite length. More precisely, suppose the base of the triangle is $[0, 1]$ and the spikes occur at points $\{x_n\}$ and have length $\frac{3}{4}(1 - x_n)$. Then if $\sum_n (1 - x_n) = \infty$, the hyperbolic geodesic that ends at the lower right corner has infinite length. The figure on the right also represents a locally connected boundary, but it has a boundary point whose preimage is an uncountable Cantor set on the circle. by Theorem **??**, this gives an example of a Cantor set with capacity zero (one can also be constructed directly).

5. Symmetry and Modulus

If γ is a path in the plane let $\bar{\gamma}$ be its refection across the real line and let $\gamma^+ = (\gamma \cap \mathbb{H}) \cup \overline{\gamma \cap \mathbb{L}}$, where $\mathbb{H}, \mathbb{L}$ denote the upper and lower halfplanes. If Γ is a path family in the plane then $\overline{\Gamma} = \{\bar{\gamma} : \gamma \in \Gamma\}$ and $\Gamma^+ = \{\gamma^+ : \gamma \in \Gamma\}$.

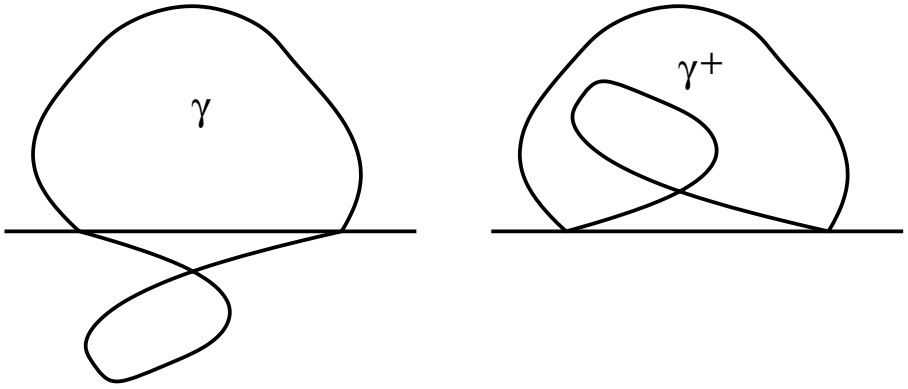


FIGURE 4. The curves γ and γ^+

LEMMA 63. *If* $\Gamma = \overline{\Gamma}$ *then* $M(\Gamma) = 2M(\Gamma^+)$.

PROOF. We start by proving $M(\Gamma) \leq 2M(\Gamma^+)$. Given a metric ρ, define $\sigma(z) = \max(\rho(z), \rho(\bar{z}))$. Then for any $\gamma \in \Gamma$,

$$\int +\gamma^+ \sigma ds \geq \int_{\gamma^+} \rho ds \geq \inf_{\gamma \in \Gamma} \int_{\gamma} \rho ds.$$

Thus if ρ admissible for Γ^+, then σ is admissible for Γ Thus, since $\max(a,b)^2 \leq a^2 + b^2$,

$$M(\Gamma) \leq \int \sigma^2 dxdy \leq \int \rho^2(z) dxdy + \int \rho^2(\bar{z}) dxdy \leq 2 \int \rho^2(z) dxdy.$$

Taking the infimum over admissible ρ's for Γ^+ makes the right hand side equal to $2M(\Gamma^+)$, proving the claim.

For the other direction, given ρ define $\sigma(z) = \rho(z) + \rho(\bar{z})$ for $z \in \mathbb{H}$ and $\sigma = 0$ if $z \in lhp$. Then

$$\begin{aligned}
\int_{\gamma^+} \sigma ds &= \int_{\gamma^+} \rho(z) + \rho(\bar{z}) ds \\
&= \int_{\gamma \cap \mathbb{H}} \rho(z) ds + \int_{\gamma \cap \mathbb{H}} \rho(\bar{z}) ds + \int_{\gamma \cap \mathbb{L}} \rho(z) + \int_{\gamma \cap \mathbb{L}} \rho(\bar{z}) ds \\
&= \int_{\gamma} \rho(z) ds + \int_{\bar{\gamma}} \rho(z) ds \\
&\geq 2 \inf_r ho \int_{\gamma} \rho ds.
\end{aligned}$$

Thus if ρ is admissible for Γ, $\frac{1}{2}\sigma$ is admissible for Γ^+. Hence, since $(a+b)^2 \leq 2(a^2+b^2)$,

$$\begin{aligned}
M(\Gamma^+) &\leq \int (\frac{1}{2}\sigma)^2 dxdy \\
&= \frac{1}{4} \int_{\mathbb{H}} (\rho(z) + \rho(\bar{z}))^2 dxdy \\
&\leq \tfrac{1}{2} \int_{\mathbb{H}} \rho^2(z) dxdy + \int_{\mathbb{H}} \rho^2(\bar{z}) dxdy \\
&= \frac{1}{2} \int \rho^2 dxdy.
\end{aligned}$$

Taking the infimum over all admissible ρ's for Γ gives $\frac{1}{2}M(\Gamma)$ on the right hand side, proving the lemma. □

LEMMA 64. *Let $\mathbb{D}^* = \{z : |z| > 1\}$ and $\Omega_0 = \mathbb{D}^* \setminus [R, \infty)$ for some $R > 1$. Let $\Omega = \mathbb{D}^* \setminus K$, where K is a closed, unbounded, connected set in $\mathbb{D}^*$ which contains the point $\{R\}$. Let Γ_0, Γ denote the path families in these domains with separate the two boundary components. Then $M(\Gamma_0) \leq M(\Gamma)$.*

PROOF. We use the symmetry principle we just proved. The family Γ_0 is clearly summytric (i.e., $\Gamma = \overline{\Gamma}$, so $M(\Gamma^+) = \frac{1}{2}M(\Gamma_0)$. The family Γ may not be symmetric, but we can replace it by a larger family that is. Let Γ_R be the collection of rectifiable curves in $\mathbb{D}^* \setminus \{R\}$ which have zero winding number around $\{R\}$, but non-zero winding number around 0. Clearly $\Gamma \subset \Gamma_R$ and Γ_R is symmetric so $M(\Gamma) \geq M(\Gamma_R) = 2M(\Gamma_R^+)$. Thus all we have to do is show $M(\Gamma_R^+) = M(\Gamma_0^+)$. We will actually show $\Gamma_R^+ = \Gamma_0^+$. Since $\Gamma_0 \subset \Gamma_R$ is obvious, we need only show $\Gamma_R^+ \subset \Gamma_0^+$.

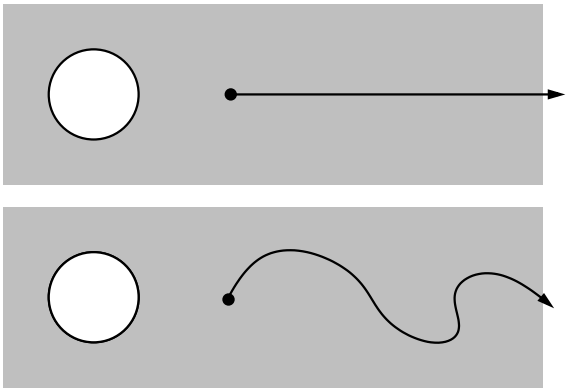


FIGURE 5. The annulus on top has smaller modulus than any other annulus formed by connecting R to ∞.

Suppose $\gamma \in \Gamma_R$. Since γ has non-zero winding around 0 it must cross both the negative and positive real axes. If it never crossed $(0, R)$ then the winding around 0 and R would be the same, which false, so γ must cross $(0, R)$ as well. Choose points $z_- \in \gamma \cap (-\infty, 0)$ and $z_+ \in \gamma \cap (0, R)$. These points divide γ into two subarcs γ_1 and γ_2. Then $\gamma^+ = \gamma_1^+ \cup \gamma_2^+$. But if we reflect γ_2^+ into the lower half-plane and join it to γ_1^+ it forms a closed curve γ_0 that is in Γ_0 and $\gamma_0^+ = \gamma^+$. Thus $\gamma^+ \in \Gamma_0^+$, as desired. □

Let $\Omega_{\epsilon,R} = \{z : |z| > \epsilon\} \setminus [R, \infty)$. Thus $\Omega_{1,R}$ is the domain considered in the previous lemma. We can estimate the moduli of these domains using the Koebe map

$$k(z) = \frac{z}{(1+z)^2} = z - 2z^2 + 3z^3 - 4z^4 + 5z^5 - \dots,$$

which conformal maps the unit disk to $\mathbb{R}^2 \setminus [\frac{1}{4}, \infty)$ and satisfies $k(0) = 0$, $k'(0) = 1$. Then $k^{-1}(\frac{1}{4R}z)$ maps $\Omega_{\epsilon,R}$ conformally to an annular domain in the disk whose outer boundary is the unit circle and whose inner boundary is trapped between the circle of radius $\frac{\epsilon}{4R}(1 \pm O(\frac{\epsilon}{R}))$. Thus the modulus of $\Omega_{\epsilon,R}$ is $2\pi \log \frac{4R}{\epsilon} + O(\frac{\epsilon}{R})$.

LEMMA 65. *Suppose $z, w \in \mathbb{D}$ and K is a compact connected set in $\mathbb{D}$ which contains both these points. Let Γ be the path family that separates K and $\mathbb{T}$. Then the modulus of this family is maximized when K is the hyperbolic geodesic between z and w in which case the modulus is $2\pi \log \frac{4}{\rho}(z, w) + O(\rho(z, w))$, where ρ denotes the hyperbolic distance.*

PROOF. By conformal invarience we may use a Möbius transformation to move z to 0 and w onto the positive axis. Applying an inversion, the path family is mapped to one as in Lemma 64, showing that the radial line from z to w maximizes the modulus. The estimate of the modulus follows from our previous remarks. □

THEOREM 66 (The Koebe $\frac{1}{4}$ Theorem). *Suppose f is holomorphic, 1-1 on $\mathbb{D}$ and $f(0) = 0$, $f'(0) = 1$. Then $D(0, \frac{1}{4}) \subset f(\mathbb{D})$.*

PROOF. This proof is from [**?**]. Recall that the modulus of a doubly connected domain is the modulus of the path family that separates the two boundary components (and is equal to the extremal distance between the boundary components). Let $R = \text{dist}(0, \partial f(\mathbb{D}))$. Let $A_{\epsilon,r} = \{z : \epsilon < |z| < r\}$ and note that by conformal invarience

$$2\pi \log \frac{1}{\epsilon} = M(A_{\epsilon,1}) = M(f(A_{\epsilon,1}).$$

Let $\delta = \min_{|z|=\epsilon} |f(z)|$. Since $f'(0) = 1$, $\delta = \epsilon + O(\epsilon^2)$. Note that $f(\mathbb{D}) \setminus D(0, delta) \supset f(A_{\epsilon,1})$, so

$$M(f(\mathbb{D}) \setminus D(0, \epsilon_2)) \geq M(f(A_{\epsilon,1})).$$

By Lemma 64

$$M(f(\mathbb{D}) \setminus D(0, \epsilon_2)) \leq M(\Omega_{\epsilon_2,R}) = 2\pi \log \frac{4R}{\epsilon_{\ 2}} + O(\frac{\epsilon_2}{R}).$$

Putting these together gives

$$2\pi \log \frac{4R}{\delta} + O(\frac{\delta}{R}) \geq 2\pi \log \frac{1}{\epsilon}.,$$

or

$$\log 4R - \log(\epsilon + O(\epsilon^2)) + O(\frac{\epsilon}{R}) \geq -\log \epsilon.$$

Taking $\epsilon \to 0$ shows $\log 4R \geq 0$, or $R \geq \frac{1}{4}$. □

COROLLARY 67. *If f is univalent on $\mathbb{D}$, then*

$$\frac{1}{4}|f'(z)|(1 - |z|^2) \leq \text{dist}(f(z), \partial\Omega) \leq |f'(z)|(1 - |z|^2).$$

PROOF. By precomposing with a Möbius transformation and postcomposing by a linear map, we may assume $z = 0$, $f(0) = 0$ and $f'(0) = 1$. Then the right hand inequality is just Schwarz's lemma applied to f^{-1}. The right side is the previous result. □

Because of Koebe's theorem we have

$$d\rho_\Omega \leq d\tilde{\rho}_\Omega \leq 4d\rho_\Omega. \tag{16}$$

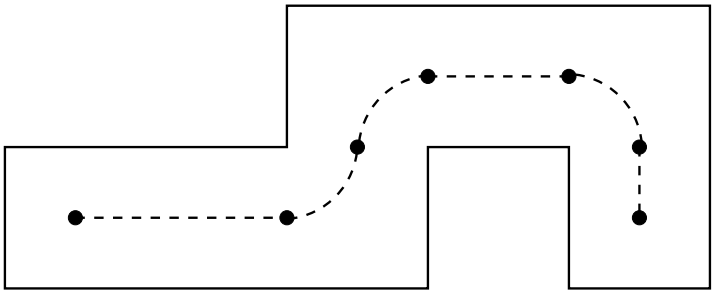


FIGURE 6. Its easy to compute the quasihyperbolic length of this curve (which is $3 + \frac{3}{2}\pi$), and a little more involved to show it is a quasi-hyperbolic geodesic, but together these facts give an estimate of its hyperbolic length.

COROLLARY 68. *Suppose Ω is simply connected, $z, w \in \Omega$. Then*

$$\rho(z, w) \geq |\log \frac{\operatorname{dist}(z, \partial\Omega)}{\operatorname{dist}(w, \partial\Omega)}|.$$

PROOF. Suppose γ is a curve in Ω connecting the two points. Then the quasi-hyperbolic length of γ is at least

$$|\int_{\operatorname{dist}(z,\partial\Omega)}^{\operatorname{dist}(w,\partial\Omega} \frac{dt}{t}| = |\log \frac{\operatorname{dist}(z, \partial\Omega}{\operatorname{dist}(w, \partial\Omega)}|.$$

By our previous remarks, the hyperbolic distance is at least $\frac{1}{4}$ of this. □

6. The distortion theorems

In this section we give the "usual" proof of Koebe's $\frac{1}{4}$ theorem, via the area theorem and deduce the sharp version of the distortion estimates.

Recall Green's theorem,

$$\iint_\Omega u\Delta v + v\Delta u dxdy = \int_{\partial\Omega} u\frac{\partial v}{\partial n} + v\frac{\partial u}{\partial n} ds, \tag{17}$$

where n denotes the inward pointing normal vector of $\partial\Omega$.

COMPLEXVERSION

We will also use Green's theorem in the following form:

$$\int_{\partial\Omega} f(x,y)dx + g(x,y)dy = \iint_{\Omega} \frac{\partial g}{\partial x} - \frac{\partial f}{\partial g}dxdy \tag{18}$$

and its simple consequence that the area of a region Ω is given by

$$\text{area}(\Omega) = \frac{1}{2}\int_{\partial\Omega} xdy - ydx = \frac{1}{2i}\int \partial\Omega \bar{z}dz. \tag{19}$$

We now come to some well known (but perhaps not as well known as the results above) estimates for univalent mappings. The basic idea is to show that a univalent map f on $\mathbb{D}$ is well approximated by its linear Taylor approximation $f(z_0)+f'(z_0)(z-z_0)$ in a hyperbolic neighborhood of z_0, with estimates that do not depend on f or z. These so called "distortion estimates" are fundamental to most arguments in geometric function theory. The first step is to prove:

THEOREM 69 (Area theorem). *Suppose $g(z) = \frac{1}{z} + b_0 + b_1 z + \dots$ is univalent in $\mathbb{D}$. Then $\sum_{n=0}^{\infty} n|b_n|^2 \leq 1$. In particular, $|b_1| \leq 1$.*

PROOF. For $0 < r < 1$ let $D_r = \mathbb{C} \setminus g(D(0,r))$. If $z = g(w)$ and $w = e^{i\theta}$ then $dw = iwd\theta$, so by (58),

$$\text{area}(D_r) = \iint_{D_r} dxdy = \frac{1}{2i}\int_{\partial D_r} \bar{z}dz = \frac{-1}{2i}\int_{\partial D(0,r)} \bar{g}(w)g'(w)dw.$$

To evaluate the right hand side note that

$$\begin{aligned} g(z) &= \frac{1}{z} + b_0 + b_1 z + \dots, \\ g'(z) &= 1\frac{1}{z^2} + 0 + b_1 + 2b_2 z + \dots, \end{aligned}$$

so that

$$\begin{aligned} \int_{|w|=r} \bar{g}(w)g'(w)dw &= i\int \bar{g}(w)g'(w)wd\theta \\ &= i\int (\frac{1}{\bar{w}} + \bar{b}_0 + \bar{b}_1\bar{w} + \dots)(-\frac{1}{w} + b_1 w + 2b_2 w + \dots)d\theta \\ &= 2\pi i(-\frac{1}{r^2} + |b_1|^2 r^2 + 2|b_2|r^4 + \dots 0 \end{aligned}$$

Thus,

$$0 \leq \text{area}(D_r) = \pi(\frac{1}{r^2} - \sum_{n=1}^{\infty} n|b_n|^2 r^{2n}).$$

Taking $r \to 1$ gives the result. □

COROLLARY 70. *If $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ is univalent on the unit disk, then $|a_2| \leq 2$.*

PROOF. Let $g(z) = (f(z^2))^{-1/2} = 1/z - a_2 z/2 + \dots$. We claim g is one-to-one. To see this suppose $g(z) = g(w)$. Then $f(z^2) = f(w^2)$, so $z = \pm w$. Note that g is odd, so $z = w$. Since $b_1 = a_2/2$, the previous result implies $|a_2| \leq 2$. □

THEOREM 71 (Koebe 1/4 theorem). *If f is univalent on $\mathbb{D}$, then*

$$\frac{1}{4}|f'(z)|(1-|z|^2) \leq \operatorname{dist}(f(z), \partial\Omega) \leq |f'(z)|(1-|z|^2).$$

PROOF. By precomposing with a Möbius transformation and postcomposing by a linear map, we may assume $z = 0$, $f(0) = 0$ and $f'(0) = 1$. Then the right hand inequality is just Schwarz's lemma applied to f^{-1}. To prove the left hand inequality, suppose f never equals w in $\mathbb{D}$. Then

$$g(z) = \frac{wf(z)}{w - f(z)} = z + (a_2 + \frac{1}{w})z^2 + \dots,$$

is univalent with $f(0) = 0$ and $f'(0) = 1$. Applying Corollary 128 to both f and g gives

$$\frac{1}{|w|} \leq |a_2| + |a_2 + \frac{1}{w}| \leq 2 + 2 = 4.$$

Thus the omitted point w lies outside $D(0, 1/4)$, as desired. □

Because of Koebe's theorem we have

$$d\rho_\Omega \leq d\tilde{\rho}_\Omega \leq 4 d\rho_\Omega. \tag{20}$$

Here $\tilde{\rho}$ is the quasi-hyperboic metric on Ω, given by $|dz|/\operatorname{dist}(z, \partial\Omega)$.

LEMMA 72. *Suppose f is univalent on $\mathbb{D}$, $f(0) = 0$ and $f'(0) = 1$. Then*

$$\frac{1-|z|}{(1+|z|)^3} \leq |f'(z)| \leq \frac{1+|z|}{(1-|z|)^3},$$

PROOF. Fix a point $w \in \mathbb{D}$ and write the Koebe transform of f,

$$F(z) = \frac{f(\tau(z)) - f(w)}{(1-|w|^2)f'(w)},$$

where

$$\tau(z) = \frac{z+w}{1-\bar{w}z}.$$

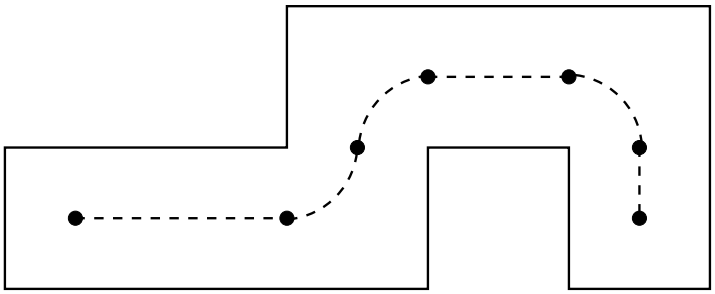


FIGURE 7. Its easy to compute the quasihyperbolic length of this curve (which is $3 + \frac{3}{2}\pi$), and a little more involved to show it is a quasi-hyperbolic geodesic, but together these facts give an estimate of its hyperbolic length.

This is univalent, so by Corollary 128, $|a_2(w)| \leq 2$. Differentiation and setting $z = 0$ shows

$$F'(z) = \frac{f'(\tau(z))\tau'(z)}{(1-|w|^2)f'(w)},$$

$$F''(z) = \frac{f''(\tau(z))\tau'(z)^2 + f'(\tau(z))\tau''(z)}{(1-|w|^2)f'(w)},$$

$$\tau'(0) = 1-|w|^2, \tau''(0) = -2(1-|w|^2),$$

$$F''(0) = \frac{f''(w)}{f(w)}(1-|w|^2) - 2\bar{w}.$$

This implies that the coefficient of z^2 (as a function of w) in the power series of F is

$$a_2(w) = \frac{1}{2}((1-|w|^2)\frac{f''(w)}{f'(w)} - 2\bar{w}).$$

Using $|a_2| \leq 2$ and multiplying by $w/(1-|w|^2)$, we get

$$|\frac{wf''(w)}{f'(w)} - \frac{2|w|^2}{1-|w|^2}| \leq \frac{4|w|}{1-|w|^2}.$$

Thus

$$\frac{2|w|^2 - 4|w|}{1-|w|^2} \leq \frac{wf''(w)}{f'(w)} \leq \frac{4|w| + 2|w|^2}{1-|w|^2}.$$

Now divide by $|w|$ and use partial fractions,

$$\frac{-1}{1-|w|} + \frac{-3}{1+|w|} \leq \frac{1}{|w|}\frac{wf''(w)}{f'(w)} \leq \frac{3}{1-|w|} + \frac{1}{1+|w|}$$

Note that

$$\begin{aligned}\frac{\partial}{\partial r}\log|f'(re^{i\theta})| &= \frac{\partial}{\partial r}\mathrm{Re}\log f'(z)\\ &= \mathrm{Re}\frac{z}{|z|}\frac{\partial}{\partial z}\log f'(z)\\ &= \frac{1}{|z|}\mathrm{Re}(\frac{zf''(z)}{f'(z)})\end{aligned}$$

Since $w = re^{i\theta}$ and $f'(0) = 1$, we can integrate to get

$$\log(1-r) - 3\log(1+r) \le \log|f'(re^{i\theta})| \le -3\log(1-r) + \log(1+r).$$

Exponentiating gives the result. □

7. The Voronoi diagram

Given a collection of sets $\mathcal{S} = \{S_k\} \subset \mathbb{C}$, called "sites", we let

$$\Omega_k = \{z \in \mathbb{C} : \mathrm{dist}(z, S_k) < \mathrm{dist}(z, S_j) \text{ for all } j \neq k\},$$

i.e., Ω_k is the collection of points which are strictly closer to S_k than to any other site. For example, if $\int$ is a finite set of distinct points, then each the $\{Omega_k\}$ are disjoint polygonal regions (possibly unbounded) whose closures fill the entire plane.

We are more interested in the case when the sides are the sides of a polygon and we only consider the part of Ω that lies inside the polygon (which we will also call Ω_k). In this case, the regions Ω_k do not "fill in" the whole interior. At each concave vertex (i.e., interior angle $> \pi$ there is a non-trivial polygonal region which are equal distant from the vertex and its two adjacent edges; thus these points are not in Ω_k for either of these edges. Such points form the set Ω_v, which we think of as the Voronoi cell associated to the vertex v. With this proviso, the Voronoi decomposition of a polygon is a division of the interior on the polygon into polygonal pieces, one for each edge and concave vertex. The boundaries of these cells for a finite graph with edges that terminate at each vertex of the polygon (one at each convex vertex and two at each concave vertex). See Figure 5.

Inside each Voronoi cell, the quasi-hyperbolic metric is easy to describe. For edge-cells, it is simply the hyperbolic metric of for the half-plane containing the cell and whose boundary line contains the corresponding edge of the polygon. For vertex

cells, it the quasi-hyperbolic metric is given by $\int_\gamma \frac{|dz|}{|z-v|}$ where $v$is the vertex. This cell has two sides which meet at v. Using the logarithmic map, we can conformally map the cell to a subset of a horizontal strip so that v maps to $-\infty$. Since $(\log z)' = \frac{1}{z}$, the quasi-hyperbolic metric in the cell is mapped isometrically the Euclidean metric in the strip.

Thus the Voronoi cells break Ω up into a finite number of pieces where it has a simply formula. On edge piece the metric is hyperbolic; on vertex pieces it is a Euclidean metric (after the log mappings).

Recall that in Davis' method we tried to iteratively improve a guessed set of z-parameters by computing the corresponding image polygon and increasing or decreasing the distance between adjacent parameters based on whether the Euclidean length of the corresponding edge (as a fraction of the total length of the polygon) was more or less than in the target polygon. The problem with this method was that we could not justify the apparent convergence in many cases because Euclidean length and harmonic measure of a boundary arc and not always monotonely related.

The quasi-hyperbolic metric allows us to compute another measure of a boundary arc that might be more directly related to the arcs harmonic measure. Choose a point z_0 on the boundary of some Voronoi cell. Each edge I of the polygon has its corresponding cell. The boundary of this cell either contains z_0 or contains a point z_I which can be connected to z_0 by a shortest path along the cell boundaries. In the first case we let ℓ_I be the interior angle of the cell Ω_I at z_0. In the second case we let ℓ_I be the product of the interior angle of Ω_I at z_I and $\exp(-\tilde{\rho}(z_0, z_I))$.

This suggests a quasi-hyperbolic version of Davis's method:

(1) Given a set of z-parameter guesses, compute the image polygon.
(2) compute the Voronoi diagram and compute the numbers ℓ_I for the image. Normalize them to have total measure 2π.
(3) Compare with the corresponding list of numbers $\hat{\ell}_I$ for the target polygon. Change the spacing between adjacent parameters by multiplying by ℓ_I/ℓ_I. Renormalize so the spacings so they sum to 2π.
(4) Repeat until the desired accuracy is attained.

I have not written code to implement this, but expect that it would work as well as Davis's method and possibly better, but is more costly to compute. Is it worth the extra work?

FIGURE of

8. Convergence of Kakutani's method

Now that we have Koebe's $\frac{1}{4}$ theorem, we can return to some unfinshed business from Chapter 5. We introduced a random process which starts at a point $z_0 \in \Omega$ and selects a new point on the circle, uniformly and randomly on the circle centered at z_0 of radius $\lambda \text{dist}(z_0, \partial\Omega)$. Repeating this process gives a random sequence of points $\{z_n\}$ which we had claimed converges exponentially to the booundary almost surely. We can now prove this.

LEMMA 73. *Assume Ω is simply connected. With notation as above, the sequence $\{\mathbf{z}_n\}$ converges to a point $z \in \partial\Omega$ almost surely. Let $d_n = \text{dist}(z_n, \partial\Omega)$. Then there are constants $0 < a, b < 1$ (independent of Ω and z_0) so that*

$$\text{Prob}(d_n \geq a^n d_0) \leq b^n.$$

In particular, after $O(\log \frac{1}{\epsilon})$ steps of the random walk, there is only an exponentially small chance that the walk is farther than ϵ from its limiting point on $\partial\Omega$. The strong law of large numbers can be applied to show that almost surely, $d_n = O(a_n)$ for large enough n.

PROOF. There is no loss in generality to assume d_0, so we do this. Let $f : \mathbb{D} \to \Omega$ be a Riemann mapping sending 0 to z_0 and define

$$v(z) = \log(|f'(z)|^{-1}) + \rho_{\mathbb{D}}(0, z) = -\log|f(z)| + \log \rho_{\mathbb{D}}(0, z).$$

since $\log \rho(0, z) \simeq 1 - |z|$ for $|z| > 1/2$, v has a negative logarithmic pole at 0 and satisifies

$$v(z) \simeq \log \frac{1}{\text{dist}(f(z), \partial\Omega)},$$

near the boundary.

We claim that v is subharmonic is a uniform way.

LEMMA 74. *If $\mathbf{z} \in \mathbb{D}$ and $W \subset \mathbb{D}$ is a subdomain which contains an ϵ-hypeerbolic ball around z, then*

$$\int_{\partial W} v(w) d\omega(z, \cdot, W) \geq v(z) + c,$$

for some $c > 0$ that depends on ϵ but not on z.

Assume this for the momment and we will finish the proof of the Lemma 73 . Suppose z_n is known. We choose z_{n+1} from a Euclidean circle of radius λd_n around z_n. By Corollary 67 this circle contains a hyperbolic ball of radius $\epsilon > 0$ around z_n where ϵ is fixed depending only on λ. The conformal preimate of this disk is a smooth domain in the disk containing $w_k = f^{-1}(z_k)$ to which Lemma 5 applies. Thus expected value of V over all choices of z_{n+1} is an additive constant larger than the expected value over z_n. Thus $\mathbf{E}(V, n) \geq cn$. After the n steps, the Kakutani process can be no closer than $(1 - \lambda)^n$ to the boundary, so the expected value of v is $\leq Cn$. □

CHAPTER 5

Quasiconformal Mappings

In this chapter we consider quasiconformal maps in more detail. In particular, we will introduce several classes of self-maps of the disk (biLipschitz, quasiconformal, quasi-isometric and quasisymmteric) and discuss the containments between these classes.

1. Compactness of K-quasiconformal maps

We want to show that the collection of K quasiconformal maps $\mathbb{D} \to \mathbb{D}$ which fix the origin is compact. This has two steps. First we have to show this collection is equicontinuous so that we can apply the Arcela-Ascoli theorem. This will follow from the symmetry principle of the previous section. Next, we have to show that a uniform limit of K-quasiconformal maps is K-quasiconformal. Since we are using the geometric definition of quasiconformality, this reduces to showing that if a sequence of generalized qudrilaterals Q_n converges to a quadrilateral Q, then the moduli of Q_n converge to the modulus of Q. We start with the equicontinuity.

LEMMA 75. *Suppose $f : \mathbb{D} \to \mathbb{D}$ is K-quasiconformal. Then*

$$\frac{1}{C}\rho(z,w)^K, \rho(z,w) \leq \rho(f(z), f(w)) \leq C \min(\rho(z,w)^{1/K}, \rho(z,w)).$$

PROOF. For the upper bound on small scales use modulus. On large scales connect by unit chain. For lower bound, use upper bound on inverse function. □

Recall that a generalized quadrilateral is a Jordan domain with four ordered, distinct points on the boundary. We say that two such domains are ϵ close if each of their boundaries lie in an ϵ neighborhood of the other and corresponding vertices lie within ϵ of each other

LEMMA 76. *Suppse R is a $1 \times r$ rectangle and $\Omega \subset R$ is ϵ close to R with $\epsilon < \frac{1}{4}r$. Then*

$$1 - O(\epsilon) \leq \frac{M(R)}{M(\Omega)} \leq 1 + O(\epsilon).$$

PROOF. Consider the metric $\frac{1}{r-2\epsilon}$. This is clearly admissible for the path family connecting the "vertical" sides of Ω and $\int \rho^2 dxdy \leq \frac{r}{(r-2\epsilon)^2} \leq \frac{1}{r}(1+O(\frac{\epsilon}{r}))$. The same esimate for the conjugate family proves the oppsite inequality. □

LEMMA 77. *Suppose Ω is a generalized quadrilateral. For any $\epsilon > 0$ there is a another quadrilateral Q' containing Q and a conformal mapping f of Q' to a rectangle R, so that $f(\Omega)$ is ϵ close to R.*

PROOF. Choose a basepoint in Ω and a domain Ω' so that the quasi-hyperbolic distance from the base point to $\partial\Omega$ in Ω' is large. Then the distnace in R will also be large which implies the boundary of $f(\Omega)$ is close to ∂R. □

LEMMA 78. *Suppose Ω is a generalized quadrilateral and $\{Q_n\}$ are quadrilaterals so that for any $\epsilon > 0$, eventually all the $Q_n's$ are ϵ-close to Ω. Then $M(Q_n) \to M(Q)$.*

PROOF. Fix $\epsilon > 0$ and choose Ω' containing Ω and a conformal map f of Ω' to a rectangle so that $f(\Omega)$ is ϵ close to R. Then for n large enough $f(\Omega_n)$ is ϵ close to $f(\Omega)$ and hence 2ϵ close to R. Thus both $M(\Omega) = M(f(\Omega))$ and $M(\Omega_n) = M(f(\Omega_n))$ are within a factor of $1 + O(\epsilon)$ of $M(\Omega)$ and hence are this close to each other. since this holds for any ϵ, we have the desired convergence. □

LEMMA 79. *If $\{f_n\}$ are K-quasiconformal maps that converge uniformly on compact sets to a homeomorphism f, then f is K-quasiconformal.*

PROOF. For any generalized quadrilateral Q, $f_n(Q)$ is eventually ϵ close to $f(Q)$, whichis a generalized quadrialateral since f is a homeomorphism. The by the previous Lemma $M(f(Q)) = \lim_n M(f_n(Q)) \leq KM(Q)$. Since this holds for all quadrilaterals, f is K-quasiconformal □

LEMMA 80. *If $\{f_n\}$ are K-quasiconformal maps of $\mathbb{D}$ to itself and $f_n \to f$ uniformly on compact subsets of $\mathbb{D}$, then f is K-quasiconformal.*

PROOF. By Lemma **??**, this collection of functions is equicontinuous and bounded, so by the Arcela-Ascoli theorem, any sequence contains a subsequence which converges uniformly on compact sets to a function f. Since the K-quasiconformal maps are bi-Hölder this map must be 1-1, hence a homeomorphism. By Lemma 79 it is K-quasiconformal. □

LEMMA 81. *If $\{\Omega_n\}$ is an exhaustion of Ω and for each n there is a K-quasiconformal map $\varphi_n : \Omega_n \to \mathbb{D}$, which maps some fixed point z_0 to 0, then we can extract a subsequence that converges to a K-quasiconformal map $\varphi : \Omega \to \mathbb{D}$.*

LEMMA 82. *Suppose Ω is a Jordan domain and for each $z \in \Omega$ let I_z be an arc of $\partial\Omega$ with*

$$\operatorname{diam}(I_z) \simeq \operatorname{dist}(z, I_z) \simeq \operatorname{dist}(z, \partial\Omega).$$

Suppose $\varphi : \Omega \to \mathbb{D}$ is a quasiconformal map. Then there is another quasiconformal map $\psi : \Omega \to$ disk with the same boundary values so that

$$|\psi'(z)| \leq C\operatorname{diam}(\psi(I_z))/\operatorname{diam}(I_z)).$$

2. Quasi-isometries

Defn QI

A map is quasi-isometry if image and inverse image of any 1-ball is contained in a C-ball.

LEMMA 83. *A K-quasiconformal map $\mathbb{D} \to \mathbb{D}$ is a quasi-isometry of the hyperbolic metric*

PROOF. □

LEMMA 84. *Image of a geodesic ray $[0, e^{i\theta}]$ stays in C-neigborhood of geodesic from $f(0)$ to $f(e^{i\theta})$.*

LEMMA 85. *Image of infinite geodesic stays inside C-neighborhood of geodesic with same endpoints.*

COROLLARY 86. *Boundary values are bi-Hö;der.*

3. Quasisymmetric maps

DEFN

LEMMA 87. *Quasi-isometry extends to quasisymmetric mapping of circle.*

LEMMA 88. *Quasiconformal maps are quasisymmetric.*

LEMMA 89. *The image of a circle $\Gamma \subset \mathbb{D}$ under a quasiconformal map satisfies the bounded turning condition.*

4. BiLipschitz maps

DEFN on metric space

LEMMA 90. *If $f : X \to Y$ is K-biLipschitz and $\gamma \subset X$ is a curve, then*

$$\ell(\gamma)/K \leq \ell(f(\gamma)) \leq K\ell(\gamma).$$

If $E \subset X$ then the 2-dimensional Hausdorff measure satisfies

$$\ell(\gamma)/^2 \leq \ell(f(\gamma)) \leq K^2\ell(\gamma).$$

PROOF. These are almost immediate from definitions. The length of a curve gamma is

$$\ell(\gamma) = \sup \sum_k (x_k, x_{k+1}),$$

where the supremum is over all partitions of γ, i.e., $x_k = \sigma(t_k)$ where $0 \leq t_1 < t_2 < \ldots t_n \leq 1$ and $\sigma : [0,1] \to \gamma$ is a parametrization. If f is biLipschitz on γ, then $f \circ \sigma$ is a parametrization of $f(\gamma)$ and

$$\ell(f(\gamma)) = \sup \sum_k (f(x_k), f(x_{k+1})) \leq K \sup \sum_k (x_k, x_{k+1}) = L\ell(\gamma).$$

Since f is a homeomorphism we can prove the reverse inequality by considering f^{-1}.

The 2-dimensional Hausdorff measure is defined as

$$\mathcal{H}_2(E) = \lim_{\delta \to 0} \inf\{\sum_k r_k^2 : E \subset \cup_k B(x_k, r_k), r_k < \delta\},$$

where the infimum is taken over all covers of E by ball of radius less than δ. If $E \subset \cup_k B(x_k, r_k)$, then $f(E) \subset \cup_k B(f(x_k), Kr_k)$, from which we get $\mathcal{H}_2(f(E)) \leq K^2\mathcal{H}_2(E)$. The opposite inequality is again proved by considering the inverse. □

When $X = \mathbb{R}^2$, the 2-dimensional Hausdorff measure is simply a multiple of the usual area measure. Thus K-biLipschitz maps of planar domains multiply area by at most K^2.

LEMMA 91. *A quasi-symmetric boundary map has a hyperbolic bi-Lipschitz extension to interior.*

PROOF. Tesselate disk by right pentagons. Each pentagon determined by ten points where bounding geodesics hit circle. For each geodesic, map endpoints forward under mapping and form new geodesic. All such cut disk into pentagons which are bounded distortion of originals. Map each pentagon forward. Gives bi-Lipschitz map.

FIGURE OF PENTAGON BEFORE AND AFTER

□

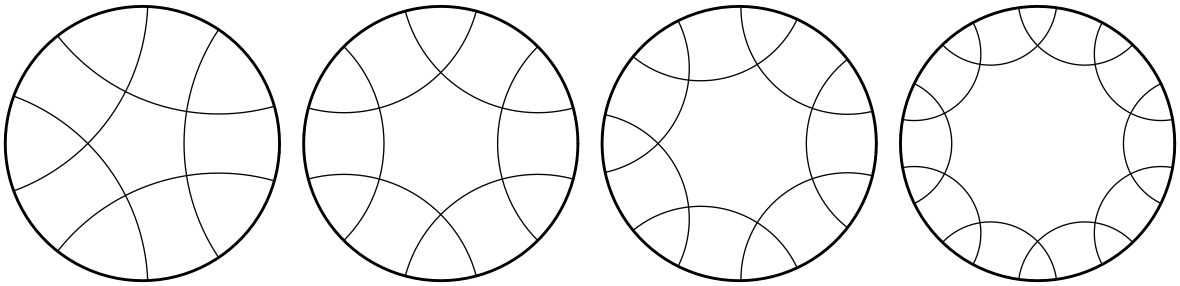


FIGURE 1. The hyperbolic disk can be tesselated by various right polygons

LEMMA 92. *A K-biLipschitz map $f : \mathbb{D} \to \mathbb{D}$ between the hypebolic metrics is K^2-quasiconformal*

PROOF. What we will really uses is the following: if $z \in \mathbb{D}$ and $r > 0$ is small enough (depending on z), then

$$|z - x| < \delta, |z - y| < \delta \Rightarrow \frac{1}{K^2(1 + O(\delta))} \frac{|z - x|}{|z - y|} \frac{|f(z) - f(y)|}{|f(z) - f(x)|} \leq K^2(1 + O(\delta)).$$

In a small hyperbolic ball the ratio on the right is the same wheter measured with Euclidean distances or hyperbolic distances, at least up to factor of $1 + O(\delta)$ which can be absored into the constants. Moreove, if we compose f with a conformal map, the bounds change by at most $1 + O(\delta)$, by the distortion theorem.

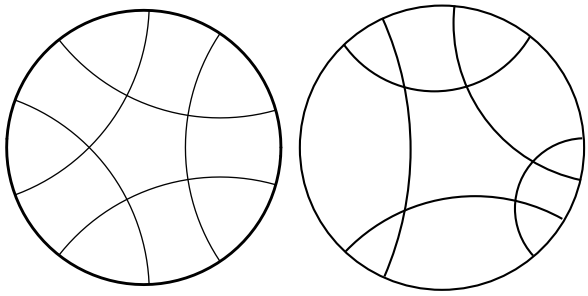


FIGURE 2. A quasisymmetric map can't distort cross ratio by too much and the angle that two geodesics cross at is determined by the cross ratio of the four endpoints on the circle. Thus a QS mapping of the circle induces a tesselation of the disk by distorted pentagons and we can define a biLipschitz extension to the disk by simply mapping each "regular" pentagon to its distorted image by a biLipschitz map. If these maps are affine along the boundaries, then we get a continuous, biLipschitz map of the hyperbolic disk to itself.

The hyperbolic metric on Ω_1 is given by $d\rho_1 = \lambda_1 ds$ for some smooth function $\lambda_1 \simeq \text{dist}(z, \partial\Omega_1)^{-1}$ (namely $\lambda_1 = |f'|/(1-|f|^2)$ where f is a conformal map of Ω_1 to $\mathbb{D}$). This function satisfies an estimate

$$1 - O(\delta) \leq \frac{\lambda_1(w)}{\lambda_1(z)} \leq 1 + O(\delta),$$

if $\rho_1(w,z) \leq \delta < 1$. A similar estimate for Ω_2. Since f is biLipschitz, the distance to the boundary only changes by at most a factor of K under f. Thus

$$\begin{aligned}
\rho_2(f(z), f(w)) &\leq (1 + O(\delta))\lambda_2(z)|f(z) - f(w)| \\
&\leq K(1 + O(\delta))\lambda_2(z)|z - w| \\
&\leq K(1 + O(\delta))\frac{\lambda_2(f(z))}{\lambda_1(z)}\rho_1(w,z) \\
&\leq K^2(1 + O(\delta))\rho_1(x,y).
\end{aligned}$$

□

Quasiconformal maps are a generalization of biLipschitz maps, i.e., maps which satisfy

$$\frac{1}{K} \leq \frac{|f(x) - f(y)|}{|x - y|} \leq K.$$

From the metric definition it is clear that any K-biLipschitz map is K^2-quasiconformal.

For K-quasiconformal self-maps of the disk, there is almost a converse. Although a quasiconformal map $f : \mathbb{D} \to \mathbb{D}$ need not be biLipschitz, it is a quasi-isometry of the disk with its hyperbolic metric ρ, i.e., there are constants A, B such that

$$\frac{1}{A}\rho(x, y) - B \leq \rho(f(x), f(y)) \leq A\rho(x, y) + B.$$

This says f is biLipschitz for the hyperbolic metric at large scales. A quasi-isometry is also called a rough isometry in some sources, e.g., [**?**], [**?**]. We will say f is a quasi-isometry with constant ϵ if we can take $A = 1 + \epsilon$ and $B = \epsilon$.

In [**?**] Epstein, Marden and Markovic show that any K-quasiconformal selfmap of the disk is a quasi-isometry respect to the hyperbolic metric with $A = K$ and $B = K \log 2$ if $1 \leq K \leq 2$ and $B = 2.37(K - 1)$ if $K > 2$. Note that small circles are asymptotically the same for the two metrics, so there is no difference between "hyperbolic-quasiconformal" and "Euclidean-quasiconformal" maps. There is a difference, however, between "hyperbolic biLipschitz" and "Euclidean biLipschitz".

THEOREM 93. *For a map $f : \mathbb{D} \to \mathbb{D}$ we have (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) where*

(1) *f is biLipschitz with respect to the hyperbolic metric.*
(2) *f is quasiconformal.*
(3) *f is a quasi-isometry with respect to the hyperbolic metric.*
(4) *There is a hyperbolic biLipschitz map $g : \mathbb{D} \to \mathbb{D}$ so that $g|_{\mathbb{T}} = f|_{\mathbb{T}}$.*

In other words, the three classes of maps (hyperbolic biLipschitz, quasiconformal, hyperbolic quasi-isometry) all have the same set of boundary values.

The boundary extension is a quasisymmetric homeomorphism, i.e., there is an $k < \infty$ (depending only on K) so that $1/k \leq |f(I)|/|f(J)| \leq k$, whenever $I, J \subset \mathbb{T}$ are adjacent intervals of equal length. Conversely, any quasisymmetric homeomorphism of $\mathbb{T}$ can be extended to a K-quasiconformal selfmap of the disk, where K depends only on k.

5. The Beltrami equation

Conversely, the "measurable Riemann Mapping theorem" says that given any such μ, there is a K-quasiconformal map f with $\mu = \overline{\partial} f / \partial f$. The Beltrami equation $\overline{\partial} f = \mu \partial f$ can be solved using a power series in μ by setting

$$f = P[\mu(h + 1)] + z,$$

and

$$h = T\mu + T\mu T\mu + T\mu T\mu T\mu + \dots,$$

where T is the Beurling transform

$$Th(w) = \lim_{r\to 0} \frac{1}{\pi} \iint_{|z-w|>r} \frac{h(z)}{(z-w)^2} dxdy,$$

and P is the Cauchy integral

$$Ph(w) = -\frac{1}{\pi} \iint h(z)(\frac{1}{z-w} - \frac{1}{z})dxdy.$$

Formally, $\overline{\partial}P$ is the identity and $\partial P = T$. So if we choose f as above, then

$$\overline{\partial} f = \mu(h+1),$$

$$\partial f = T\mu(h+1) + 1 = h + 1.$$

Hence, $\overline{\partial} f/\partial f = \mu(1+h)/(1+h) = \mu$, as desired. To make the argument rigorous requires L^p estimates on these operators as described in [**?**], Chapter V.

CHAPTER 6

Schwarz-Christoffel iterations

In this chapter we consider a very simple iteration on the space of Schwarz-Christoffel parameters (n-tuples of the circle modulo Möbius transformations) which seems to work well in practice, but is not proven to converge in all cases. The method can be interpreted as a version of Newton's method for root finding in which the derivative is assumed to be the identity. We will then see how to improve the speed of covergence by using approximations of the Jacobian instead; first by discrete approximation and then by the faster, but less accurate, method of Broyden updates.

1. The space of n-tuples

Suppose Ω is polygonal with vertices $\mathbf{v} = \{v_1, \ldots v_n\}$ and $f : \mathbb{D} \to \Omega$ is conformal. The Schwarz-Christoffel formula ("SC-formula" below) says

$$f(z) = A + C \int_0^z \prod_{k=1}^{n} (1 - \frac{w}{z_k})^{\alpha_k - 1} dw,$$

where $\alpha\pi = \{\alpha_1\pi, \ldots, \alpha_n\pi\}$, are the interior angles of the polygon and $\mathbf{z} = \{z_1, \ldots, z_n\} \subset \mathbb{T}$ map to the corresponding vertices (these are called the SC-parameters). For fixed angles, evaluating the SC-formula defines a map S from n-tuples in $\mathbb{T}$ to n-tuples in $\mathbb{C}$. In fact, S is a well defined map from $\mathbb{T}^n_*$ (n-tuples of distinct points on $\mathbb{T}$ modulo Möbius transformations) into $\mathbb{C}^n_*$ (complex n-tuples modulo Euclidean similarities). Moreover, we can identify $\mathbb{T}^n_*$ with $\mathbb{R}^{n-3}$ as follows: fix a triangulation of the n points on $\mathbb{T}$, and for each pair of adjacent triangles record the logarithm of the cross ratio of the four vertices (the cross ratio is positive if we take the correct ordering of the four points). The original n-tuple can be recovered, up to Möbius transformations, from these $n - 3$ real values so $\mathbb{T}^n_* = \mathbb{R}^{n-3}$. This allows us to apply linear algebra to n-tuples on $\mathbb{T}$.

Suppose we have a explicit way of guessing the SC-parameters for a polygon, i.e., a map $G : \mathbb{C}^n_* \to \mathbb{T}^n_*$. Then $F = G \circ S$ gives a map $\mathbb{R}^{n-3} \to \mathbb{R}^{n-3}$ and the correct

SC-parameters for Ω solve $F(\mathbf{z}) = \mathbf{z}_0$ (where $\mathbf{z}_0 = G(\mathbf{v}) \in \mathbb{T}^n_*$) and hence are fixed by

$$\mathbf{z}_{k+1} = \mathbf{z}_k - A^{-1}(F(\mathbf{z}_k) - \mathbf{z}_0). \tag{21}$$

where A is an $(n-3) \times (n-3)$ matrix. Taking the derivative matrix $A = DF$ gives Newton's method (or take a discrete approximation to DF). If F is close to the identity, it may be much faster to just take $A = I$ (the identity). A compromise is to use Broyden updates, i.e., start with $A = I$ and multiply A by a rank one matrix at each step, chosen to optimize the approximation to DF given the evaluations of F made so far (see [**?**]). This is often fastest in practice. If $\mathbf{w} = \{w_1, \ldots, w_n\}, \mathbf{z} = \{z_1, \ldots, z_n\} \in \mathbb{T}$, define

$$d_{QC}(\mathbf{w}, \mathbf{z}) = \inf\{\log K : \exists\ K\text{-QC } h : \mathbb{D} \to \mathbb{D} \text{ such that } h(\mathbf{z}) = \mathbf{w}.\}$$

FMA is locally convergent with respect to d_{QC}. Is this true for some SC-iteration, i.e., can we choose G and A so that (21) always converges to the SC-parameters? For example, what about the following SC-iterations that are used in pratice?

2. Davis's iteration

Define G by $\arg(z_{k+1}) - \arg(z_k) = 2\pi|v_{k+1} - v_k|/\ell(\partial\Omega)$, i.e., parameter spacing is proportional to the edge lengths of the polygon. Davis's method is (21) with this G and $A = I$. See [**?**], [**?**]. The idea behind Davis' method is that for fixed angles, longer edges should have larger harmonic measure. This is false in general (see Figure 1). Howell [**?**] showed Davis's method in $\mathbb{T}^n$ locally diverges for this polygon, but my own experiments indicate this example converges in $\mathbb{T}^n_*$, i.e., the n-tuples diverge unless we renormalize at each step by Möbius transformations. Is this always the case? Prob Does Davis's method converge? Does it converge locally w.r.t. d_{QC}?

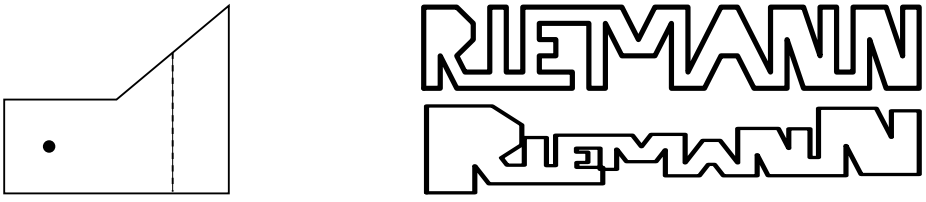


FIGURE 1. Left: longer edges can have less harmonic measure. Right: a polygon and the SC-image using the CRDT guessed parameters.

3. Newton's method

4. Broyden updates

CHAPTER 7

Tree-of-disk maps

We want to find a conformal map from a polygon to the unit disk. Möbius transformations are conformal, so could we build such a map from pieces that a Möbius? In this chapter we will show how to build maps from $\partial\Omega$ to $\mathbb{T}$ which are piecewise Möbius, easy to describe and compute explicitly and which approximate the boundary values of a conformal map in a quasiconformal sense. This approximation is not arbitrarily close, but it does lie within a quasiconformal neighborhood of the true conformal map and the size of the neighborhood is independent of the domain. Thus such maps may serve as good "initial guesses" for an iterative algorithm to find the conformal. Indeed, one of the methods we discuss was proposed by Driscoll and Vavasis exactly as a starting guess for their CRDT iteration (which we will discuss later in this chapter). The other example we discuss comes from the medial axis of a polygon and has close connections to 3-dimensional hyperbolic geometry.

1. The general set up

Suppose $\mathcal{D} = \{D_k\}$ is a finite collection of n disks in the plane (non necessarily distinct; we may also allow disks of zero radius, i.e., points, in some cases). Suppose also that these disks have been given the structure of a finite rooted tree, i.e., we have a distinguished element D_0 and a each disk D_k has a single parent D_k^* which is adjacent to it and closer to the root (in the tree distance). Finally, assume that for each non-root disk we are given a Möbius transformation $\tau_k : D_k \to D_k^*$. By composing these maps along the unique path from D_k to the root D_0 we obtain a Möbius transformation $\sigma_k : D_k \to D_0$. Clearly this map can be constructed from the τ_k's in time $O(n)$ by starting with the identity map $\sigma_0 : D_0 \to D_0$ and using

$$\sigma_k = \sigma_{k^*} \circ \tau_k,$$

to extend the map from a disk to its children (in this chapter, k^* denotes the index of the parent of D_k).

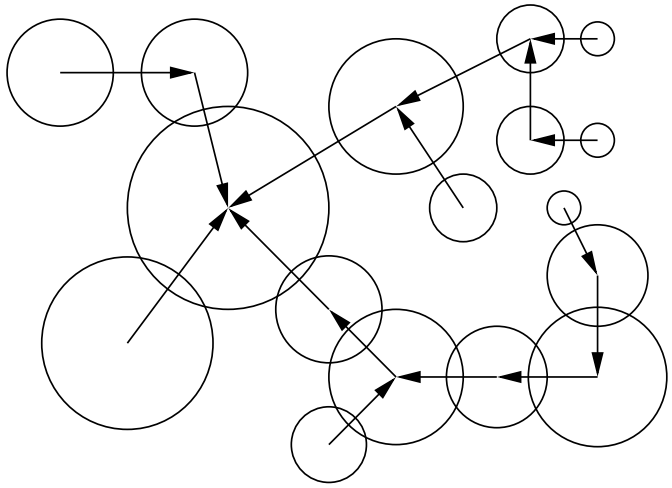


FIGURE 1. A tree-of-disks is a collection of disks in the plane, that form the vertices of a rooted tree. We are also give a Möbius transformation τ_D from each disk D to its parent D^*. By composing these maps we obtain a map ι_D from each disk to the root disk. Given a polygon, we wish to find an associated tree of disks which contains the vertices of the polygons and so that the ι map, restricted to the vertices of P, is a good approximation to a conformal map.

In this chapter we consider two approximations to the Riemann map which can be written as tree-of-disks maps. In the first example the tree comes from the triangulation tree of a polygon and was invented by Driscoll and Vavasis as part of their CRDT algorithm (although they used different method to describe it). In this case the disks are the circumdisks of triangles in a triangulation of the polygon. Such disk must overlap, and the τ maps we use are elliptic rotations around the intersection points. The second example uses disks and tree structure coming from the medial axis. Here the disks need not overlap, and more general transformations are used.

2. Triangulations

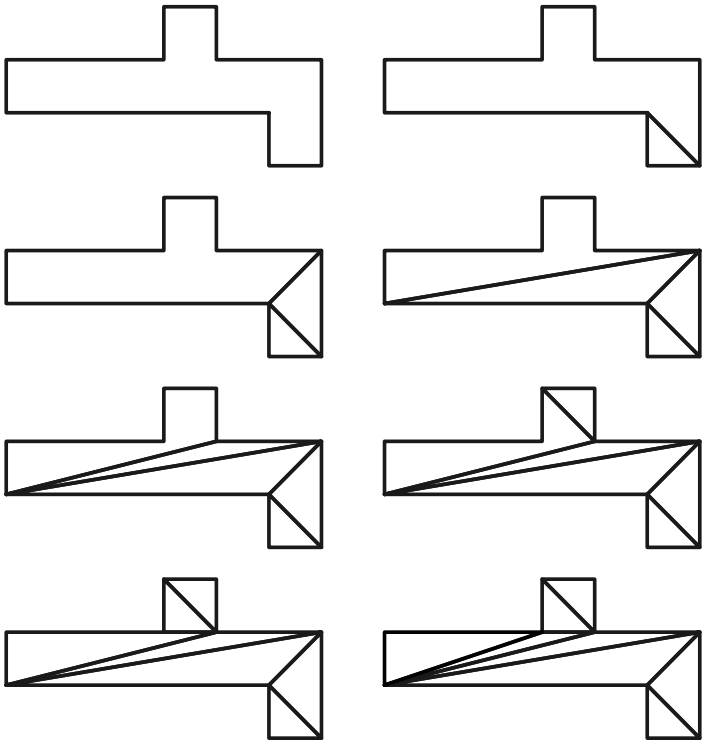


FIGURE 2. A triangulation of a simple polygon. We show the triangles as they are added one at a time, by the algorithm described in the text.

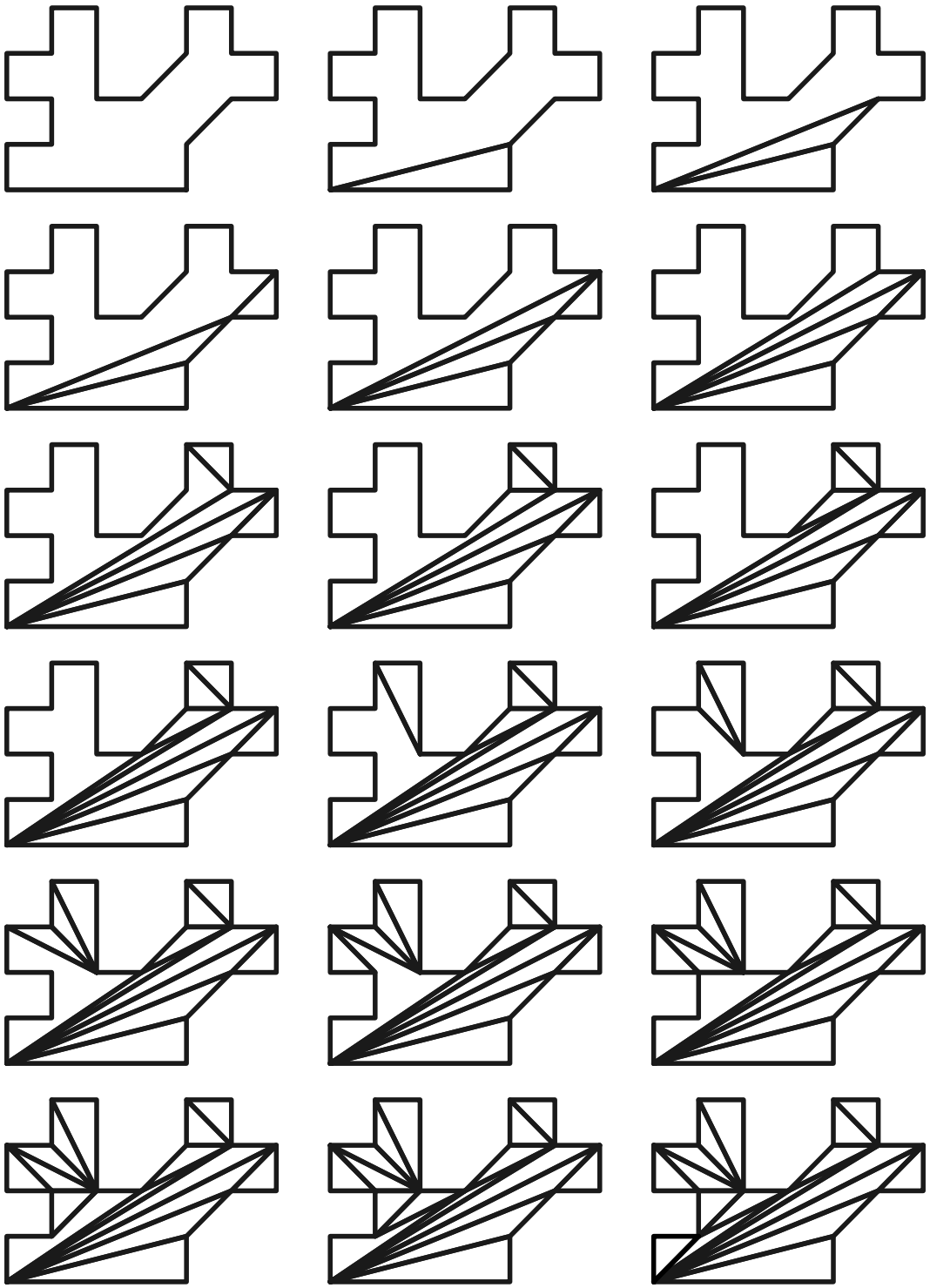


FIGURE 3. Another triangulation of a simple polygon.

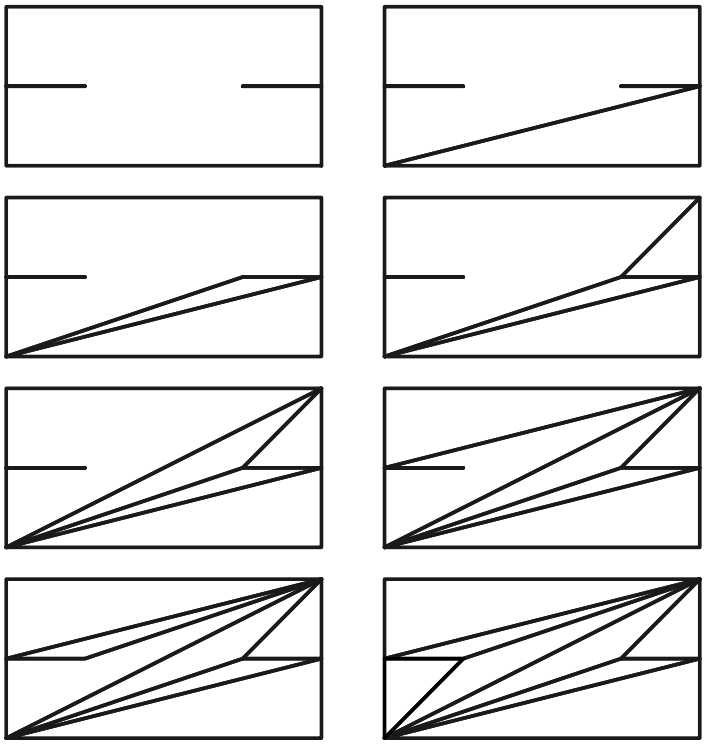


FIGURE 4. Triangulation of a non-simple polygon.

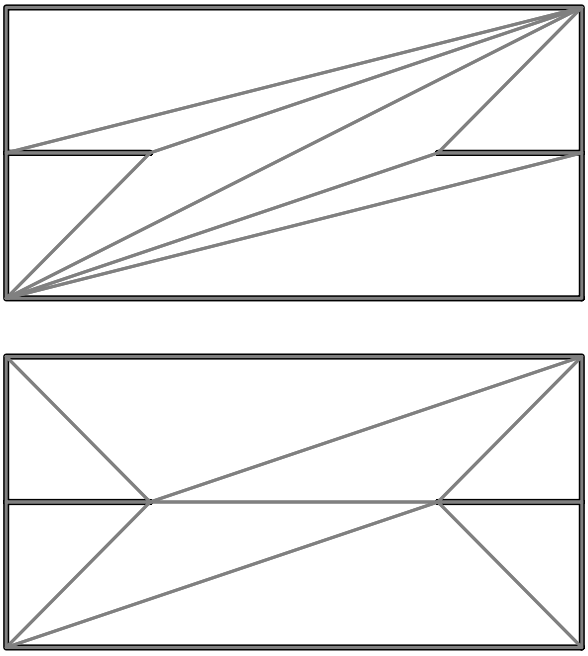


FIGURE 5. Delaunay triangulation obtained from general triangulation by flipping diagonals. Each figure shows one pass through the list of diagonals.

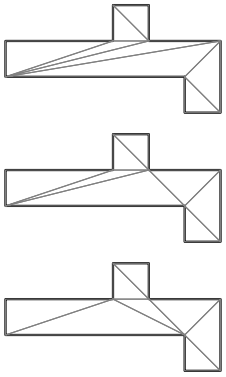


FIGURE 6. Delaunay triangulation obtained from general triangulation by flipping diagonals. Each figure shows one pass through the list of diagonals.

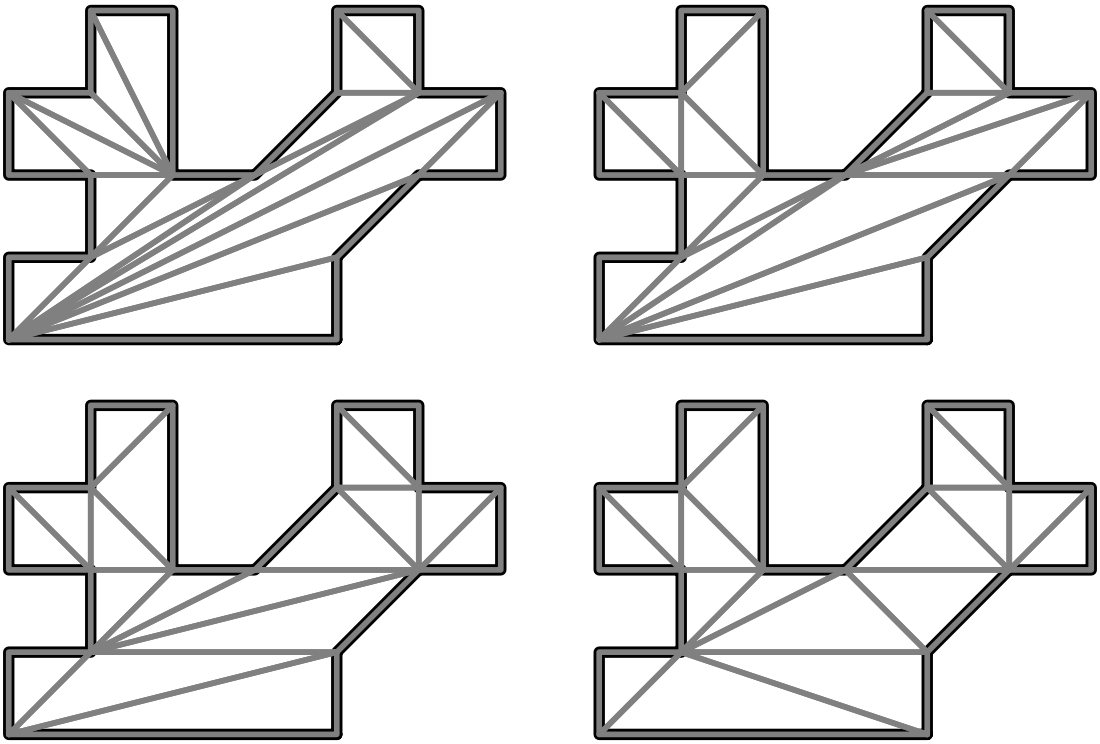


FIGURE 7. Delaunay triangulation obtained from general triangulation by flipping diagonals. Each figure shows one pass through the list of diagonals.

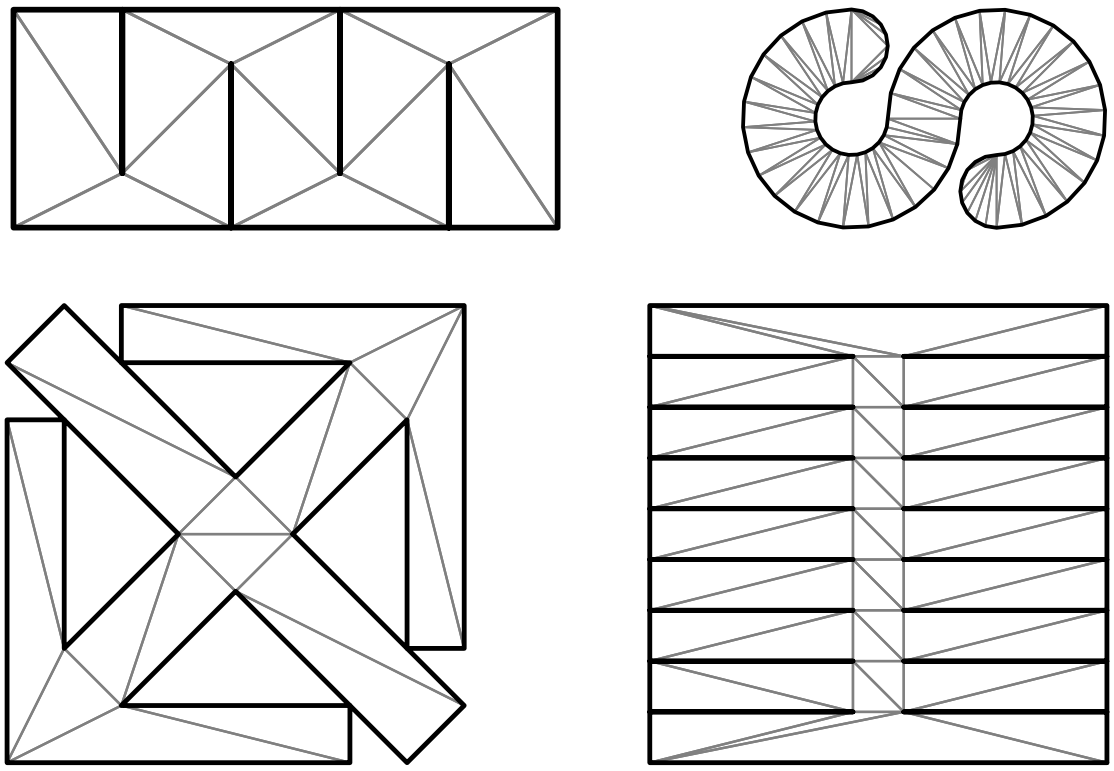


FIGURE 8. Delaunay triangulations of some more polygons.

3. Delaunay triangulations and CRDT

Every polygon P has a triangulation $\mathcal{T}$; this is collection of triangles formed by removing line segments that connect vertices of P within the interior of P. The triangles form the vertices of a tree if we define triangles to be adjacent iff they share one of the removed edges (i.e., share a boundary edge which is in the interior of P). Each triangle T is contained in a unique disk D whose boundary circle contains the three vertices of T; this is called the circumdisk. Note that circumdisks for adjacent triangles overlap since they contain at least two points in common (the endpoints of the shared edge of the two triangles). It is possible that the two disks are actually identical, as in the case of a triangulation of a square by a diagonal. In this case, we may wish to identity the two disks as a single disk. Alternatively, we may consider them as distinct, but require the map between them be the identity.

When adjacent disks are distinct then their boundaries overlap at exactly two points and there is an "obvious" Möbius map from one to the other given by an elliptic rotation around these two points. Suppose D_1, D_2 are the disks and we want to map $D_2 \to D_1$. Note that $D_2 \setminus D_1$ is a crescent, and has an interior angle θ at each vertex. The desired map is given by

$$\tau_{a,b,\theta}(z) = \frac{(be^{i\theta} - a)z + ab(1 - e^{i\theta})}{(e^{i\theta} - 1)z + (b - ae^{i\theta})}. \tag{22}$$

The formula can be easily derived by sending a and b to 0 and ∞ by the map $w = \tau(z) = (z-a)/(z-b)$, then multiplying by $e^{i\theta}$ and then applying the inverse map $z = \tau^{-1}(w) = (bw - a)/(w - 1)$. Geometrically, a crescent is foliated by circular arcs orthogonal to both boundaries and this elliptic transformations identifies endpoints of leaves. See Figure 10.

Thus to any rooted triangulation of a polygon, we can associate a tree-of-disks where all the edge maps are elliptic. We can compose the edge maps to form an ι map from each disk to the root. Since each vertex of P in on the boundary of some triangle in the triangulation, it is also on the boundary of some disk in our tree-of-disks and hence the ι map can be applied to it (one should check that if a vertex is on the boundary of more than one disk, then we get the same image regardless of which ι map we apply, but this is easy since the τ map corresponding to a triangulation edge e fixes the endpoints of e).

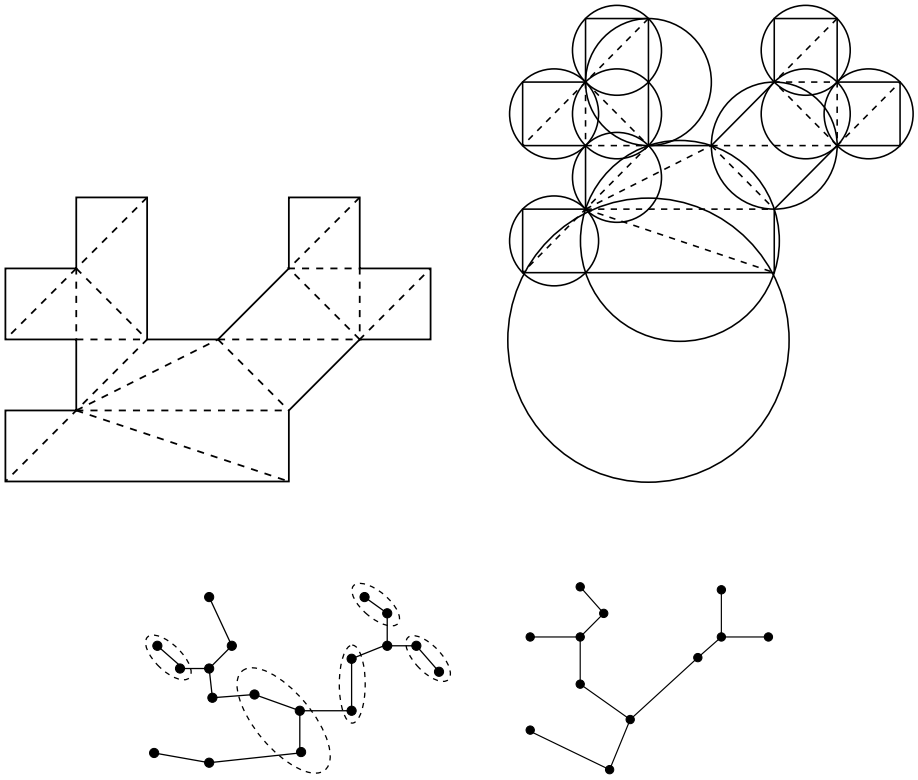


FIGURE 9. On the top we show a triangulated polygon and the corresponding collection of circumdisks. On the bottom we show the tree corresponding to the triangulation and the tree corresponding to the circumdisks, obtained from the triangulation tree by identifying adjacent vertices with identical corresponding disks (vertices to be identified are grouped by the dashed lines).

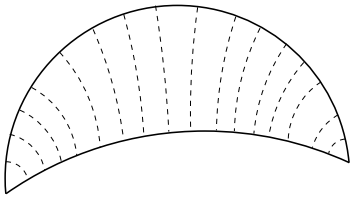


FIGURE 10. The orthogonal foliation of a crescent

We would like to claim that the ι map we have constructed is uniformly close to the a conformal map, i.e.,

$$d_{QC}(\iota(\mathbf{v}), \mathbf{z}) < C,$$

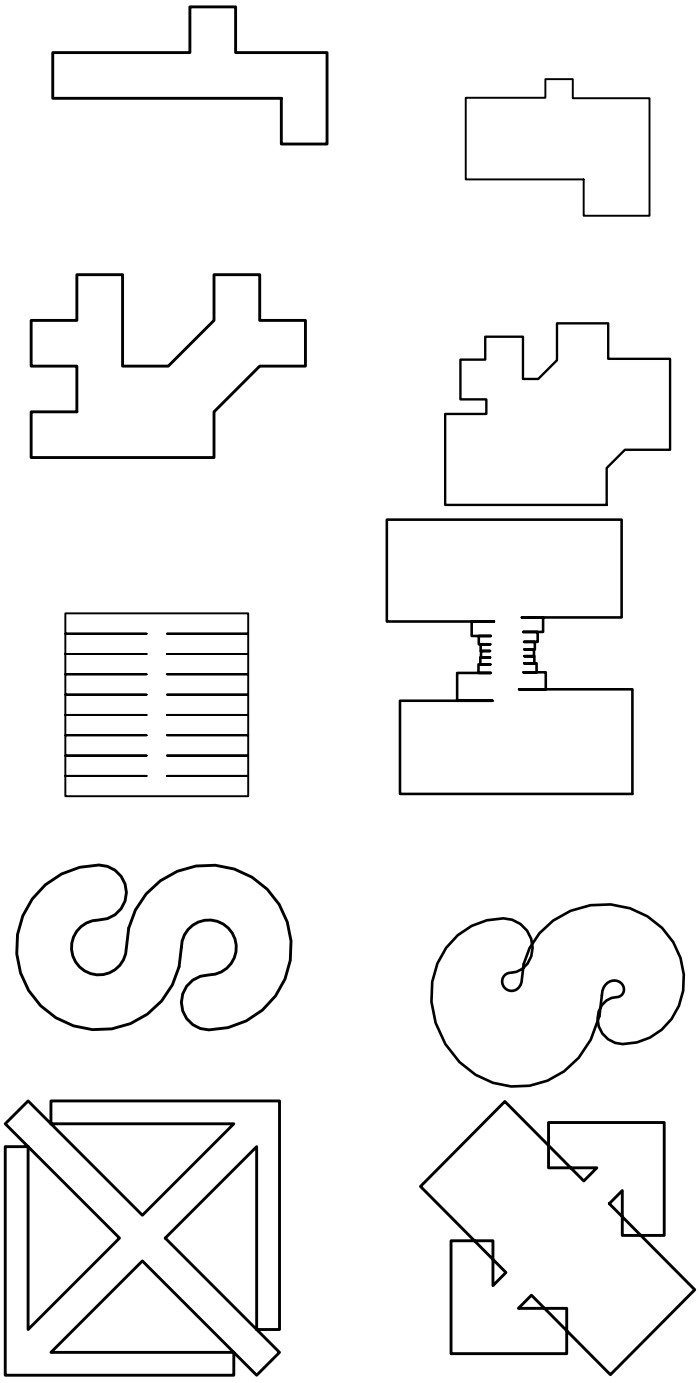


FIGURE 11. On the left are various polygons and on the right is the corresponding Schwarz-Christoffel image using the parameters given by the CRDT intial guess. We have not performed the first step of CRDT by adding extra vertices. Note that in some cases the SC image is not simple.

for some universal C independent of n or the polygon. Such an estimate is true, but we need to impose some extra conditions:

(1) the triangulation should be a Delaunay triangulation.
(2) extra vertices (of angle π) are added to any edges that are "too long".

We will explain each of these conditions in turn.

A polygon can have more than one triangulation, but there always exists a special triangulation called the Delaunay triangulation with the property that when any two of the triangles T_1, T_2 meet along an interior edge e, the two angles opposite e sum to π or less. An equivalent condition is that the open disk D_1 whose boundary contains the three vertices of T_1 does not contain T_2 and similarly for D_2 and T_1. See Figure 5.

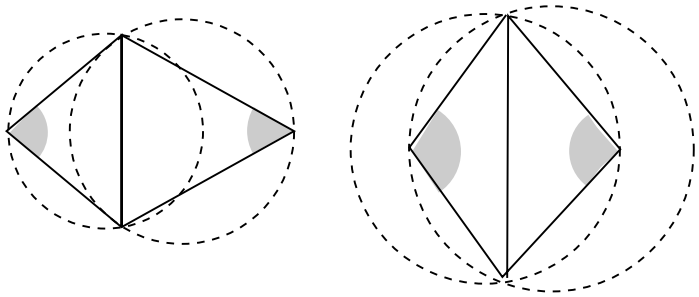


FIGURE 12. The triangles on the left satisfy the Delaunay condition, but the ones on the right do not.

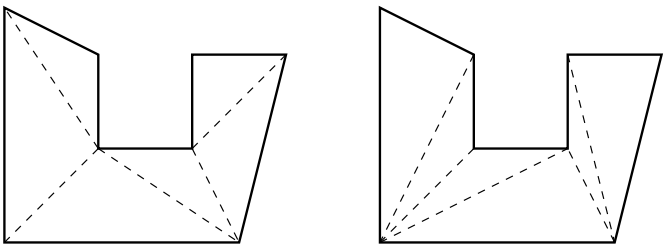


FIGURE 13. A Delaunay and non-Delaunay triangulation. We have draw a circumcircle on the right which fails the Delaunay condition.

Every polygon has a Delaunay triangulation which is unique except when adjacent triangles define four vertices on a common circle: then we may "flip" the diagonal of the resulting quadrilateral and get another Delaunay triangulation. Note however, that the corresponding family of disks and the tree-of-disk map on the vertices does

not depend on the choice of Delaunay triangulation. Delaunay triangulations have many nice properties and have been intensively studied (e.g., they are dual to Voronoi diagrams; they maximize the minimum angle in the triangulation [**?**]; they minimize the largest circumcircle [**?**], [**?**]). The basic facts can be found in various sources such as [**?**], [**?**], [**?**], [**?**].

The second condition said that we may need to add extra vertices to P to make sure that its Delaunay triangulation consists of "roundish" triangles. The precise result we need is

LEMMA 94. *We can add extra vertices of angle π to the edges of P so that the new polygon has a triangulation with the property that every triangle*

(1) *is isosceles with the base adjacent to its parent and base angle $\leq \pi/4$,*
(2) *is isosceles with any angle and is a leaf of the triangulation tree with the base adjacent to its parent or*
(3) *has all angles in $[\theta, \pi - \theta]$ for some $\theta > 0$.*

PROOF. We follow the construction from [**?**], which has two steps. In the first step, every vertex with interior angle $\leq \pi/4$ is "chopped off" by taking the largest isosceles triangle T in P formed with v and subsets of its two adjacent edges and adding two new vertices to P at the midpoints of the two sides of T. A new polygon P' is formed by replacing v by these two new vertices and the edge between them. This edge is protected: no new vertices may be added to it later. The second step is iterative. For each edge e in the current polygon compute its length L and the minimal distance D in P from e to any vertex which is not one of its endpoints (distance is the path distance within the polygon). If $D < L/(3\sqrt{2})$, the edge e is split into three equal edges and the process continues. See Figure **??**. The filled dots indicate the original vertices and the open dots the vertices added by the algorithm.

Driscoll and Vavasis give a lower bound r_0 on the shortest edge that can be produced: r_0 is the minimum over all unprotected edges of the path distance in P from that edge to any non-adjacent edge. Since every step reduces the length of some edge by a third, this proves the process terminates in a finite number of steps (and gives an estimate for the number of vertices that have been added). Thus the resulting polygon has a triangulation in which every triangle T is either (1) isosceles

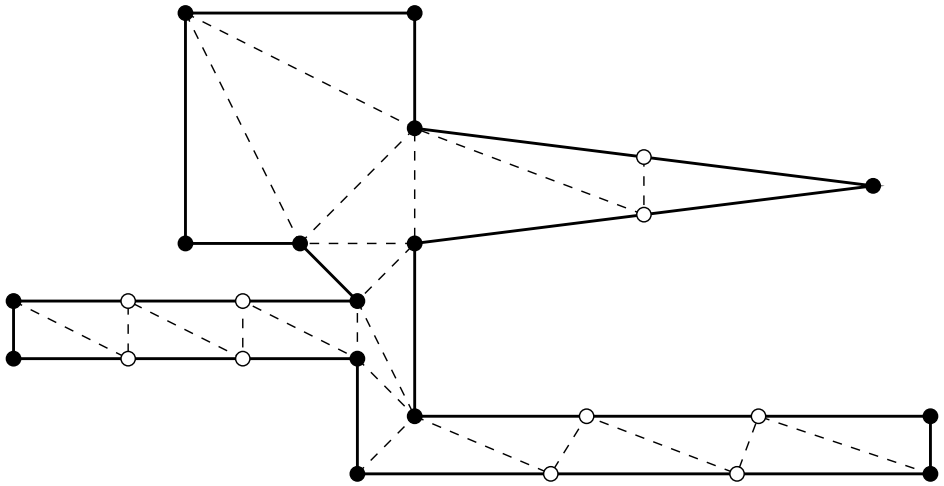


FIGURE 14. Adding extra points to a polygon in the CRDT algorithm to remove small angles from the triangulation

with angle $\leq \pi/4$ or (2) $D \geq L/(3\sqrt{2})$ for any side of T. The later condition easily implies the angles of T are bounded away from 0, so Lemma **??** holds. □

THEOREM 95. *Assuming we take a Delaunay triangulation and add extra vertices to O as described above, there is a $C < \infty$ so that the Driscoll-Vavasis guess w satisfies $d_{QC}(\mathbf{w}, \mathbf{z}) \leq C$.*

As Driscoll and Vavasis note in their paper [**?**], the number of new points needed can be arbitrarily large, depending on the geometry of P (if P has a long, narrow corridor, then many new vertices have to be added). It might be better to have a tree-of-disks map that had the uniform approximation property described in Theorem 95, but which also comes with a uniform time bound to compute it. Such a map will be described in the next section.

In their paper [] Driscoll and Vavasis did not express the initial guess of CRDT as a tree-of-disks map. Instead they described it using the domain decomposition idea described in Section 5 of Chapter 5. We review their definition briefly and then show it is the same as what we have done.

Recall that given four distinct points a, b, c, d in the plane we define their cross ratio as

$$\mathrm{cr}(a, b, c, d) = \frac{(d-a)(b-c)}{(c-d)(a-b)}.$$

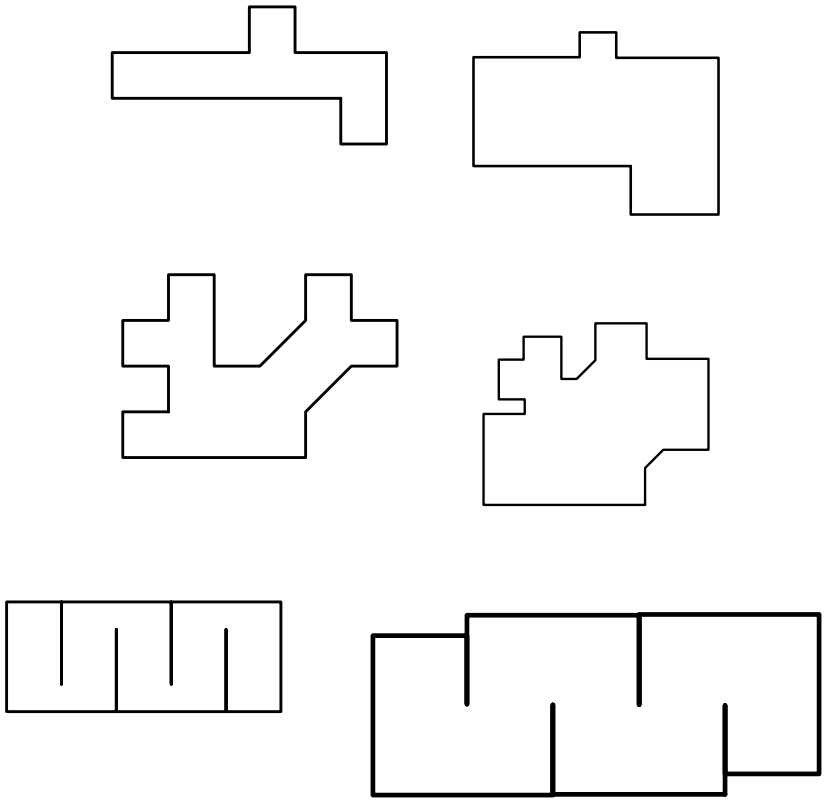


FIGURE 15. Three target polygons on the left and the SC image using the initial guess of the CRDT algorithm.

Note that $z \to \mathrm{cr}(a, b, c, z)$ is the unique Möbius transformation which sends a to 0, b to 1 and c to ∞. Suppose P is a simple n-gon and that $\mathcal{T} = \{T_1, \ldots T_{n-2}\}$ is a Delaunay triangulation of P. Let D_k be the circumcircle associated to each triangle T_k, $k = 1, \ldots, n-2$. Choose some root triangle for the triangulation and map its vertices to any three points in $\mathbb{T}$ with the correct orientation. In general, suppose Q is the quadrilateral formed by two adjacent triangles T_1 and T_2 (which have vertices v_1, v_2, v_3 and v_1, v_3, v_4 in counterclockwise order respectively). Also suppose that we have already defined $w_1, w_2, w_3 \in \mathbb{T}$. Then w_4 is uniquely determined by the condition

$$\mathrm{cr}(w_1, w_2, w_3, w_4) = -|\mathrm{cr}(v_1, v_2, v_3, v_4)|. \tag{23}$$

It is easy to see by induction that this uniquely determines the points $\mathbf{w}$ up to a Möbius transformation of the circle.

Now suppose D_1, D_2 are the circumdisks associated to T_1, T_2 and $\tau_{D_2} : D_2 \to D_1$ is the elliptic Möbius transformation that fixes the two points $\partial D_1 \cap \partial D_2$. After conjugating by a Möbius transformation η which sends $v_1 \to 0$, $v_2 \to 1$ and $v_3 \to \infty$, the elliptic map τ_{D_2} is conjugated to the Euclidean rotation around 0 which sends the image to v_4 onto the negative real axis. Thus

$$\mathrm{cr}(v_1, v_2, v_3, \tau_{D_2}(v_4)) = -|\mathrm{cr}(v_1, v_2, v_3, v_4)|.$$

The Möbius transformation $\iota_{D_1} : D_1 \to D_0$ is then applied to the four points $\{(v_1, v_2, v_3, \tau_{D_2}(v_4)\}$ to give $\{\iota(v_1), \iota(v_2), \iota(v_3), \iota(v_4)\}$. Since Möbius transformations do no change cross ratios

$$\mathrm{cr}(\iota(v_1), \iota(v_2), \iota(v_3), \iota(v_4)) = -|\mathrm{cr}(v_1, v_2, v_3, v_4)|.$$

Therefore $\mathbf{w} = \iota(\mathbf{v})$ up to a Möbius transformation of the circle.

4. The CRDT iteration

As noted above Driscoll and Vavasis developed their initial guess for the Schwarz-Christoffel parameters merely as a first step to an iteration that would attempt to approximate the actual parameters arbitrarily well. We now describe this part of the their CRDT algorithm.

Given an n-gon, compute a rooted Delaunay triangulation (which has $n-2$ triangles) and form the collection of $n-3$ quadrilaterals $\{Q_j\}_{j=1}^{n-3}$ of adjacent triangles. Let σ_j be the cross ratio of the four vertices of Q_j (to be precise, assume the vertices are ordered counterclockwise staring with the vertex in the parent triangle which is opposite the side shared with the child triangle). Let $c_j = \log|\sigma_j|$. This defines map

$$G : \mathbb{C}^n \to \mathbb{R}^{n-3},$$

which takes the vertices of the polygon to the list of cross ratios.

Define a map from $\Phi : \mathbb{R}^{n-3} \to \mathbb{C}^n$ using the Schwarz-Christoffel map as follows. From $n-3$ real numbers c_j, find $\sigma_j = \exp(c_j)$. Then, using induction, find a triangulated n-gon inscribed in the unit circle so that corresponding quadrilaterials have cross ratios σ_j. Then use these points as the z-parameters and the fixed set of α's to evaluate the Schwarz-Christoffel formula

$$\int_0^{z_k} \prod_{k=1}^{n} (1 - \frac{w}{z_k})^{\alpha_k - 1} dw,$$

to find each of the n image vertices.

We want to find a set of parameters for the Schwarz-Christoffel formula which gives a map (after linear renormalizing) onto our target polygon Ω_0. In other words we want to solve $\Phi(\mathbf{z}) = \mathbf{v}$. This is the same as solving $G(\Phi(\mathbf{z})) = G(\mathbf{v})$ (at least if we can show G is 1-1, which it is). Define another map $F : \mathbb{R}^{n-3} = \mathbb{R}^{n-3}$, by $F = G \circ \Phi - G(\mathbf{v})$. Then the equation we want to solve is

$$F(\sigma) = 0.$$

We do this by defining a sequence

$$\sigma_0 = G(\mathbf{v}),$$

$$\sigma_{k+1} = \sigma_n - F(\sigma_k).$$

Driscoll and Vavasis report that in experiments, this iteration converges linearly, i.e.,

$$\|F(\sigma_{k+1})\|_2 \leq \lambda \|F(\sigma_k)\|_2,$$

for some $\lambda < 1$. However, there is currently no proof, or even explanation, for why this holds.

The linear convergence observed in experiments is too slow for practical use, so Driscoll and Vavasis us a non-linear equation solver NESOLVER due to Behrens. This is the same as used by Driscoll's SC-Toolbox. It uses approximations of F' to implement a Newton type iteration. Driscoll and Vavasis report in [**?**] that this also gives linear convergence, but at a better rate than the simple iteration.

5. The medial axis

The medial axis of Ω consists of the centers of all disks in Ω whose boundary hits $\partial\Omega$ in two or more points. See Figure 22. It is a one dimensional object that is used in computer science to encode the shape of 2-dimensional objects and was introduced by Blum in 1967 [**?**] (the same set, with a different name, also appears in a 1945 paper of Erdös [**?**]). There is a large literature describing its mathematical properties and numerous applications including [**?**], [**?**], [**?**], [**?**], [**?**], [**?**]. It is a theorem of Chin, Snoeyink and Wang [**?**], [**?**] that the medial axis of polygon with n vertices can be computed in $O(n)$ steps. This depends on a deep and difficult result of Chazelle [**?**] that a polygon can be triangulated in linear time. However, special cases, such as

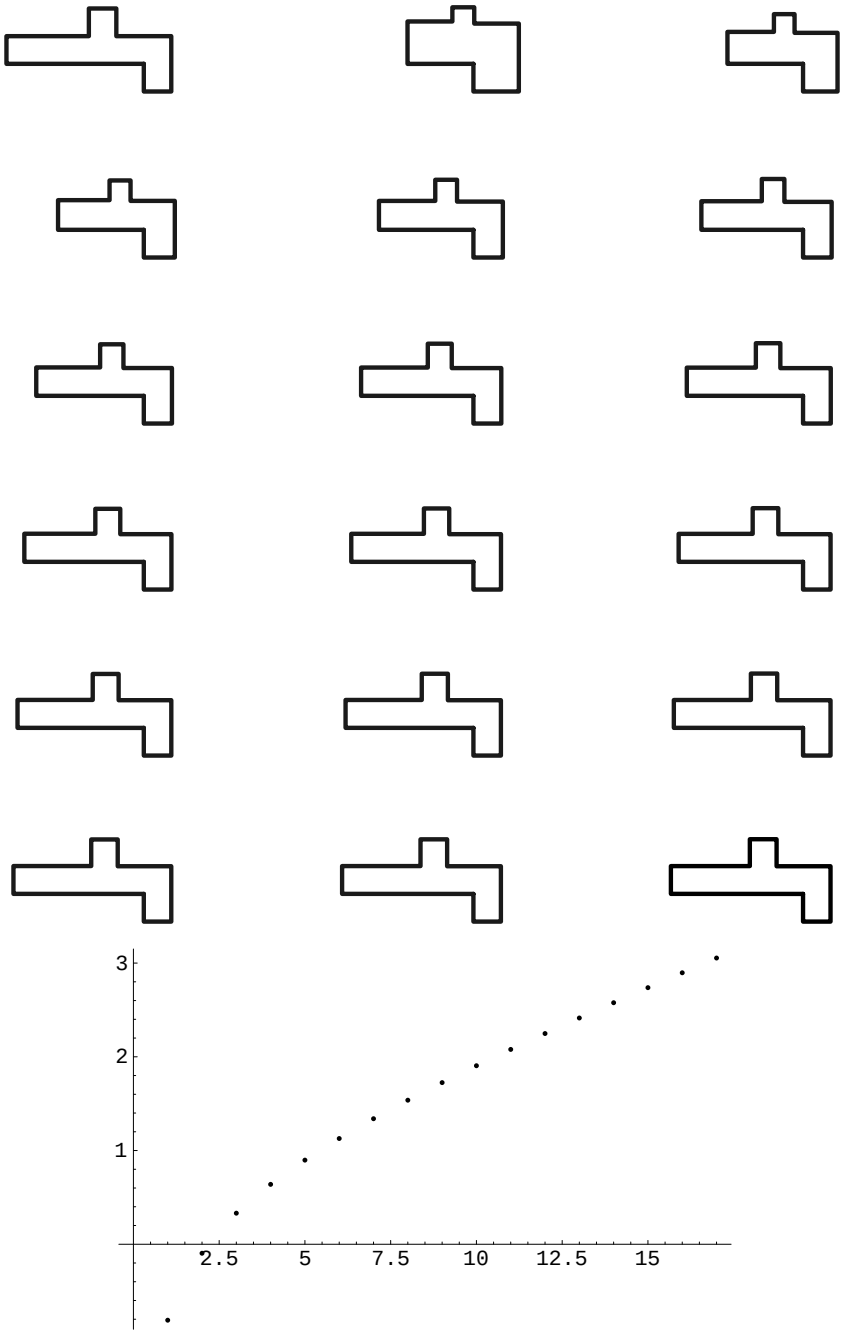


FIGURE 16. Simple CRDT.

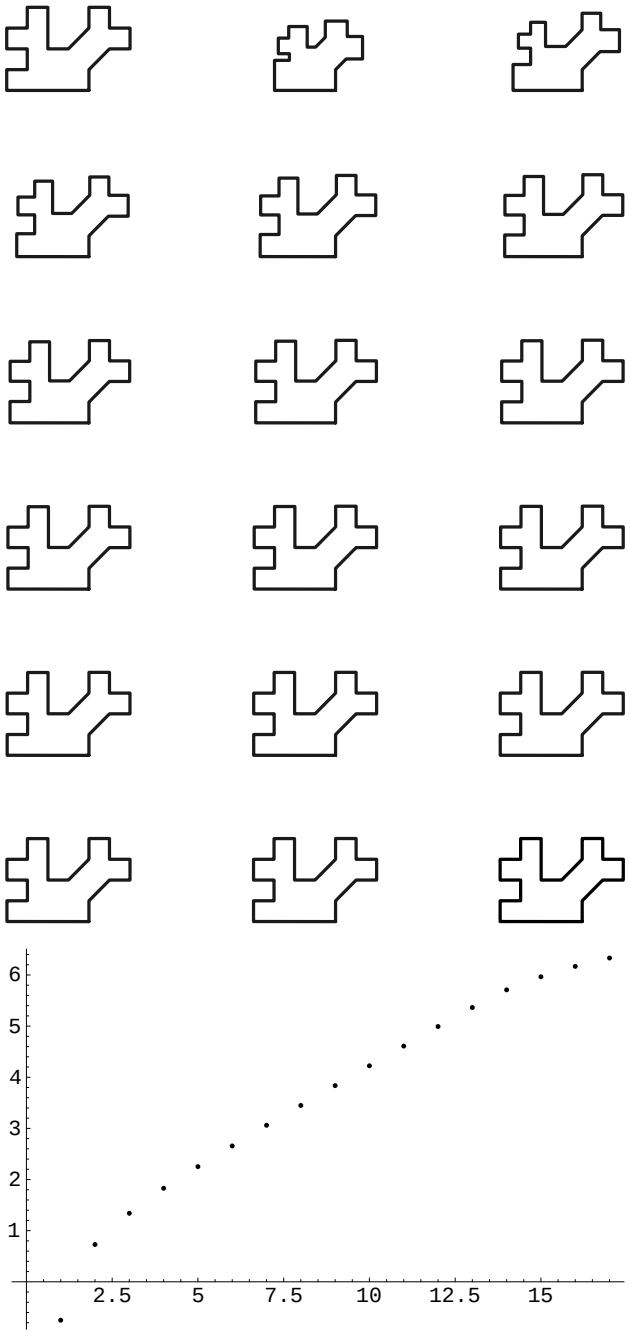


FIGURE 17. Simple CRDT.

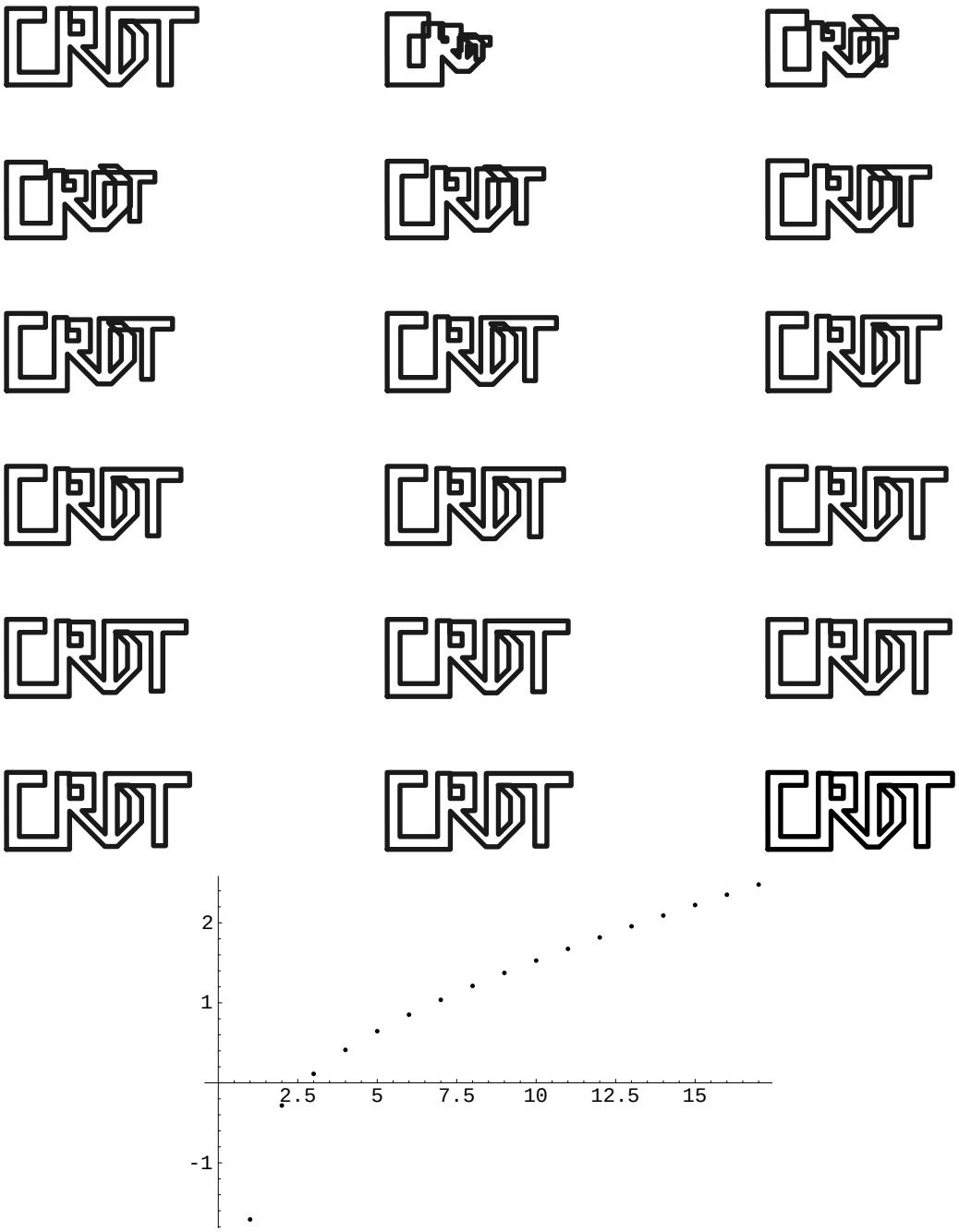


FIGURE 18. On the top is the target polygon and 17 iterations of the simple CRDT iteration, reaching a QC error of at most 1.08388. On the bottom is the graph of the log of the QC errors.

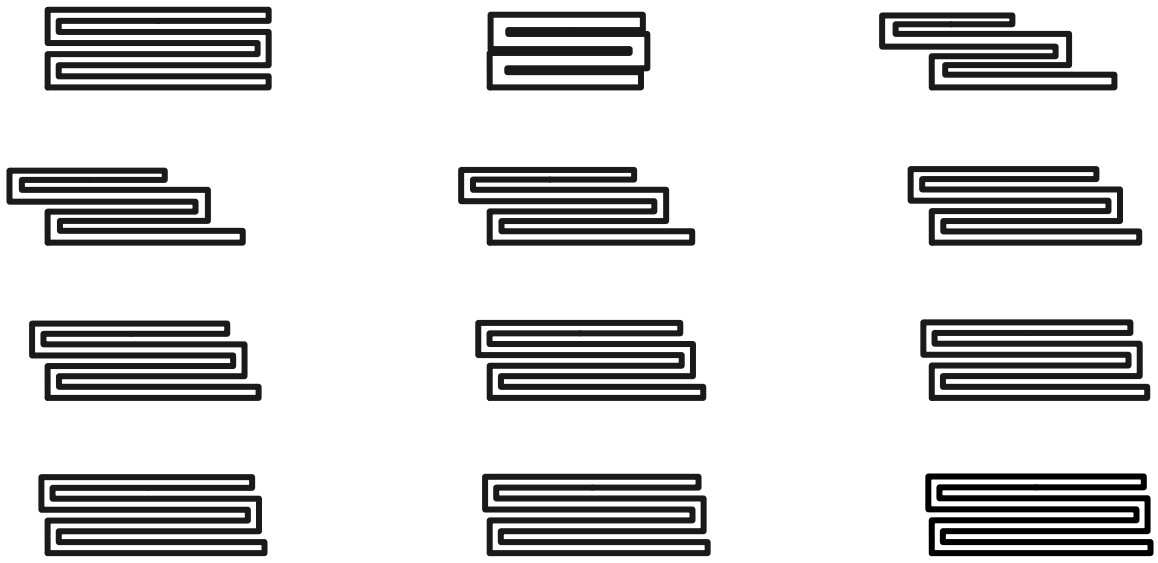


FIGURE 19. On the top is the target polygon and 11 iterations of the simple CRDT iteration. The polygon contains 4 1×20 rectangles, so the aspect ration of the opposite thin sides is about 80 to 1. Even from the center, the separation of some Schwarz-Christoffel parameters must be less than $\simeq \exp(-\pi 40) \simeq 10^{-55}$ which is far smaller than machine precision $\approx 10^{-16}$. The remarkable aspect of CDRT is that it can compute the conformal map from the disk to such a region accurately regardless.

convex polygons are easier (see [**?**]) and $O(n \log n)$ methods have been implemented for general polygons.

How does the medial axis of $\partial\Omega$ give a map from $\partial\Omega$ to the boundary of a disk? If Ω is a finite union of disks, then its medial axis is a finite tree; vertices correspond to disk that hit the boundary in three or more points. We can rewrite Ω as the union of disks $\{D_k\}$ corresponding to vertices of this tree. Choose one of these, D_0, as the root. Then each non-root disk has a parent disk (the one closer to the root) and if we remove the parent from the disk, we are left with a crescent. Thus Ω may be written as the union of the root disk and a finite union of crescents.

Every crescent has two natural foliations by circular arcs: one by arcs passing through the two vertices (we call this the medial axis foliation) and the other by arcs perpendicular to the first (the medial axis flow). Following leaves of the medial axis flow gives an identification between the two boundary arcs of the crescent (this is

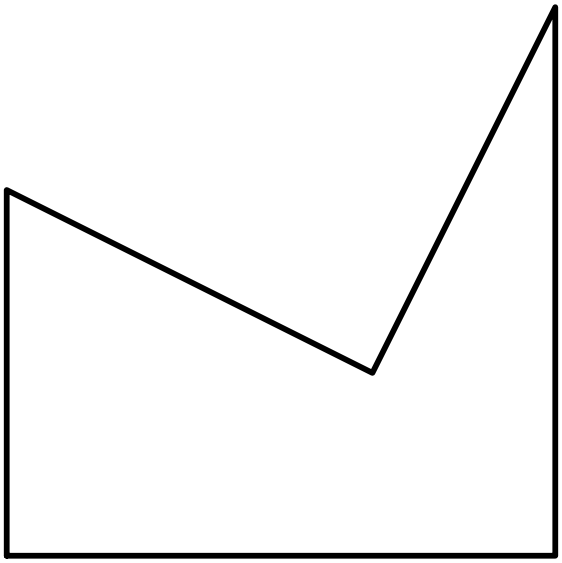


FIGURE 20. A pentagon used in Figure 21

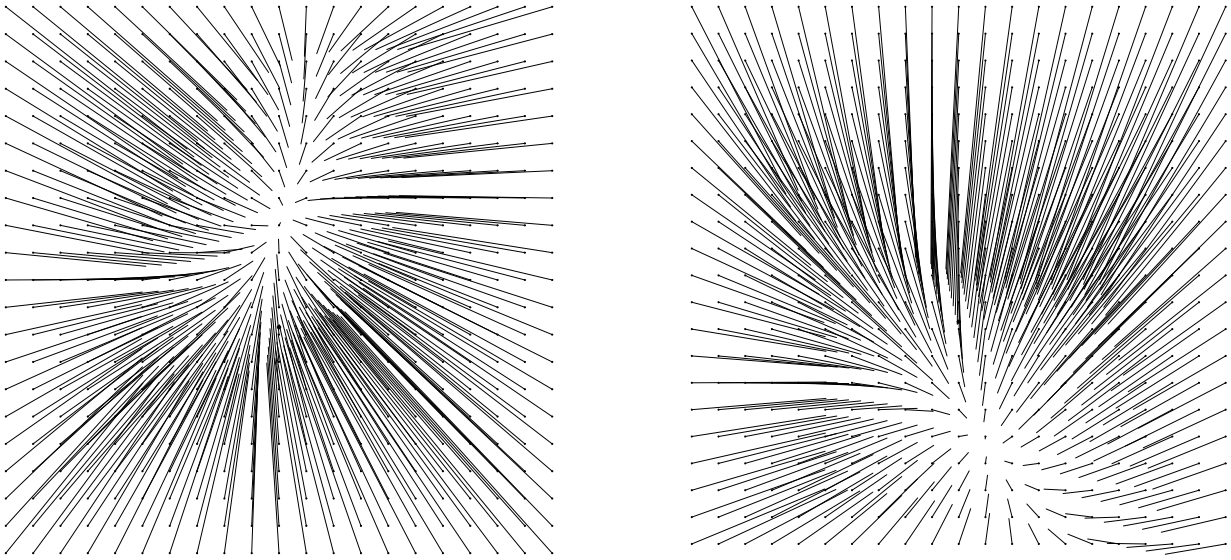


FIGURE 21. On the left is the composition of the SC map and the CRDT map for the pentagon shown in Figure 20. On the right is the simple CRDT iteation map $\mathbf{z} \to \mathbf{z} - (H(\mathbf{z}) - \mathbf{z}_0)$. It appears from the picture that iterations of this map will converge to a fixed point. In both pictures the coordinates are logs of cross ratios and take values in $[-5, 5]^2$.

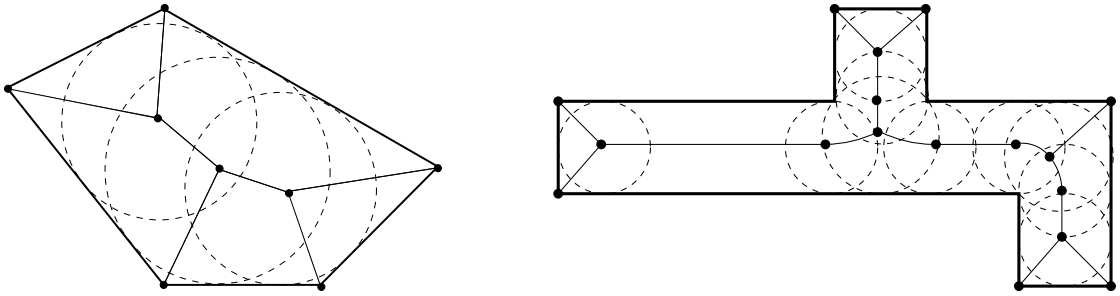


FIGURE 22. Examples of medial axes of polygons

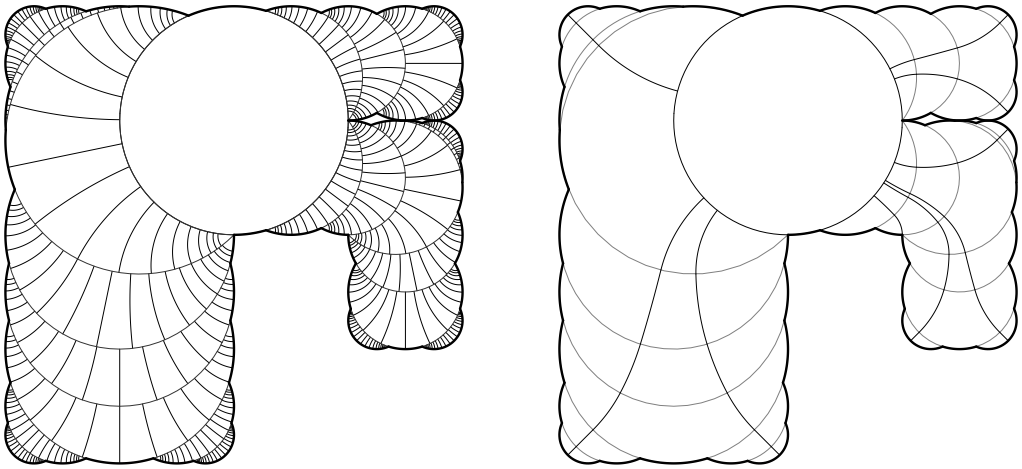


FIGURE 23. A finite union of disks written a union of a root disk D_0 and several crescents. Each crescent is foliated by circular arcs orthogonal to the boundary and following the foliation gives a map from $\partial\Omega$ to ∂D_0.

the same as applying an elliptic Möbius transformation which fixes each vertex and rotates one boundary arc to the other). Since Ω is written as a union of a root disk and a finite union of crescents, we can compose the maps on each crescent and get a map from $\partial\Omega$ to ∂D_0. We call this the medial axis flow from $\partial\Omega$ to ∂D_0. See Figure 23.

If Ω is bounded by an n-gon then the medial axis is still a finite tree and has $O(n)$ vertices and edges. The edges correspond to 1 parameter families of disks which meet $\partial\Omega$ at exactly two points and come in three types:

(1) a line segment that is equidistant from two vertices
(2) a line segment that is equidistant from two edges
(3) a parabolic arc that is equidistant from a edge and a vertex.

For each edge of the medial axis, consider the union of medial axis disks centered on that edge, minus the disk corresponding to the endpoint closer to the root. We call these subdomains "generalized crescents" (in case (1) it is an actual crescent). Together with the root disk, the generalized crescents decompose Ω into a finite number of pieces and on each piece there is a foliation by boundary arcs of medial axis disks and a corresponding orthogonal flow; the medial axis flow for that piece. Figure 24 shows this foliation and flow for different types of generalized crescents, Figure 25 shows the decomposition, foliation and flow for a polygon and Figure 26 shows just the foliation and flow for two more examples.

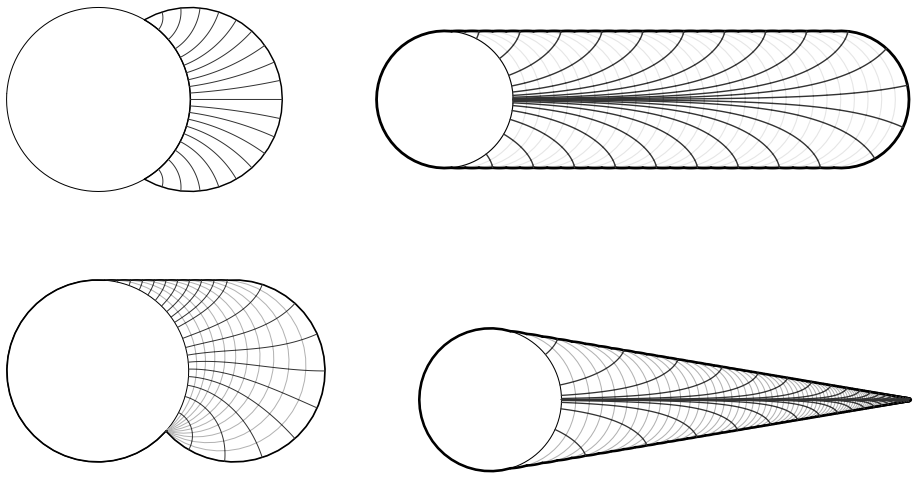


FIGURE 24. The medial axis foliation and flow corresponding to different types of medial axis edges. The white disk in each picture corresponds to the endpoint closer to the root.

To visualize where the vertices of the polygon are mapped we simply follow the medial axis flow from $\partial\Omega$ to ∂D_0. However, to actually compute these images, we start by defining a map from each medial axis vertex disk to its parent.

Suppose e is an edge of the medial axis, D, D^* are the disks corresponding to its endpoints and D^* is the parent of D (i.e., is closer to the root). Let Ω_e be the

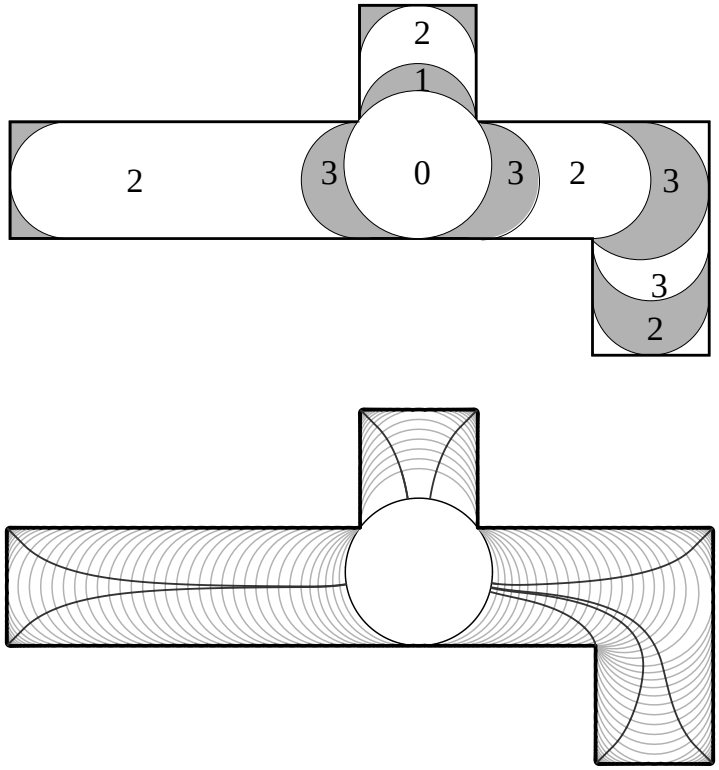


FIGURE 25. This shows the medial axis flow for the polygon illustrated on the right of Figure 22. The top picture shows the a root disk and the generalized crescents, labeled by their type. The bottom figure shows the medial axis flow. The landing points of the vertices are the parameter guesses.

corresponding generalized crescent. The medial axis flow in Ω_e defines a map from $\partial D \cap \partial\Omega_e$ to $\partial D^* \cap \partial\Omega_e$. For type 1 edges this is just an elliptic Möbius transformation (with fixed points $\partial D \cap \partial D^*$), and we shall see that it is a Möbius transformation $\tau_D : D \to D^*$ in all cases. Moreover, we shall give explicit formulas τ_D in terms of D, D^* and the type of edge. The medial axis flow map from D to the root, $\iota_D : D \to D_0$ is the composition of the maps τ_{D_k}, where $\{D_k\}$ is a path from D to D_0. This map can easily be computed in linear time by induction and the formula

$$\iota_D = \iota_{D^*} \circ \tau_D,$$

if the disks are arranged in a list starting with the root (for which ι_{D_0} is the identity map) and so that every disk comes later that its parent on the list. Since each map is a Möbius transformation, we only have to record a 2×2 matrix for each disk and the composition above corresponds to matrix multiplication.

For each concave vertex v of the polygon (i.e., those with interior angle $\leq \pi$), we will choose a medial axis disk D with $v \in \partial D$ and define $\iota(v) = \iota_D(v)$. If v is a convex vertex then it is the endpoint of a medial axis edge whose other endpoint corresponds to a disk D. We will map v to the closest point v^* of ∂D and then define $\iota(v) = \iota_D(v^*)$. This defines ι at every vertex and gives the parameter guesses $\mathbf{w}$ used in Theorem **??**.

THEOREM 96. *There is a $C < \infty$ (independent of P) so that the ι map satisfies $d_{QC}(\iota(\mathbf{v}), \mathbf{z}) \leq C$.*

Thus the medial axis flow provides an $O(n)$ algorithm for computing guesses for the Schwarz-Christoffel parameters that are guaranteed to be within a uniform distance of the correct parameters. We will prove this later without an explicit estimate of C, but a more detailed analysis (carried out in [**?**]) shows that we can take $C = \log 8$ and an example from [**?**] shows $C > \log 2.1$. However, there may be other, still undiscover3ed, tree-of-disk constructions which give fast approximations with even better bounds. Can the reader find one?

Figure 26 shows two more examples of the medial axis flow in a polygon. Figures 28 - **??** each show four pictures: the target polygon in the upper left, the medial axis flow in the upper right, the Schwartz-Christoffel image using the medial axis flow parameters in the lower left and the Schwarz-Christoffel image using equispaced parameters in the lower right. Comparing the latter two, we see that the medial axis flow always gives a better approximation and often gives quite a good approximation.

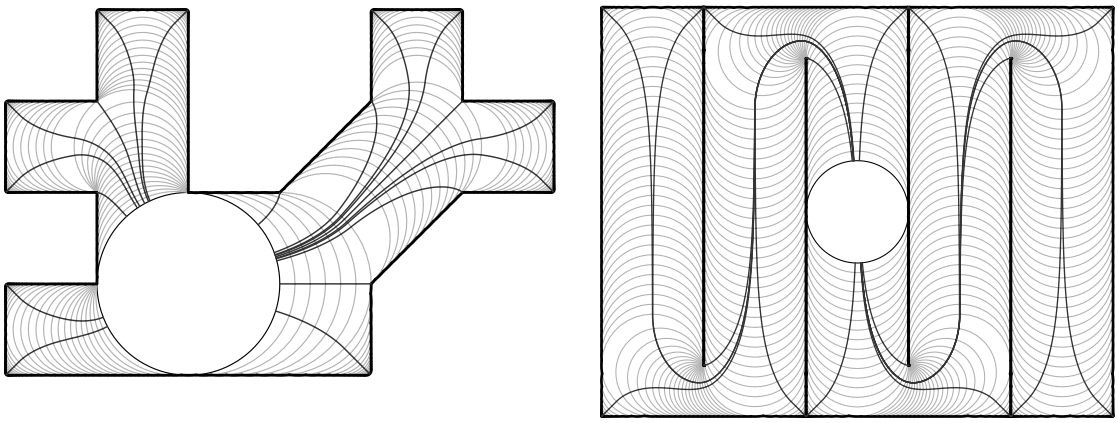


FIGURE 26. More examples of the medial axis foliation and flow.

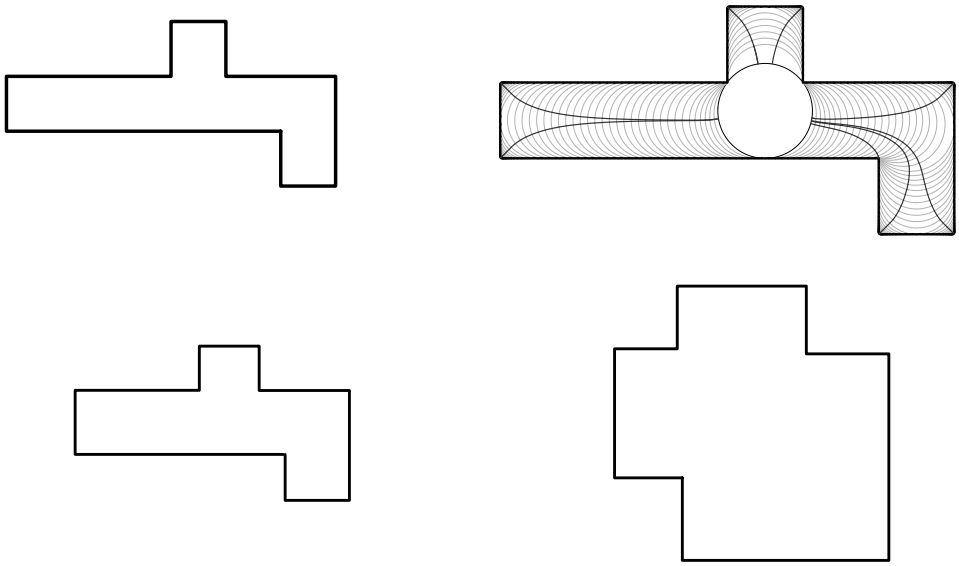


FIGURE 27.

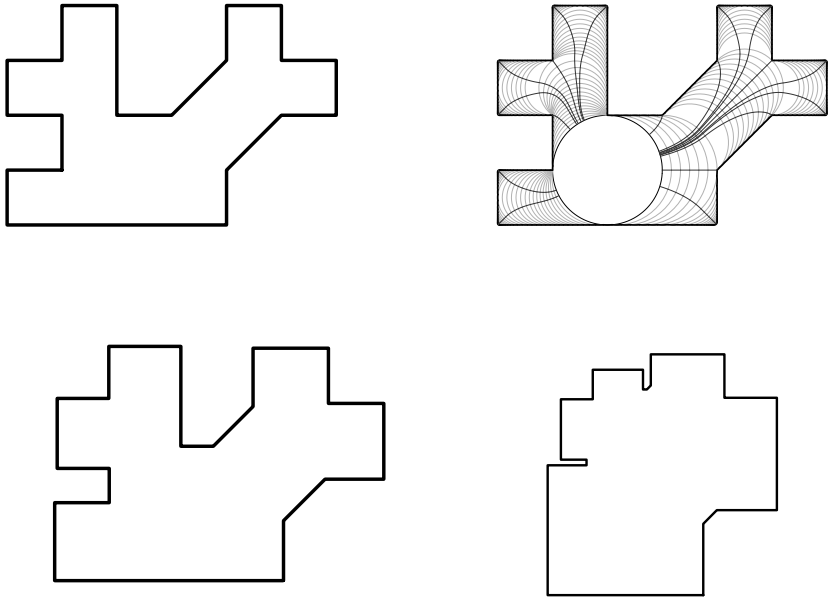


FIGURE 28.

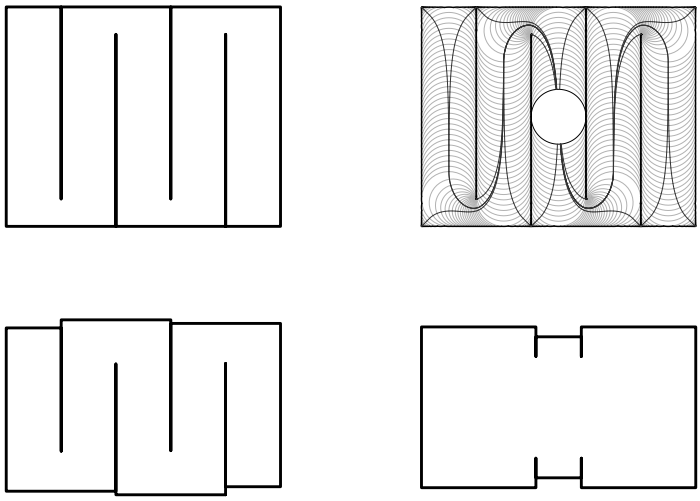


FIGURE 29.

6. Formulas for the ι map

As noted earlier, to compute ι for a polygon, we only have to compute the Möbius transformation for each generalized crescent, i.e., compute the maps between disks corresponding to vertices of the medial axis. Let $\mathcal{D}$ denote the disks that correspond to vertices of the medial axis and assume we have chosen a root disk, D_0. Below, we will describe how to map a disk $D_1 \in \mathcal{D}$ to its parent $D_2 \in \mathcal{D}$. The map from D_1 to the root disk can then be computed by composing with the corresponding map for the parent (which we may assume has already been computed). To map the vertices of $\partial\Omega$ to ∂D_0, map there are two cases. For concave vertices v (interior angle $\geq \pi$) choose a medial axis vertex disk D so that $v \in \partial D$ and apply the map ι_D to v. If v lies on the boundary of more than one such disk, each of the corresponding maps must agree at v, so it does not matter which disk we choose. Each convex vertex (interior angle $< \pi$) it the endpoint of a type 2 edge of the medial axis. The other endpoint is a vertex of the medial axis and we map v onto the boundary of the corresponding disk D by simply mapping it to the closest point of ∂D and then applying ι_D to map the point to ∂D_0.

We now consider each type of medial axis edge.

Case 1 (point-point bisector): If the two disks are $D_1 = D(z_1, r_1)$ and $D_2 = D(z_2, r_2)$ with D_2 being the parent, then the desired map is just the unique elliptic Möbius transformation which fixes the two points, a, b of $\partial D_1 \cap \partial D_2$ and maps $\partial D_1 \setminus D_2$ onto $\partial D_2 \cap \overline{D_1}$ by an elliptic rotation τ of angle α, the interior angle of the crescent. If $r = |z_1 - z_2|$, law of cosines implies

$$\cos\alpha = \beta = \frac{r^2 - r_1^2 - r_2^2}{-2r_1 r_2},$$

and hence $e^{i\alpha} = \beta + i\sqrt{1-\beta^2}$. Similarly,

$$a = z_2 + \frac{r_2}{r} e^{i\theta}(z_1 - z_2), \tag{24}$$

$$b = z_2 + \frac{r_2}{r} e^{-i\theta}(z_1 - z_2), \tag{25}$$

where $\cos\theta = (r_1^2 - r^2 - r_2^2)/(-2rr_2)$. Thus τ is given by the formula $\tau(z) = \sigma^{-1} \circ R_\alpha \circ \sigma$, where $R_\alpha(z) = e^{i\alpha} z$ and $\sigma(z) = (z-b)/(z-a)$ sends a to ∞ and b to zero.

Case 2a (edge-edge bisector, parallel case): Consider Figure 30. Suppose the two endpoint disks have centers z_1, z_2 and common radius $r_1 = r_2$. We normalize

so that these disks become $D_1 = D(0,1)$ and $D_2 = D(A,1)$ using the map

$$\eta(z) = (z - z_1) \cdot \frac{|z_2 - z_1|}{(z_2 - z_1)r_1}.$$

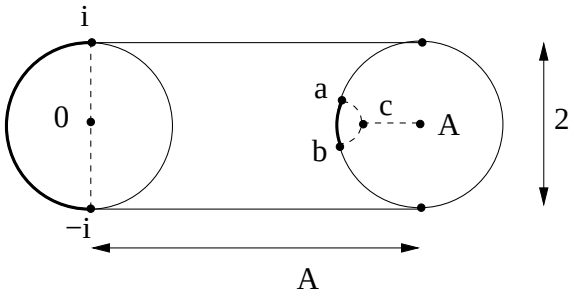


FIGURE 30. Case 2a: between parallel lines

Let $D(x_k, 1)$ be a collection of disks of radius 1 centered at points $x_k = Ak/n$ for $k = 0, \ldots mn$ and let $\tau_k : D(x_k, 1) \to D(x_{k-1}, 1)$ be the elliptic tranformation that fixes the two points where the bounaries intersect. Let $\sigma_n : D_2 \to D_1$ be the composition of these maps. The map $D_2 \to D_1$ we want is the limit of these maps as $n \to \infty$. Since this map is a Möbius tranformation and is symmetric with respect to the real axis it must map $A + 1$ to 1 and $A - 1$ to -1 and is the composition of a Euclidean translation from D_2 to D_1 followed by a hyberbolic self-map of D_1 which fixes ± 1, i.e., it is of the form

$$z \to z - A \to \mu(z - A),$$

where μ is of the form

$$\mu(z) = \frac{z - a}{1 - az},$$

for some $a \in (0, 1)$. The map μ is determined by a and a can be determined by the derivative of μ at $z = 1$ using

$$\mu'(z) = (\frac{z - a}{1 - az})' = \frac{1 - a^2}{(1 - az)^2},$$

which implies

$$\mu'(1) = \frac{1 - a^2}{(1 - a)^2} = \frac{1 + a}{1 - a},$$

or

$$a = \frac{\mu'(1) - 1}{\mu'(1) + 1}.$$

So it is enough to compute $\mu'(1)$.

By the chain rule this will be the same as $[\tau_n'(A)]^n$, so we really only need compute $\tau_n'(A)$. Clearly it suffcies to consider two disks of unit radius centered at $\pm\epsilon$ where $\epsilon = \frac{1}{2}A/n$. The elliptic map has fixed points at $\pm i\lambda = \pm i\sqrt{1-\epsilon^2}$ and maps $1+\epsilon$ to $1-\epsilon$. We can write this map as a composition of three maps,

$$z \to \nu(z) = \frac{z - i\lambda}{z + i\lambda},$$

which maps the vertices to 0 and ∞, a rotation around the origin and finally by ν^{-1}. By the chain rule the absolute value of the derivative of the elliptic map at $1+\epsilon$ is the product of the absolute values of thederivatives of the threee maps in the composition. The rotation has derivative of modulus 1, so contributes nothing. Thus

$$|\tau_n'(A)| = |\nu'(1+\epsilon)|/|\nu'(1-\epsilon)|.$$

By explicit calculation

$$\nu'(z) = \frac{2i\lambda}{(z+i\lambda)^2},$$

so

$$|\nu'(1+\epsilon)|/|\nu'(1-\epsilon)| = \frac{|1-\epsilon+i\lambda|}{|1+\epsilon+i\lambda|}.$$

By the Pythagorean theorem

$$|1-\epsilon+i\lambda| = sqrt(1-\epsilon)^2 + \lambda^2 = \sqrt{1-2\epsilon+\epsilon^2+1-\epsilon^2} = \sqrt{2-2\epsilon},$$

and similarly,

$$|1+\epsilon+i\lambda| = \sqrt{2+2\epsilon}.$$

Thus

$$|\tau_n'(A)| = \frac{1-\epsilon}{1+\epsilon} = 1 - 2\epsilon + O(\epsilon^2).$$

Therefore, since $\epsilon = \frac{1}{2}A/n$,

$$|\tau_n'(A)|^n = (1 - 2\frac{A}{2n} + O(n^{-2}))^n = e^{-A}(1+o(1)).$$

Taking the limit as $n \to \infty$ gives $\mu'(1) = e^{-A}$, which gives

$$a = \frac{1-e^{-A}}{1+e^{-A}}.$$

Of course, we could experiment with other possible values as well. We shall see in the next chapter that there is a strong reason based on hyperbolic geometry for thinking that

$$a = \frac{1 - e^{-\pi A/2}}{1 + e^{-\pi A/2}},$$

would be an even better choice. If we make this change, we will refer to the resulting tree-of-disks map as the "modified ι map".

Case 2b (edge-edge bisector, non-parallel case): The situation is shown in Figure 31. We will assume D_1 is smaller than D_2; the opposite case is handled similarly. Suppose the endpoint disks are $D(z_1, r_1)$ and $D(z_2, r_2)$ with $r_1 < r_2$ and normalize by a linear map η so that they become $D_1 = D(1, B)$ and $D_2 = D(y, yB)$ for some $B > 0, y > 1$. This can be done with the map

$$\begin{aligned} s &= r_1|z_1 - z_2|/(r_2 - r_1) \\ z_3 &= z_1 + s(z_1 - z_2)/|z_2 - z_1| \\ \eta(z) &= (z - z_3)/(z_1 - z_3). \end{aligned}$$

Thus $B = r_1/|z_1 - z_3|$, $y = |(z_2 - z_3)/(z_1 - z_3)|$. The points labeled a, b in Figure 31 are mapped to c, d respectively.

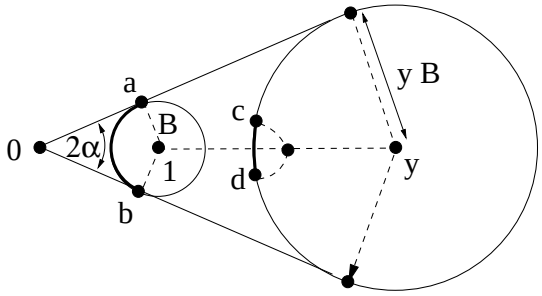


FIGURE 31. Case 2b: between non-parallel lines

As in the parallel case, the map $D_2 \to D_1$ we want is a limit of maps which are each compositions of elliptic transformations along a chain of disks running from D_2 to D_1. In this case the disks are of the form $D(x_n, r_n)$ where $x_n = y^{k/n}$ and $r_n = Bx_n$. We can also think of the desired map as the linear map $D_2 \to D_1$ given by $z \to z/y$, followed by a hyperbolic self-map of D_1 that fixes $1 \pm B$ and is determined by its derivative at $1 + B$. Again by the chain rule, this derivative is $\tau_n'(1 + B)^n$. As before,

this elliptic can be written as a composition of map sending the fixed points to $0, \infty$, a rotation and the inverse of the first map. Also as before, the rotation contributes nothing, and the deriviative is y^2/x^2 where y is the distance from $y(1+B)$ to the fixed points and x is the distance from $x_{n-1}(1+B)$ to the fixed points. See Figure 5. set $\epsilon = \frac{1}{n}\log y$. By the law of cosines

$$x^2 = B^2 + B^2 - 2B^2\cos(\frac{\pi}{2} + \alpha) = 2B^2(1+\sin\alpha),$$

and

$$y^2 = B^2 + (\epsilon + (1+\epsilon)B)^2 + 2B(\epsilon + (1+\epsilon)B)^2\sin\alpha.$$

Compute the derivative of y^2 with respect to ϵ to get

$$\frac{d}{d\epsilon}y^2 = 2(\epsilon + (1+\epsilon)B)(1+B) + 2B(1+B)\sin\alpha,$$

which at $\epsilon = 0$ becomes

$$2B(1+B)\sin\alpha,$$

so near $\epsilon = 0$,

$$y^2 = x^2 + 2B(1+B)\sin\alpha\epsilon + O(\epsilon^2).$$

This implies

$$|\tau_n'| = \frac{y^2}{x^2} = 1 + (1+\frac{1}{B})\epsilon + O(\epsilon^2).$$

Thus

$$|\tau_n'|^n = (1 + (1+\frac{1}{B})\frac{\log y}{n})^2(1+o(1)) = e^{\log y(1+\frac{1}{B})}(1+o(1)) = y\cdot y^{1/B}(1+o(1)).$$

This means the self-map of D_1 hasa derivative $y^{1/B}$ at $1+B$ and hence is given by

$$z \to \frac{1-a}{1-az},$$

where

$$a = 1 - y^{1/B}1 + y^{1/B} = 1 - e^{-\log y/B}1 + e^{-\log y/B} = 1 - e^{-A}1 + e^{-A}.$$

Thus the desired map $\tau : D_1 \to D_2$ is

$$\tau(z) = yB\sigma(\frac{1}{B}(z-1)) + y,$$

where

$$\sigma(z) = \frac{z-a}{1-az}.$$

If $r_1 > r_2$ then the mapping is almost same, except that now $0 < y < 1$, so $A = \log y/r$ is negative. Thus we define $a = -(1 - e^A)/(1 + e^A)$ and proceed as before.

As in the case of parallel edges, consideration of the hyperbolic metric will later lead us to think we should replace A by $A = \frac{\pi}{2\alpha} \log y$. If we make this change we will call the resulting map the modified ι map, as before.

Case 3 (point-line bisector): This case is pictured in Figure 32. Here we have a parabolic edge of the medial axis which is the bisector of a line segment and a point (a vertex of the polygon). There are actually two cases depending on the orientation of the picture: as we traverse the medial axis edge towards the root, the vertex can either be on our left or on our right. We will assume it is one the right, as illustrated in Figure 32, and the other case is handled similarly.

Suppose a is the vertex in question (computed from D_1 and D_2 as in (24)) and L is the line. Then L is tangent to D_1 and D_2 at points c, d given by

$$c = z_1 + r_1 e^{i\theta} \frac{z_2 - z_1}{r}, \qquad d = z_2 + r_2 e^{i\theta} \frac{z_2 - z_1}{r},$$

where $\cos\theta = (r_1 - r_2)/r$ and $r = |z_2 - z_1|$, as before. Let a^* be the reflection of a across L. This is given by

$$a^* = c + (\bar{a} - \bar{c}) \cdot (d - c)/(\bar{d} - \bar{c}).$$

Then D_1 can be mapped to D_2 by an elliptic Möbius transformation which fixes a and a^* and sends L to itself. This elliptic element rotates around a by some angle θ. We will think of this as a composition of n separate rotations, each by angle θ/n. Applying each of these rotations in turn produces a sequence of disks $\{B_k\}$ intermediate between D_1 and D_2. Let $\tau_k : B_k \to B_{k+1}$ be the elliptic map that fixes the two points of intersection $\{a, b_k\} = \partial B_k \cap \partial B_{k+1}$. The generalized crescent being considered is the limit of the union of these disks and so $\tau_n \circ \cdots \circ \tau_1$ converges to the desired $\tau : D_1 \to D_2$ as $n \to \infty$.

To understand this composition better, we will conjugate by a Möbius transformation that converts the maps τ_k into linear maps. To do this, we have to map one of the fixed points to ∞. Consider the map

$$\eta(z) = (-i)(\frac{z - a^*}{z - a})(\frac{c - a}{c - a^*}).$$

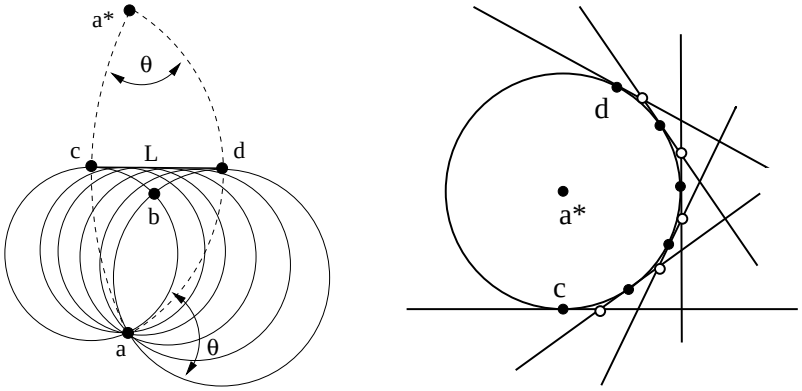


FIGURE 32. Case 3: equidistant from a point and a line. On the left is the family of circles passing through a and tangent to L. On the right we have conjugated a to ∞ and L to the unit circle. The converted τ_k maps are rotations around the white dots.

This sends $a \to \infty$, $a^* \to 0$ and $c \to -i$. Moreover, the line L maps to the unit circle and each of the disks B_k is sent to a half-plane tangent to the unit circle. The maps τ_k must therefore be sent to rotations by angle θ/n around the points where the boundaries of the kth and $(k+1)$st half-planes meet. In the limit, the composition of these maps is simply "rolling" a line counterclockwise around the unit circle, i.e., is $\sigma(z) = e^{i\theta}(z - \theta)$. Thus $\tau = \eta^{-1} \circ \sigma \circ \eta$.

If we are in the other case, when the vertex is on the left, then we define the point a using (25) instead of (24), Similarly c, d are now defined using $-\theta$ in place of θ. Then define η as before. The σ map now rolls a line clockwise around the unit circle, i.e., $\sigma(z) = (z + \theta)e^{-i\theta}$. We have now completed the proof of Theorem **??**.

Estimates of hyperbolic distance lead to a modification of ι in this case as well. Take $\sigma(z) = e^{\pm i\theta}(z \mp 2\theta)$ for the modified map.

If we put our guesses into the Schwarz-Christoffel formula we get a locally 1-1 map of the disk to a polygonal region (possibly overlapping itself) which has all the correct angles and which can be mapped to the desired region by a quasiconformal map with uniformly bounded constant and with vertices mapping to vertices. Figure 33 shows five such polygons and the corresponding regions using the ι map and modified ι map to guess the parameters. (The curved "S" shape is actually a 100-gon.) Figures **??** gives some more comparisons using the unmodified ι map.

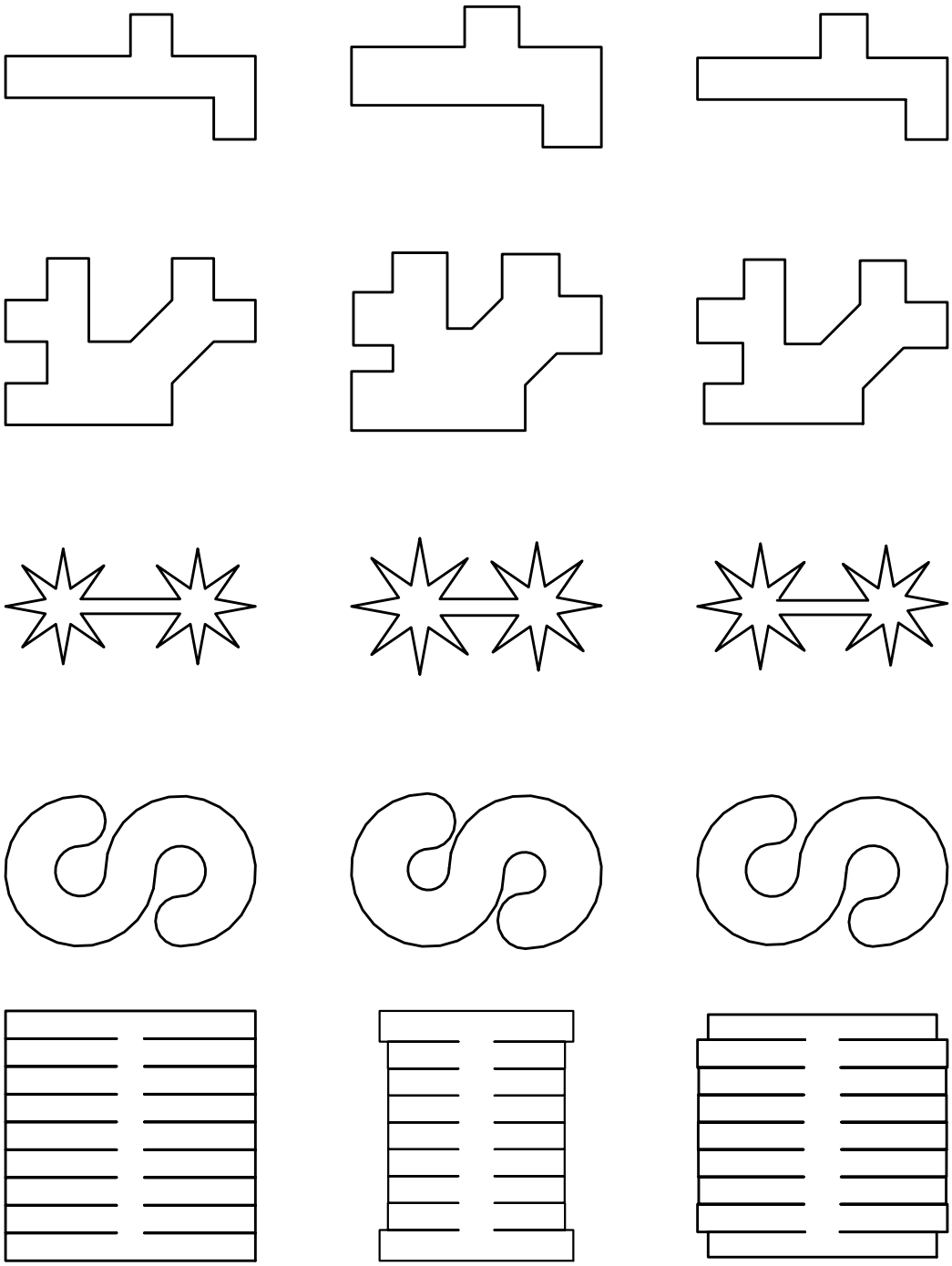


FIGURE 33.

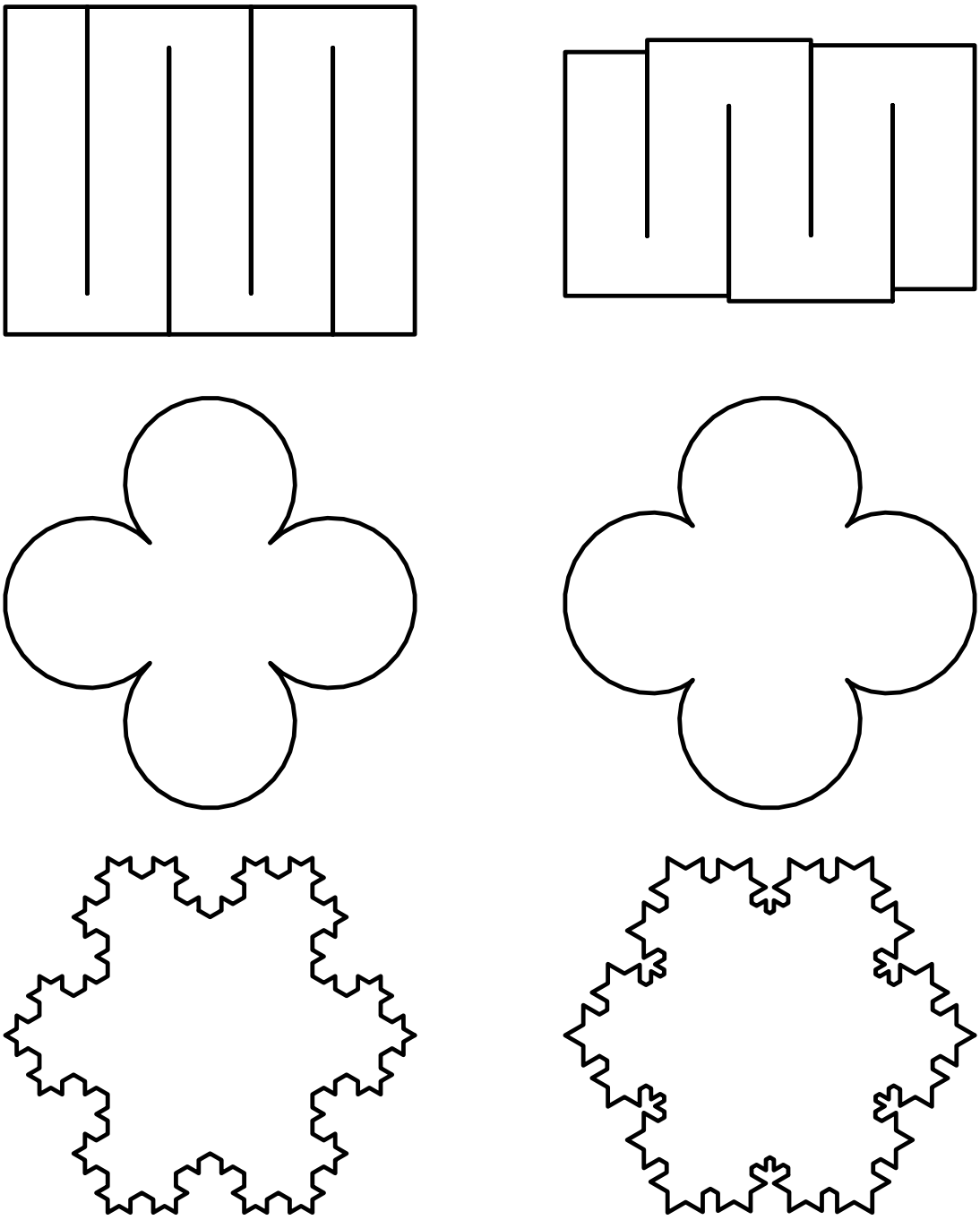


FIGURE 34.

7. ι decreases length

If we restrict the map to $\partial\Omega = \partial S_\Omega$, we have defined ι as a composition of elliptic Möbius transformations on each circular arc in $\partial\Omega$. Note that the crescents that we use are always of the form $W = D_2 \setminus D_1$ and that we are mapping the edge $\partial W \cap \partial D_2$ to the edge $\partial W \cap \partial D_1$. Thus we are in the case of the following lemma.

LEMMA 97. *Suppose Ω is a crescent which lies on one side of the line L passing through the two vertices. Let γ_1, γ_2 be the circular arcs in $\partial\Omega$ with γ_1 between γ_2 and L. If τ is an elliptic Möbius transformation fixing the two vertices and mapping γ_2 to γ_1 then $|\tau'(z)| \leq 1$ on γ_2.*

PROOF. To see this suppose $\tau(z) = (az + b)/(cz + d)$ where $ad - bc = 1$ (which we can always assume by normalizing). Then a simple calculation shows $|\tau'(z)| < 1$ iff $|1/c| < |z + d/c|$. Note that $-d/c = \tau^{-1}(\infty)$. By normalizing by a Euclidean similarity, we may assume the vertices are 1 and -1 and the crescent lies in the upper half-plane. See Figure 35. Then $-d/c$ is on the negative imaginary axis and $|\tau'(z)| < 1$ outside a circle C centered at $-d/c$ passing through -1 and 1 (since the derivative of an elliptic transformation has modulus one at the two fixed points). Let γ be the arc of this circle between 1 and -1 which lies in the upper half-plane. We claim that γ_2, the upper edge of our crescent, lies above γ.

Suppose the elliptic transformation is a rotation by θ around the points $-1, 1$. Since γ_2 and its image are both in the upper half-plane, $\theta < \pi$. Therefore $-d/c$ lies on a circle which makes angle $\pi - \theta$ with the segment $[-1, 1]$. See Figure 35. Hence the isosceles triangle with base $[-1, 1]$ and vertex $-d/c$ has two base angles of $\psi = (\pi - \theta)/2$ and the circle C makes angle $\pi/2 - \psi = \theta/2$ with $[-1, 1]$. Since γ_1 lies in $\mathbb{H}$, γ_2 makes angle of at least θ with $[-1, 1]$ and hence lies above C. □

This implies that for finite unions of disks, the map $\iota : \partial\Omega \to \partial D_0$ can be chosen to have derivative < 1 on $\partial\Omega$ (except possibly at the vertices). By taking approximations and passing to the limit we obtain:

LEMMA 98. *If Ω is simply connected Jordan domain and contains the unit disk, then we can choose ι so that $\iota : \partial\Omega \to \mathbb{T}$ does not increase the length of any boundary arc.*

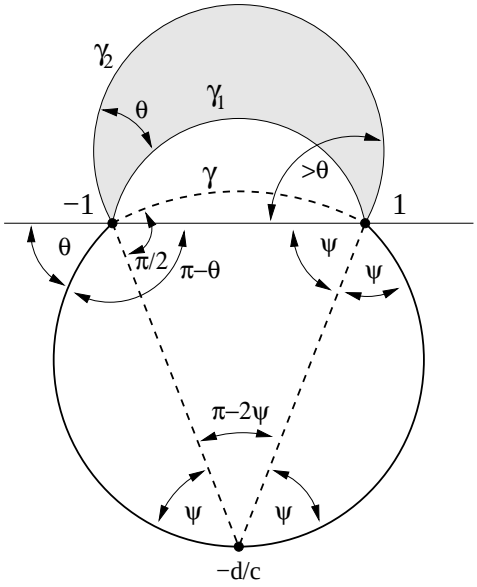


FIGURE 35. The setup in Lemma 97. We prove that γ_2 lies above γ by showing that γ makes angle $\theta/2$ with $[-1, 1]$, but γ_2 makes angle $> \theta$ with the same segment.

8. Uniform bounds for tree-of-disk maps

Earlier in this chapter we have claimed that both tree-of-disks constructions, the CRDT initial guess and the medial axis flow, give mappings of the vertices of a polygon onto the unit circle which are within a uniformly bounded quasiconformal distance of the the true conformal prevertices. In this section we will give a proof of this for the medial axis, and only indicate the additional steps that are needed for the CDRT case. For the full argument, see [**?**]. A second proof for the medial axis case will be sketched in the next section, where we describe the connection between the medial axis flow and certain ideas from hyperbolic geometry.

THEOREM 99. *There is a $C < \infty$ (independent of P) so that the ι map satisfies $d_{QC}(\iota(\mathbf{v}), \mathbf{z}) \leq C$.*

The medial axis flow is the limit of flows on domains which are finite unions of disks and where the flow from each disk to its parent is simply along the orthogonal foliation of the crescent $D^* \setminus D$. Therefore it is enough to prove the result for such domains (if we have a uniform bound, we can then pass to the limit).

Suppose we have two adjacent disks D_1, D_2. Assume these disks are distinct and the boundaries intersect at exactly two points a, b. Let γ_1 and γ_2 be the hyperbolic geodesics connecting a and b in D_1 and D_2 respectively. Then $\gamma_1 \cup \gamma_2$ forms the boundary of a crescent C in $D_1 \cup D_2$ with some angle θ. This is called the "normal crescent" since its edges are perpendicular to the boundary of $D_1 \cup D_2$. If $D_1 = D_2$ then the normal crescent has angle 0 and consists of a single geodesic arc. The disk D_1 nay be adjacent to several other disks. For each such disk form the normal crescent and remove it from D_1. These leaves several components, but only one it adjacent to all the removed crescents. This is called the gap associated to D_1. when we do this for every disk in our collection we partition our domain into a finite union of gaps and crescents; one gap for each disk and one crescent for each pair of adjacent disks. See Figure 37 for an example of this "gap/crescent" decomposition.

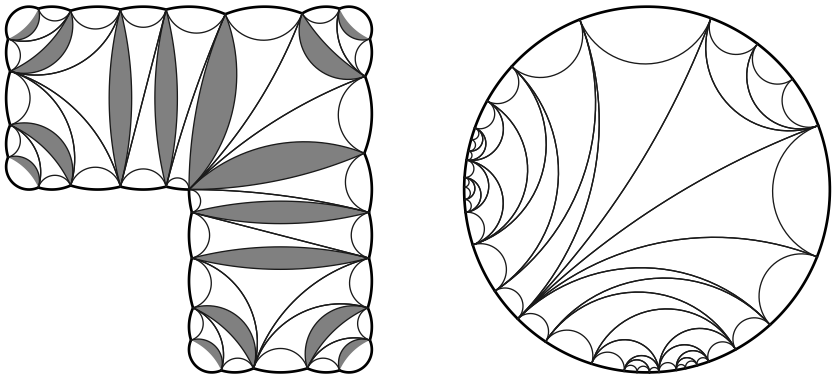


FIGURE 36. A gap/crescent decomposition for a tree of disks. When we collapse the crescents by elliptic maps the gaps come together to form a disk. This defines a continuous map φ from the tree of disks to the root disk. We have drawn arcs to indicate crescents with zero angle; this occurs when adjacent triangles have all four vertices on a single circle. We have further subdivided the gaps by adding "boundary crescents" whose edges complete the hyperbolic triangulation of the disk.

Assume that the the m disks in our tree are enumerated as $\{D_k\}$ so that D_0 is the root and so that any disk has a lower index than any of its children. Define a sequence of surfaces $\{R_k\}$ inductively by setting $R_0 = D_0$ and defining R_k by attaching the crescent $D_k^* \setminus D_k$ to R_{k-1}. Define a map $\Gamma_k : R_k \to R_{k-1}$ by extending the map $\Gamma : D_{k-1} \cup D_k \to D_{k-1}$ as the identity on the rest of R_k. Then $\Phi : \Gamma_1 \circ \cdots \circ \Gamma_m$ is

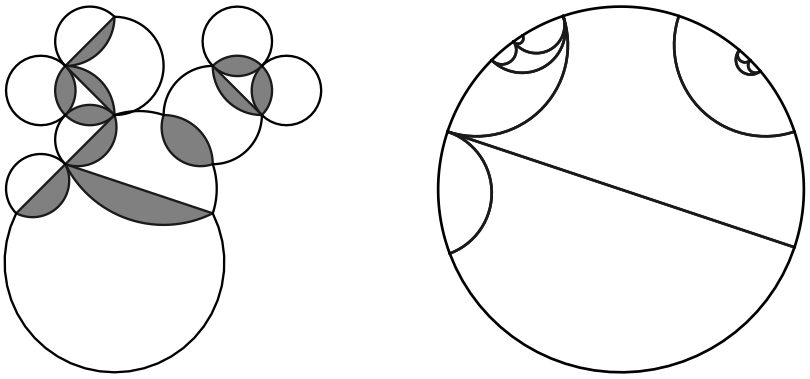


FIGURE 37. On the left is a gap/crescent decomposition for the domain corresponding to the polygon in Figure 9. On the right is the Φ image of the decomposition. Note that every crescent has been collapsed to a hyperbolic geodesic and each gap has been mapped into the disk by a Möbius transformation.

a mapping from $R = R_m$ to $D_0 = R_0$. It is equal to the identity on the gap of D_0, is equal to some Möbius transformation on the gap of every other disk and collapses each normal crescent to a hyperbolic geodesic of D_0 (these are called the bending geodesics, a term coming from hyperbolic geometry). See Figure 37. Note that Φ is not a homeomorphism, since arcs in the normal crescents are collapsed to points, but at least it does map the interior of R to the interior of D_0. We shall see that Φ is actually close to a homeomorphism in a precise sense.

LEMMA 100. *Suppose G is a gap in R and $z \in G$. Let D be the disk associated to G. Then*

$$\frac{1}{\sqrt{2}}\text{dist}(z, \partial R_p) \leq \text{dist}(z, \partial D) \leq \text{dist}(z, \partial R_p).$$

PROOF. The right hand inequality is obvious since $D \subset R$. To prove the left hand inequality, note that if $z = 0$ or if the radial projection of z is in ∂R then we have equality. Otherwise, the radial segment through z hits a geodesic γ in D bounding a crescent. Replacing z with this point of intersection decreases $\text{dist}(z, \partial D)$ more than $\text{dist}(z, \partial R_p)$ and for a point on γ, the inequality holds by a simple computation. □

On each gap G, φ is an isometry from the hyperbolic metric of the corresponding disk D to the unit disk. We have just seen that this metric (restricted to G) is

comparable to the hyperbolic on R and hence φ is Lipschitz on G. On the crescents φ is clearly Lipschitz, and thus it Lipschitz on all of R.

To prove φ is a quasi-isometry, we need to prove the opposite direction

$$\rho_{\mathbb{D}}(\varphi(z), \varphi(w)) \geq C\rho_R(z, w) - C.$$

It suffices to show that for some $\epsilon > 0$ the preimage under φ of a ball of hyperbolic radius ϵ in $\mathbb{D}$ has uniformly bounded hyperbolic diameter in R. If this is true, then any two points distance $d > \epsilon$ apart in $\mathbb{D}$ can be connected by a geodesic, and this geodesic can be cut into $\simeq d/\epsilon$ parts, each of which has preimage of diameter C. Thus the two original points have all preimages within $(C/\epsilon)d$ of each other, which is the desired lower bound.

The key step to proving this is the following.

LEMMA 101. *There is a $\epsilon_0 > 0$ and $C < \infty$ so that if $\epsilon < \epsilon_0$ then the following holds. Suppose $\{\tau_j\}_1^M$ is a finite collection of elliptic transformations such that each τ_j rotates by angle θ_j, that $\sum_j |\theta_j| \leq L$ and that each τ_j has one fixed point in $D(0, \epsilon)$ and one fixed point outside $D(0, 1/\epsilon)$. Then $|\tau_1 \circ \cdots \circ \tau_M(w) - w \exp(i \sum_j \theta_j)| \leq C\epsilon L$ for any w with $\frac{1}{2} \leq |w| \leq 2$.*

PROOF. Each elliptic transformation preserves a family of circles and by the hypothesis on the fixed points, the circles in these families of diameter between $1/4$ and 4 differ by at most angle $C\epsilon$ from circles concentric with the origin. Thus if τ_j rotates by angle θ_j it will differ from a rotation of angle θ_j around the origin by at most $C\epsilon\theta_j$. See Figure 38. Summing gives the desired result. □

Suppose $\epsilon < \epsilon_0/2$ and consider a ball B of hyperbolic radius ϵ in $\mathbb{D}$. Renormalizing by a Möbius transformation we may assume it is centered at the origin. If is it contained inside the image of one gap $G \subset D \subset R$, there is nothing to do, since the preimage has smaller size in R. Thus B must hit one or more bending geodesics (i.e., images of crescents under φ). Let $\{\gamma_j\}$ be an enumeration of these and let θ_j be the angle measure of the corresponding crescents in R. Since each of these hits B, the endpoints of all the γ's must be clustered in two balls of (Euclidean) radius δ (which tends to zero with ϵ) on opposite side of the circle. Without loss of generality we may assume this balls are centered at $-i$ and $+i$.

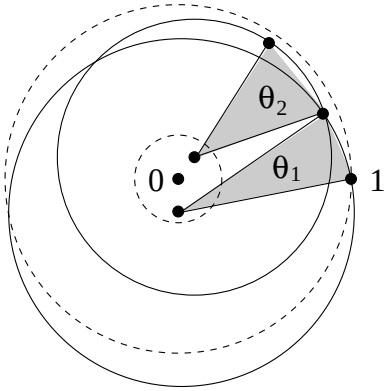


FIGURE 38. If the fixed points are ϵ-near 0 and ∞ the elliptic transformations look like Euclidean rotations up to error $C\epsilon$.

By applying a Möbius transformation taking the unit disk to the upper half-plane we can instead assume these geodesics are in $\mathbb{H}$ and each has one endpoint in $[0, \delta]$ and the other outside $[1/\delta, \infty]$. These geodesics also come with a natural left to right ordering and each has an associated elliptic Möbius transformation which rotates by angle θ_j around the endpoints of γ_j. If we apply these maps from rightmost to leftmost, letting the map act on everything to the right of the corresponding geodesic, we get the map φ on the gaps.

Let B_2 be the ball of radius $1/2$ centered at i. If $\sum_j \theta_j = L \geq 2\pi$ then Lemma 101 implies the preimage of B_2 in R covers the unit circle more than once at some point. Since our domain is planar, this is a contradiction, so we must have $L < 2\pi$. But this means that the preimage of B is a connected set in R which can be covered by a uniformly bounded number of balls of radius $1/4$ which are more than distance $1/4$ from ∂R. Each element of this cover therefore has bounded hyperbolic radius in R and hence the preimage of B had uniformly bounded radius.

Thus φ is a quasi-isometry with uniformly bounded constants and thus there is a K-quasiconformal map $R \to \mathbb{D}$ (with uniformly bounded K) that has the same boundary values. This completes the proof of Theorem **??**

The proof given here is not explicit. An argument of Epstein and Marden's in [**?**] implies a quasiconformal constant $K \approx 82.6$ and they conjectured $K_0 = K = 2$ is correct. In [**?**] it is proven that one can take $K = 7.82$. More recently Epstein and

Markovic [**?**] have shown that $K > 2.1$, by showing the K is at least this large when Ω is the complement of a certain logarithmic spiral.

The proof presented in this section generalizes to proving the initial guess of the CRDT algorithm also gives values that are within a uniform QC distance of the true prevertices. The argument is more complex however, mostly because we have to introduce a possibly non-planar Riemann surface in place of the planar, finite union of disks used above. The main steps are as follows:

(1) Given a polygon P with interior Ω, we define a Riemann surface R so that $\Omega \subset R$ in a natural way and the vertices of P lie in $\partial\Omega \cap \partial R$. We then define two maps $\Psi : \Omega \to R$ and $\Phi : R\mathbb{D}$. The map Ψ is the identity on the vertices of P and the map Φ maps them to the CRDT guesses. Thus the composition $\Phi \circ \Psi$ is a map $\Omega \to \mathbb{D}$ which extends the CRDT initial guess to the whole interior. To prove the claim we have to show this map has a uniformly bounded QC constant, so it suffices to show each of the two maps has uniformly bounded constant.

(2) We build the surface R as follows. Given the polygon, compute a rooted Delaunay triangulation and associate to each triangle its circumdisk. Assume the disks are listed as $D_0, D_1, \ldots$ with D_0 the root and each disk occurring later in the list than its parent. Build a sequence of surfaces inductively starting with $R_0 = D_0$, the root disk. For $k = 1$ we attach the crescent $D_1 \setminus D_0$ to R_0. This gives a surface which is still a planar domain, i.e., a union of two disks. At the kth step, we attach the crescent $D_k \setminus (D_k)^*$ to R_{k-1} (by assumption the parent of D_k occurred earlier in the list so the boundary of this crescent is already part of the boundary of R_{k-1}). However, this crescent is considered disjoint from R_{k-1}, and this means that R_{k-1} may not be a planar domain. $R = R_n$, where n is the number of elements in our triangulation.

(3) The map $\Psi : \Omega \to R$ is fairly easy. One each interior triangle of the triangulation, the map is the identity. On boundary triangles the map Ψ "blows up" the triangle to the union of itself and one or two crescents attached along its non-interior edges. This is completely explicit and has uniformly bounded

QC constant as long as the step one of CRDT has been done (i.e., extra vertices added to P to make sure it has a nice triangulation).

(4) Constructing the map $\Phi : R \to \mathbb{D}$, is more complicated, but completely analogous to the construction for the planar union of disks discussed in this section. Define a map $\partial R_k \to \partial R_{k-1}$ by collapsing the crescent $D_k \setminus (D_k)^*$ along the orthogonal foliation lines. Extend this map to the interior using the gap/crescent decomposition.

(5) Prove that Lemma 5 holds for the surface R. This requires some property of R that say it is not too far from being a planar domain. In [] it is shown that any curve in R which projects to a circle can cover that circle at most 3-to-1. This property, combined with an extremal length argument is used to prove the desired estimate on harmonic measure. This show that map $\Phi : R \to \mathbb{D}$ is Lipschitz, just as before. The proof of the opposite direction (the lower bound in the definition of quasi-isometry) is exactly the same as before, except now the number 2π is replaced by 6π since a circle in R can't cover a circle more than 3 times.

9. The factorization theorem and Brennan's conjecture

we have proved that any simply connected polygonal domain Ω which contains a disk of radius 1 has a quasiconformal map $\phi : \Omega \to \mathbb{D}$ so that both the quasiconformal constant is uniformly bounded and $|\varphi'| \leq 1$ on $\partial\Omega$. From this it follows that we can find another quasiconformal map (also with a uniformly bounded QC constant) so that $|\psi'|$ is bounded over all of Ω. By passing to the limit this must be true for all simply connected domains.

COROLLARY 102. *If Ω is any simply connected domain then there is a Lipschitz homeomorphism from Ω with its internal path metric to the unit disk with its usual Euclidean metric.*

Recall that a domain Ω is called quasiconvex if there is a $C < \infty$ so that any two points $x, y \in \Omega$ can be joined by a path in Ω of length at most $C|x - y|$ (i.e., the internal path metric is comparable to the Euclidean metric). The following is an immediate corollary of the previous result

COROLLARY 103. *If Ω is quasiconvex, there is a Lipschitz homeomorphism from Ω to the disk (with respect to the Euclidean metric on both domains).*

If Ω is a quasidisk, then it is quasiconvex. Moreover, in this case there is also a biLipschitz reflection across $\partial\Omega$, so it is easy to prove the following:

COROLLARY 104. *If Γ is a bounded quasicircle then there is a quasiconformal, Euclidean Lipschitz map of the plane which maps Γ to the unit circle.*

COROLLARY 105. *[The factorization theorem] There is a universal $K < \infty$ so the following holds. Suppose $f : \mathbb{D} \to \Omega$ is conformal. Then $f = g \circ h$, where $h : \mathbb{D} \to \mathbb{D}$ is a K-quasiconformal self-map of the disk and $g : \mathbb{D} \to \Omega$ is expanding in the sense that $|g'(z)| > C|f'(0)|$ for all $z \in \mathbb{D}$.*

The proof will actually show that g has the property that

$$\min_{Q} |g'(z)| \geq C \max_{T(Q)} |g'(z)|,$$

for any Carleson square Q and its top half, $T(Q)$. Thus $|g'|$ almost behaves as if it were increasing near the boundary.

This is a pretty result that quantifies the idea that a conformal map of the disk can have an unlimited amount of expansion, but only a bounded amount of contraction. Many problems and results in geometric function theory seek to quantify this various different ways. One of the most famous such problems is Brennan's conjecture.

Suppose Ω is a simply connected plane domain and $F = f^{-1} : \Omega \to \mathbb{D}$ is a conformal map. It is obvious that $\int_\Omega |F'|^2 dxdy = \text{area}(\mathbb{D}) = \pi$ so that $F' \in L^2(\Omega, dxdy)$, but it is not clear what other L^p spaces F' must belong to. Gehring and Hayman (unpublished) showed that $F' \in L^p$ for $p \in (\frac{4}{3}, 2]$ and showed the lower bound is sharp. Metzger [**?**] improved this to $p \in (\frac{4}{3}, 3)$. In 1978 James Brennan [**?**] improved this by showing one can take $p \in (\frac{4}{3}, p_0)$ for some $p_0 > 3$ and conjectured that $p_0 = 4$ is possible (this is sharp since the Koebe function mapping $\mathbb{D} \to \mathbb{C} \setminus [\frac{1}{4}, \infty)$ gives an $F' \notin L^4$). If one prefers to consider maps $f : \mathbb{D} \to \Omega$ then it is easy to check by change of variables that $f \in L^p$ it is equivalent to

$$\int_{\mathbb{D}} |f'|^{2-p} dxdy < \infty.$$

The best estimate (so far as I know) is currently due to Bertilsson [**?**], [**?**] who showed $p_0 \geq 3.422$. This is a slight improvement of the earlier result of Pommerenke [**?, ?**], that $p_0 \geq 3.399$.

In addition to its intrinsic interest, the Brennan conjecture has interesting consequences (e.g. see Section **??**) and is currently under intense investigation. Some recent papers on the Brennan conjecture include the work of Carleson and Makarov [**?**], Hurri-Syrjänen and Staples [**?**], Volberg and Zdunik [**?**]. Moreover, the Brennan conjecture is now just a special case of the more general "universal spectrum conjecture", e.g., [**?**], [**?**]. Recall Astala's recent (and remarkable) proof of the area distortion conjecture for quasiconformal maps [**?**]. One consequence of Astala's result is that if h is a K-quasiconformal map of the disk to itself, then $|h'|$ is in weak L^p where $p = 2K/(K-1)$. A function F is said to be in weak L^p if

$$\text{area}(\{z : |F(z)| > \lambda\}) \leq \frac{C}{\lambda^p}.$$

In particular h' is in every L^p space with $p < 2K/(K-1)$.

Let $f : \mathbb{D} \to \Omega$ be conformal and let $f = g \circ h$ be the factorization given by Corollary 105. If the theorem holds with constant K then $h^{-1} : \mathbb{D} \to \mathbb{D}$ is also K-quasiconformal and so by Astala's theorem, $(h^{-1})'$ is in $L^p(\mathbb{D})$ for every $2 \leq p < 2K/(K-1)$. Thus for $p > 2$ and $w = u + iv = h(x + iy)$,

$$\begin{aligned}
\int_{\mathbb{D}} |f'(z)|^{2-p} dxdy &\leq \int_{\mathbb{D}} |h'(z)|^{2-p} |g(h(z))|^{2-p} dxdy \\
&\leq |Cf'(0)|^{2-p} \int_{\mathbb{D}} |h'(z)|^{2-p} dxdy \\
&\leq |Cf'(0)|^{2-p} \int_{\mathbb{D}} |(h^{-1})'(w)|^p dudv,
\end{aligned}$$

which is finite if $p < 2K/(K-1)$. Thus, if Sullivan's theorem holds for every $K > 2$, then the Brennan conjecture is true. In order to improve Bertilsson's result we would have to prove Sullivan's theorem with

$$K = \frac{3.422}{3.422 - 2} \approx 2.4064,$$

which is much better than currently known estimates. Moreover, Marden and Markovic constructed a domain for which the quasiconformal constant of the ι map is strictly larger than 2, so we can't prove Brennan's conjecture for all domains using just this

map. On the other hand, we don't need to consider just the ι map: any locally Lipschitz, 2-quasiconformal map to the disk will do. Does one always exist?

CHAPTER 8

Domes and scaling

In the last chapter we saw that the medial axis flow from $\partial\Omega$ to ∂D (where $D \subset \Omega$ is some medial axis disk) gives a uniform approximation to the boundary values of the Riemann map $\Omega \to D$. This fact was originally discovered because of the connections between the medial axis and certain objects in 3-dimensional hyperbolic geometry and the original proof of the uniform approximation property of the medial axis flow was as a corollary to a theorem on Sullivan, Epstein and Marden about the boundaries of hyperbolic 3-manifolds. In this chapter we will explain the connection, reprove certain facts from the previous chapter from this perspective and derive a continuation method that allows one to converts a locally convergent algorithm for finding conformal maps into a globally convergent one.

1. The dome of a domain

So far we have presented the medial axis flow as a purely two dimensional object. However, its connection to conformal mappings arose from observations about three dimensional hyperbolic geometry and in order to give explicit formulas for the corresponding tree-of-disks map and to prove the uniform bounds for the resulting guesses, it is convenient to explain the connection between the medial axis and 3-dimensional geometry. We start by recalling the definition of the dome of a planar domain.

Given a closed set E in the plane, we let $C(E)$ denote the convex hull of E in the hyperbolic upper half-space, $\mathbb{R}^3_+ = \mathbb{R}^3_+$. This is the convex hull in $\mathbb{R}^3_+$ of all the infinite hyperbolic geodesics which have both endpoints in E (recall these are exactly the circular arcs in $\mathbb{R}^3_+$ which are orthogonal to $\mathbb{R}^2 = \partial\mathbb{R}^3_+$). One really needs to take the convex hull of the geodesics ending in E and not just the union of these geodesics; for example, if E consists of three points, then there are three such geodesics and these form the "boundary" of an ideal triangle whose interior is also in the convex hull of E.

The complement of $C(E)$ is a union of hyperbolic half-spaces. There is one component of $\mathbb{R}^3_+ \setminus C(E)$ for each complementary component Ω of E and this component is the union of hemispheres whose bases are disks in Ω (also include half-planes and disk complements if Ω is unbounded). For example, when E is the boundary of a square, the lower and upper boundaries of $C(E)$ are illustrated in Figure 1.

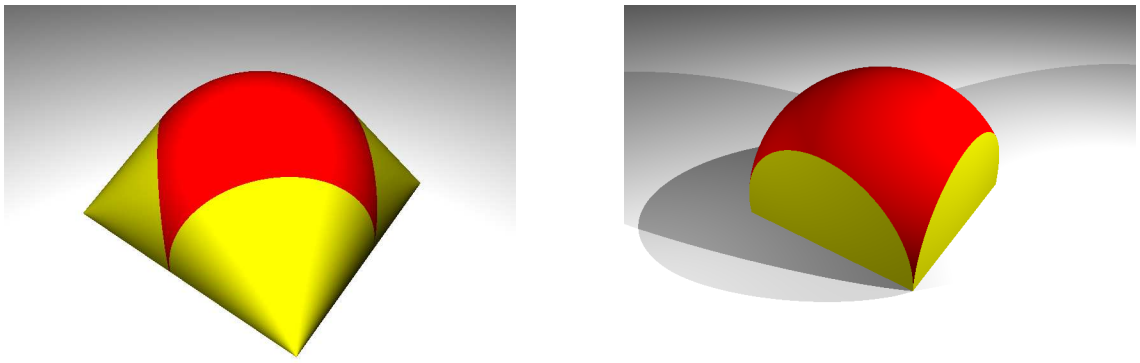


FIGURE 1. The lower and upper boundaries of the hyperbolic convex hull of the boundary of a square (left and right figures respectively). The lower boundary consists of one geodesic face (dark) and four Euclidean cones (lighter). The upper boundary has five geodesic faces (one hemisphere and four vertical). The outside of the square is a finitely bent domain, but the inside is not.

LEMMA 106. *Suppose S_Ω is the dome of a simply connected, hyperbolic plane domain Ω. Then for every $x \in S_\Omega$ there is an open hyperbolic half-space H disjoint from S_Ω so that $x \in \partial H \cap S_\Omega$. For any such half-space, $\partial H \cap S_\Omega$ contains an infinite geodesic, and its base disk (or half-plane) has boundary which hits $\partial\Omega$ in at least two points.*

PROOF. Let $W = C(\Omega^c)$ be the hyperbolic convex hull of Ω^c, so $S_\Omega = \partial W$. By definition, W is the intersection of all closed half-spaces which contain it, and from this it is easy to see that any boundary point on W is on the boundary of some closed half-space which contains W. Thus x is also on the boundary of the complementary open half-space H (which must be disjoint from W). The base of H on $\mathbb{R}^2$ is a half-plane or a disk and by conjugating by a Möbius transformation, if necessary, we assume it is the unit disk $D = \mathbb{D}$ and that H contains the point

$z = (0, 0, 1) \in S_D$. Clearly ∂D hits $\partial\Omega$ in at least one point, for otherwise its closure would be contained in another open disk in Ω, whose dome would be strictly higher than S_D, contradicting that $z \in S_D \cap S_\Omega$. In fact, ∂D must hit $\partial\Omega$ in at least two points. For suppose it only hit at one point, say $(1, 0) \in \mathbb{R}^2$. Then for $\epsilon > 0$ small enough the disk $D(-2\epsilon, 1+\epsilon)$ would also be in Ω and its dome would strictly separate z from S_Ω. Thus ∂D hits $\partial\Omega$ in at least two points and the geodesic in $\mathbb{R}^3_+$ between these points lies on the $\partial H \cap S_D$, as desired. □

Thus each point on the dome is also on the dome of a disk in Ω whose boundary hits $\partial\Omega$ in at least two points, i.e., the dome is the union of hemispheres corresponding to medial axis disks.

The dome is easiest to visualize when Ω is a finite union of disks, e.g., see Figure 2. Such a domain will be called "finitely bent" because the dome consists of a finite union of geodesic faces (each contained on a geodesic plane in $\mathbb{R}^3_+$, i.e., a Euclidean hemisphere or vertical plane) which are joined along infinite geodesics called the bending geodesics. Such a dome is the analog of a convex, unbounded, polyhedral surface in Euclidean space.

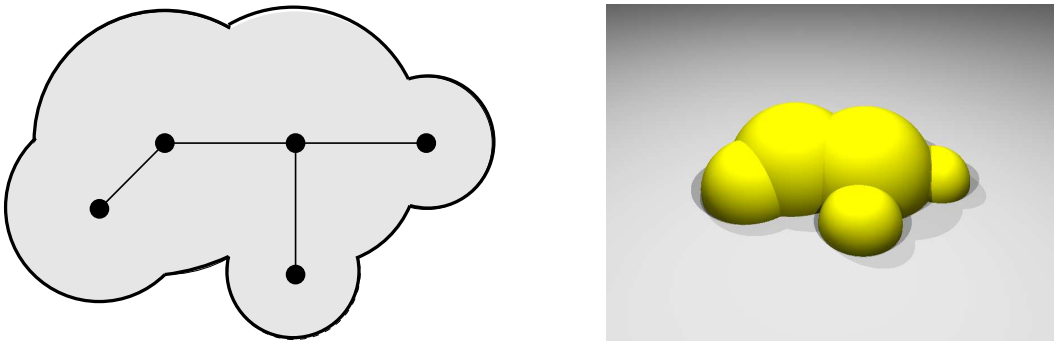


FIGURE 2. A finitely bent domain, its medial axis and its dome

When we are given a finitely bent domain Ω we shall always assume we are given a complete list of disks in Ω whose boundaries hit $\partial\Omega$ in at least three points. Then every face of the dome corresponds to a hemisphere that has one of these disks as its base. This is slightly different than just giving a list of disks whose union is Ω; in Figure 3 we show a domain which is a union of four disks $\Omega = D(1, 1) \cup D(i, 1) \cup D(-1, 1) \cup D(-i, 1)$ but which contains a fifth disk, $D(0, \sqrt{2})$ which also corresponds to a face on the dome of Ω.

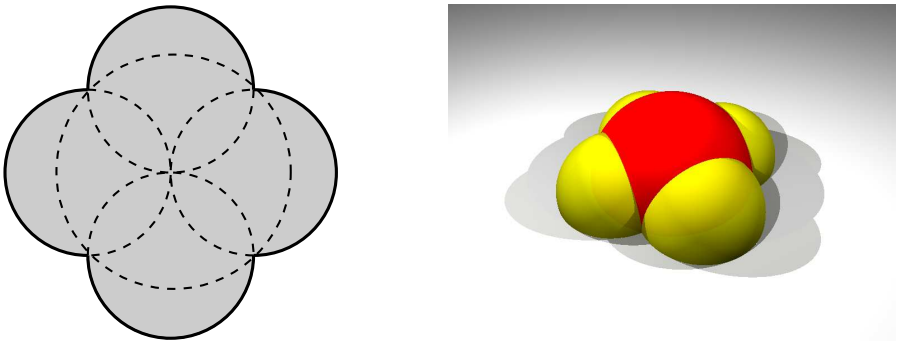


FIGURE 3. A domain which is a union of four disks, but which has five faces on the dome because of a "hidden" maximal disk.

The faces of the dome of a finitely bent domain form the vertices of a finite tree, with adjacency defined by having an infinite geodesic edge in common. This induces a tree structure on the maximal disks in the base domain: disks which hit exactly two boundary points are interior points of edges of the tree and disks which it three or more points are the vertices (of degree 1 or ≥ 3). A "hidden" maximal disk (i.e., one that corresponds to some face of the dome, but which does not share a boundary arc with Ω) must have degree at least three in this tree.

LEMMA 107. *For any tree the number of vertices of degree three or greater is less than the number of degree one vertices.*

The proof is easy and left to the reader (remove a degree one vertex and use induction). So if Ω can be written as a union of n disks in any way, there are at most $2n$ vertices of the medial axis.

To illustrate these ideas we show a few polygons, along with their medial axes and their domes. The dome of a polygon is naturally divided into kinds of pieces: (1) a hyperbolic geodesic face corresponding to a vertex of the medial axis of degree three or more (2) a cylinder or cone corresponding to sweeping a hemisphere along a bisector of two edges or (3) sweeping a hemisphere along the parabolic arc of a point-edge bisector. Disks corresponding to the interiors of point-point bisector edges do not contribute to the dome since the union of the two disks at the endpoints of this edge contain all the disks corresponding to the interior points.

In the dome of a convex polygon, only the first two types of pieces can occur. These are illustrated in Figure 4. The third type of medial axis arc can occur in non-convex domains, as illustrated in the polygonal "corner" in Figure 5.

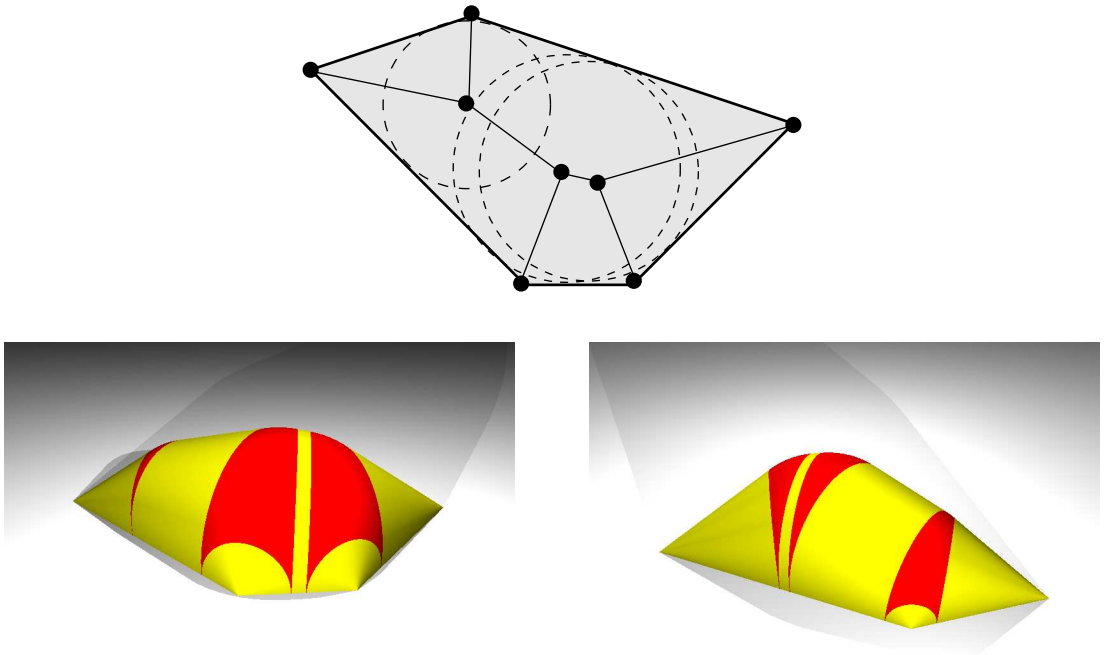


FIGURE 4. The medial axis and dome of a convex region. This dome has three geodesic faces which are shaded darker (these correspond to vertices of the medial axis); the lighter parts of the dome are Euclidean cones which correspond to edges of the medial axis. The dome is shown from two different directions

The medial axis also suggests a way of approximating any domain by a finite union of disks; simply take a finite subset of the medial axis so that the corresponding union of medial axis disks is connected. The medial axis of such a union consists of one vertex for each geodesic face in the dome and straight lines connecting the vertices corresponding to adjacent faces. A polygon, its medial axis and a finitely bent approximation are shown in Figure 6. In Figure 7 we show the domes of the polygon and its approximation.

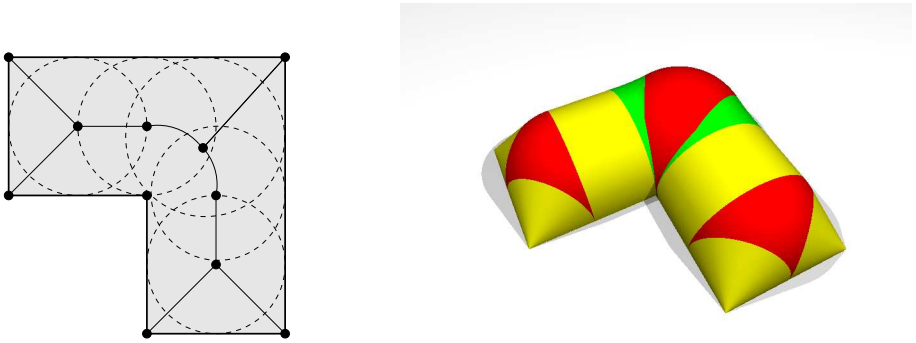


FIGURE 5. The dome of a "corner". The darkest shading are geodesic faces (vertices of the medial axis); the lightest are Euclidean cones or cylinders (edge-edge bisectors in the medial axis). The medium shading illustrates the third type of medial axis edge that can occur: the parabolic bisector of a point and a line.

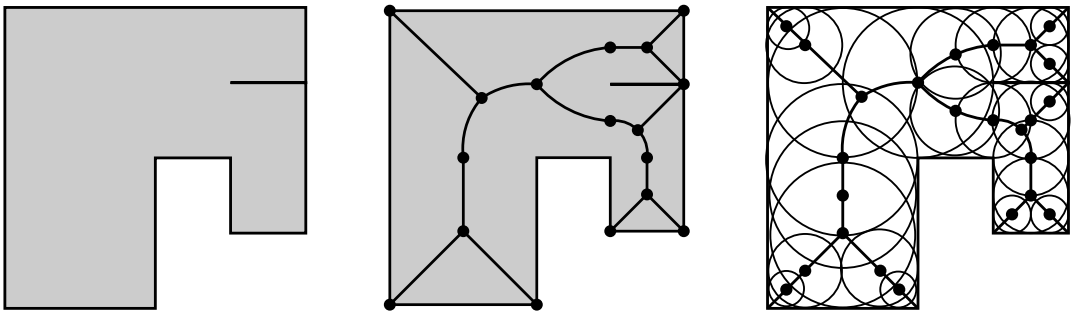


FIGURE 6. A (non-simple) polygon, its medial axis and a finitely bent approximation.

2. The Sullivan-Epstein-Marden theorem

The two main results about the dome of Ω say that (1) it is isometric to the hyperbolic disk and (2) it is "almost isometric" to the base domain Ω. More precisely, equip the dome with the hyperbolic path metric ρ_S (shortest hyperbolic length of a path connecting two points and staying on the surface).

THEOREM 108 (Thurston, [**?**]). *Suppose Ω is a simply connected plane domain (other than the whole plane or the complement of a circular arc) and let S be its dome. Then (S, ρ_S) is isometric to the hyperbolic unit disk. We will denote the isometry by $\iota : S \to \mathbb{D}$.*

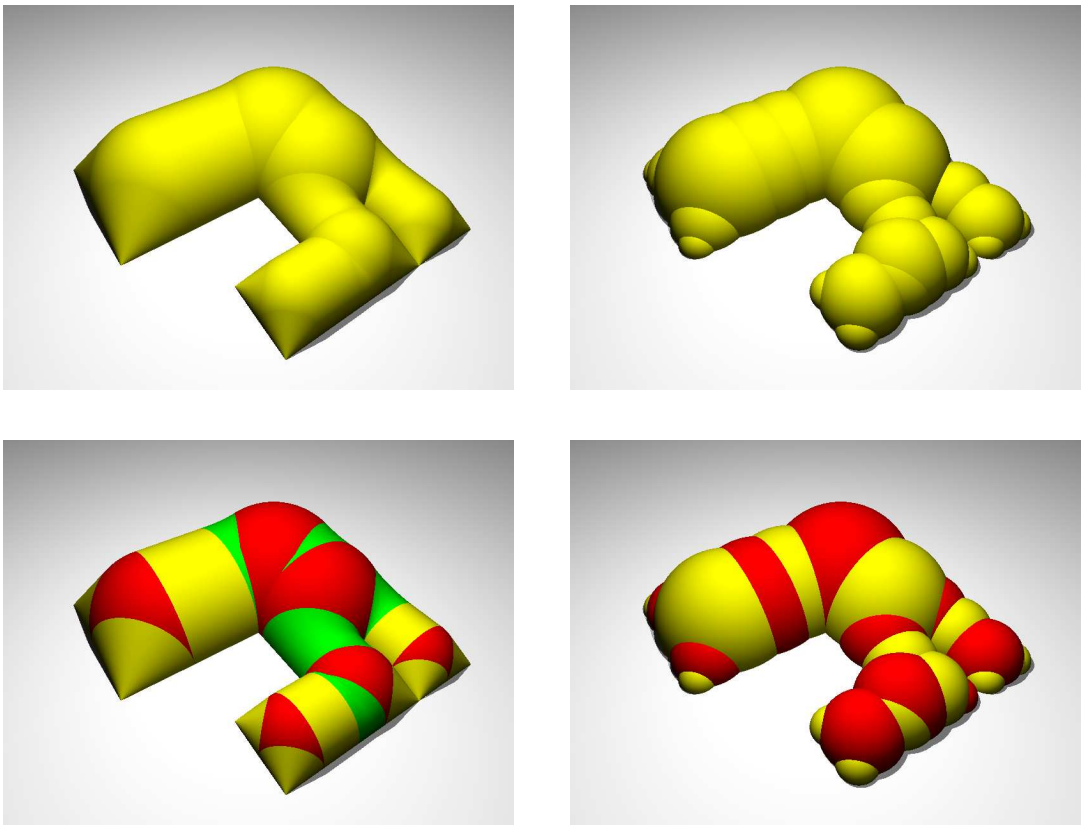


FIGURE 7. On the top left is the dome of the polygon P_2 and on the top right is the dome of the finitely bent approximation Ω_2. Below each, we have redrawn the domes, but with different sections shaded differently. For P_2, regions corresponding to different edges of the medial axis colored differently. For Ω_2 the dome is a union of geodesic faces (which form the vertices of a tree) and adjacent faces are shaded in alternating colors.

THEOREM 109 (Sullivan [**?**], Epstein-Marden [**?**]). *Suppose Ω is a simply connected plane domain (other than than the whole plane or the complement of a circular arc). There is a K-quasiconformal map $\sigma : \Omega \to S$ which extends continuously to the identity on the boundary (K is independent of Ω).*

In fact, there is a biLipschitz map between Ω and its dome (each with their hyperbolic metric; see Theorem 93), but we will only use the quasiconformal version of the result. We place the additional restriction that Ω is not the complement of a circular arc because in that case the convex hull of $\partial\Omega$ is a hyperbolic half-plane and

the dome should be interpreted as two copies of this half-plane joined along its edge with bending angle π. In order to simplify the discussion here, we simply omit this case (with the correct interpretations the results above still hold in this case; this is discussed in complete detail in Section 5 of [**?**]).

Both these theorems have their origin in the theory hyperbolic of 3-manifolds. Such a manifold M is a quotient of the hyperbolic half-space, $\mathbb{R}^3_+$, by a discrete group G of isometries. The orbit of any point under this group accumulates only on the boundary of the half-space and the accumulation set (which is independent of the orbit except in trivial cases) is called the limit set Λ. The complement Ω of Λ in the boundary of hyperbolic space is called the ordinary set. The group G acts discontinuously on Ω and $\partial_\infty M = \Omega/G$ is called the "boundary at infinity" of M. This is a Riemann surface (possibly with branch points). The manifold M contains closed geodesics and the closed convex hull of these is called the convex core of M and denoted $C(M)$. The lift of the convex core to $\mathbb{R}^3_+$ is the hyperbolic convex hull of the limit set and its boundary is the dome of the ordinary set. Thus $\partial C(M)$ is just the quotient of this dome by the group G. Theorem 108 implies that the boundary of $C(M)$ is a surface of constant negative curvature, i.e., is isomorphic to the hyperbolic disk modulo a group of isometries. Theorem 109 says that $\partial_\infty M$ and $\partial C(M)$ are homeomorphic, indeed, are biLipschitz (and hence quasiconformal) images of each other with respect to their hyperbolic metrics. This fact was needed in the proof of Thurston's hyperbolization theorem for 3-manifolds that fiber over the circle. The proof of Theorem 108 for finitely bent domains simply consists of observing that if we deform the dome by bending it along a bending geodesic, we don't change the path metric at all. Moreover, a finite number of such deformations converts a finitely bent dome into a hemisphere, and this is obviously isomorphic to the hyperbolic disk. More precisely, we are using the following simple lemma.

LEMMA 110. *Suppose two surfaces S_1, S_2 in $\mathbb{R}^3_+$ are joined along a infinite hyperbolic geodesic and suppose σ is an elliptic Möbius transformation of $\mathbb{R}^3_+$ which fixes this geodesic. Then a map to another surface which equals the identity on S_1 and equals σ on S_2 is an isometry between the path metric on $S_1 \cup S_2$ and the path metric on the image.*

PROOF. This becomes obvious is one normalizes so that the geodesic in question becomes a vertical line and σ becomes a (Euclidean) rotation around it, since it is then clear that the length of any path is left unchanged. □

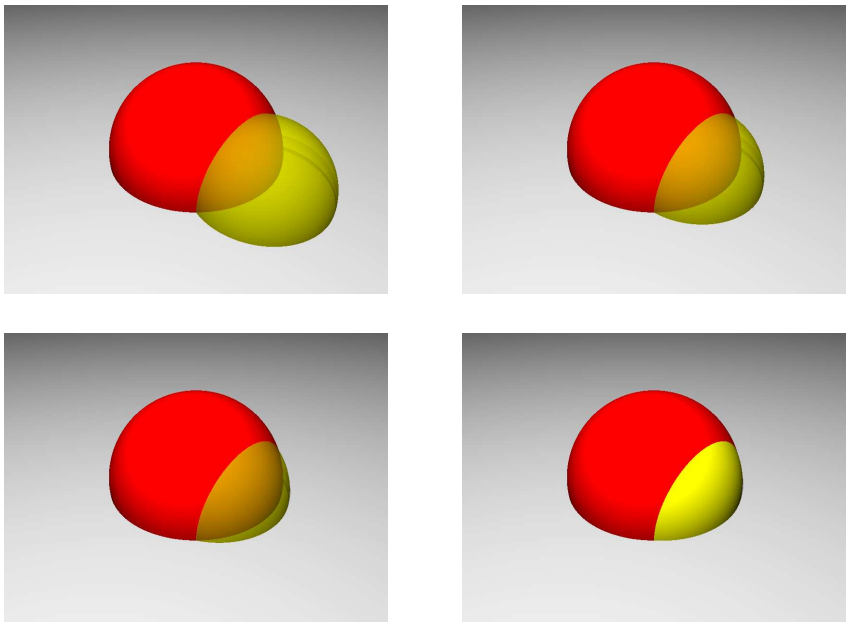


FIGURE 8. A dome consisting of two geodesic faces joined along an infinite geodesic. By bending the dome along the geodesic we get a one-parameter, isometric family of surfaces ending with a hemisphere, which is obviously isometric to the hyperbolic disk.

Theorem 108 then follows by taking a finitely bent surface and "unbending" it one geodesic at a time, i.e., we can map it to a hemisphere by a series of maps, each of which is an isometry by the lemma. Since a hemisphere is isometric to the disk, we are done. In Figure 2 we illustrate the bending along a geodesic for a dome with two faces.

This proof gives us a geometric interpretation of the map $\iota : \partial\Omega \to \partial\mathbb{D}$. The disks making up a finitely bent domain have a tree structure and if Ω is finitely bent then we fix a root disk D_0 and write $\Omega = D_0 \cup_j D_j \setminus D_j^*$, where D_j^* denotes the parent disk of D_j. This gives $\Omega \setminus D_0$ as a union of crescents. See Figure **??**. We call these "tangential" crescents since one edge of the crescent follows $\partial\Omega$ near each vertex (and to differentiate them from the "normal" crescents we will introduce later).

Each crescent in the tangential crescent decomposition has an "inner edge" (the one in the boundary of D_j^*) and an "outer edge" (the other one) and there is a unique elliptic Möbius transformation which maps the outer edge to the inner one, fixing the two vertices of the crescent (this is just the restriction to the plane of the Möbius transformation of $\mathbb{R}^3_+$ which removes the bending along the corresponding bending geodesic). The map $\iota : \partial\Omega \to \partial\mathbb{D}$ is the composition of these maps along a path of crescents which connects an arc on $\partial\Omega$ to an arc on $\partial\mathbb{D}$. An alternate way to think of this is to foliate each crescent $D_j \setminus D_j^*$ by circular arcs which are orthogonal to both boundary arcs. This gives a foliation of $\Omega \setminus D_0$ by piecewise circular curves which connect $x \in \partial\Omega$ to $\iota(x) \in \partial\mathbb{D}$. This is exactly the same as the medial axis flow for the base domain. Thus restricted to $\partial\Omega = \partial S$ the ι map and the medial axis flow map are identical.

Theorem 109 implies that the mapping $\iota : \partial\Omega \to \partial\mathbb{D}$ has a quasiconformal extension to a map $\Omega \to \mathbb{D}$ which is K-quasiconformal with a bound K which is independent of Ω. This is the same as Theorem **??**, indeed, this is way this result was first proved.

3. The retraction map onto the dome

We need to show that there is a quasi-isometry from the hyperbolic metric on Ω to the hyperbolic path metric on its dome S which extends continuously to the identity on the boundary. It turns out that there is an "obvious" map $R : \Omega \to S$; the nearest point retraction. Recall that S is the boundary of a convex region in $\mathbb{R}^3_+$. Thus every point in its complement has a nearest point on S (just think of expanding a hyperbolic ball around a point until the first time it hits S). This map extends continuously to Ω: for each $z \in \Omega$ expand a ball in $\mathbb{R}^3_+$ tangent to the boundary at until it first hits S at a point $R(z)$. This point is both on the tangent ball B at z and on a hemisphere with base in Ω which contains B. Since B can only hit the boundary of the hemisphere at one point, it can only hit S at one point, and so the point of first contact is unique.

We claim that $R : \Omega \to S$ is a quasi-isometry with constants that are independent of Ω. We will actually show that

$$\frac{1}{A}\rho_\Omega(w,z) - B\rho_S(R(z),R(w)) \leq A\rho_\Omega(z,w),$$

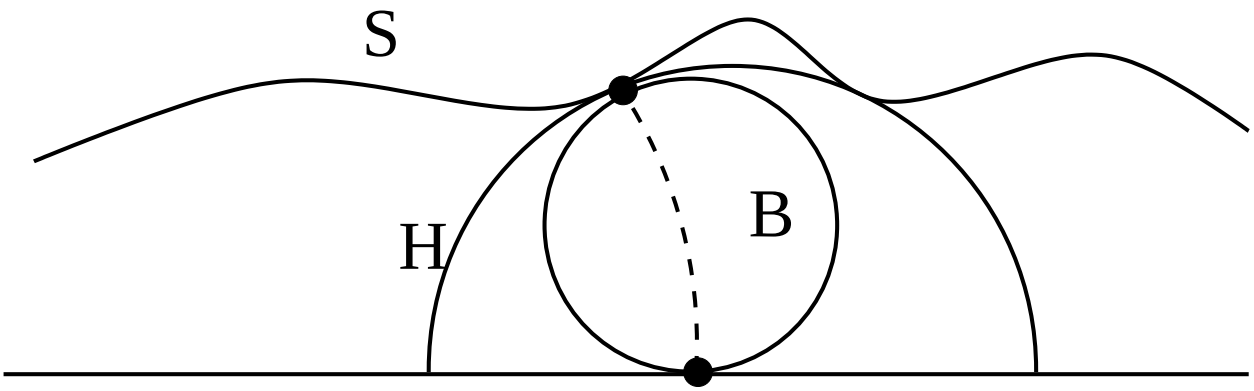


FIGURE 9. A horoball can only touch a hemisphere at one point and hence can touch the doem at only one point.

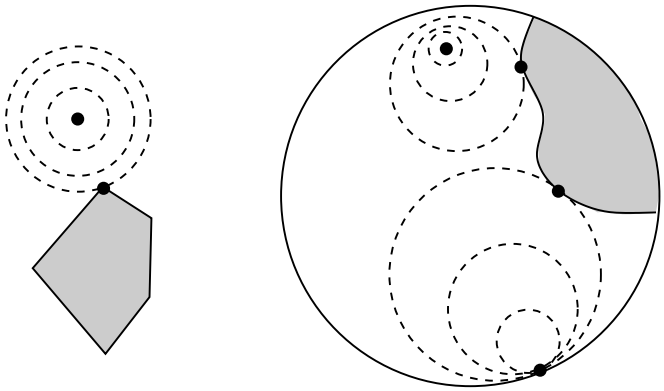


FIGURE 10. On the left we show the nearest point retraction in Euclidean space: we expand a Euclidean ball around a point until it first hits the set. If the set it convex, then this first contact point is unique. On the right is the analogous situation for hyperbolic space. For a point inside hyperbolic space we taking expanding hyperbolic balls until they hit the desired set. The natural extension to the boundary is to take expanding horoballs (balls tangent to the boundary).

so that the retraction map is Lipschtitz (hence contiuous). To prove the upper bound, we claim it suffices to show that

$$\rho_\Omega(z,w) \le \epsilon \quad \Rightarrow \quad \rho_S(R(z),R(w)) \le A\epsilon. \tag{26}$$

If this is true, then given any pair of points we can connect them by a geodesic and choose points $z = z_1, z_2, \ldots, z_{n+1} = w$ so that each is less than distance ϵ from the

next. Then

$$\sum_{k=1}^{n} \rho_{\mathbb{H}}(R(z_k), R(z_{k+1})) \leq A \sum_{k} \rho_{\Omega}(z_k, z_{k+1}) \leq A\rho(z, w),$$

and the left hand side gives an upper bound for the path distance between $R(z)$ and $R(w)$ on S as $\epsilon \to 0$. For the lower bound it also suffices to check that

$$\rho_S(R(z), R(w)) \leq 1 \quad \Rightarrow \quad \rho_{\Omega}(z, w) \leq B.$$

Similarly, to prove the lower bound it suffice to prove

$$\rho_S(R(z), R(w)) \leq 1 \quad \Rightarrow \quad \rho_{\Omega}(z, w) \leq B. \tag{27}$$

If this holds and $R(z)$ and $R(w)$ between n and $n+_1$ apart with respect to the path metric on S, then we can connect $R(z)$ and $R(w)$ by a geodesic and divide this path into at most $n+1$ segments of length $\leq$. By (27), the preimage of each segment has length $\leq B$ and so the distance between z and w is at most $B(n+1) \leq (B+1)\rho_S(R(z), R(w))$. Thus

$$\frac{1}{B+1}\rho_{\Omega}(z, w) - \frac{1}{B+1} \leq \rho_S(R(z), R(w)),$$

as desired.

We start with some easy inequalities. We normalize so that $\infty \notin \Omega$. Since W is a union of hemispheres, its complement (the region "above" S) is an intersection of hyperbolic half-spaces and hence is hyperbolically convex. Thus the geodesic from any point of S to ∞ is in the complement of W, i.e., the vertical ray from any point of S to ∞ does not hit W. Thus S is a graph.

Let $z \in \Omega$ and let $R(z) \in S$ be its image under the nearest point retraction. Let r be the radius of the horoball B that is tangent to $\mathbb{R}^2$ at z and hits S at the point $R(z)$. Because $\infty \notin \Omega$ and because the region above the dome is hyperbolically convex, the vertical ray from any point of the dome to ∞ must lie above the dome. Thus $R(z)$ must be on the top half of B. Thus

$$r \leq \operatorname{dist}(R(z), \mathbb{R}^2) \leq 2r. \tag{28}$$

Similarly,

$$\sqrt{2}r \leq \operatorname{dist}(R(z), \mathbb{R}^2) \leq 2r. \tag{29}$$

If $\text{dist}(z, \partial Omega) < r$, then the vertical line from the nearest boundary point to z must hit B, which contradicits the convesity of the dome complement. Hence $r \leq \text{dist}(z, \partial Omega)$. Finally, if $\text{dist}(z, \partial\Omega) > 2r$, then the hemisphere for the disk centered at z with this radius separates B from S, which is impossible. Thus

$$r \leq \text{dist}(z, \partial\Omega) \leq 2r. \tag{30}$$

In particular,

$$r \simeq \text{dist}(z, \partial\Omega) \simeq \text{dist}(R(z), \mathbb{R}^2) \simeq \text{dist}(R(z), \mathbb{R}^2).$$

Now suppose $z, w \in \Omega$ and $\rho_S(z, w) \leq 1$. This means $\rho_{\mathbb{R}^3_+}(z, w) \leq 1$ as well. Suppose $\text{dist}(R(z), \mathbb{R}^2) = r$ and let γ be the geodesic from z to w. Each point of γ is within distance 1 of $R(z)$ in the path metric on S and hence also in the hyperbolic metric on $\mathbb{R}^3_+$ and hence the geodesic stays between heights r/e and er above $\mathbb{R}^2$. That means that $E = R^{-1}(\gamma) \subset \Omega$ is a connected set so that every point $x \in E$ satisfies

$$\text{dist}(x, \partial\Omega) \simeq \text{dist}(x, R(z)) \simeq r.$$

Therefore E can be covered by a connected union of Euclidean squares Q from a grid so that

$$\text{diam}(Q) \simeq \text{dist}(Q, \partial\Omega) \simeq \text{dist}(Q, R(z)).$$

However, there can only be $O(1)$ such squares and each has uniformly bounded hyperbolic diameter. Thus E has uniformly bounded diameter, which is (**??**).

To prove (**??**), fix any point $z \in \Omega$ and a small $\epsilon > 0$. By applying a Möbius transformation we can assume $z = 0$ and $R(z) = p = (0, 0, 1)$. This means that the unit disk, $\mathbb{D}$, is contained in Ω but $\partial\mathbb{D}$ hits $\partial\Omega$. Let H be the hyperbolic half-space (i.e., Euclidean hemisphere) with base $\mathbb{D}$. Let $w \in \Omega$ and assume the quasi-hyperbolic distance from w to z is $\leq \epsilon$. This means that $|z - w| \leq \epsilon$ as well. Let B be the maximal horoball in $\mathbb{R}^3_+$ which is tangent to $\mathbb{R}^2$ at w and which is contained in W, i.e., B touches S at the point $R(w)$. Clearly $R(w) \in V = \partial B \setminus H$. How large is V? V is a spherical cap whose diameter is attained along the vertical Euclidean plane passing through 0 and w. See Figure 5 for a picture of this slice.

In this slice, H becomes the unit half-disk in the upper half-plane and B becomes a disk of radius r tangent to $\mathbb{R}$ at $\epsilon > 0$. Then $\partial Hand\partial B$ intersect at two points a, b as labeled. Note that the perpendicular bisector of $[a, b]$ is a diameter of B and

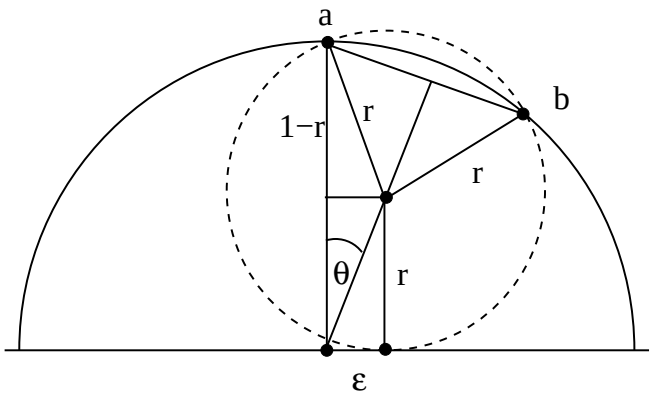


FIGURE 11. Proof that the nearest point retraction is Lipschitz.

passes through both the center of B and the origin. The arc of the unit circle between these points has length 2θ where $\tan\theta = \epsilon/r$. We can compute r by applying the Pythagorean theorem to the shaded right triangle in Figure 5 to write

$$(1-r)^2 + \epsilon^2 = r^2,$$

which implies $r = \frac{1}{2}(1-\epsilon^2)$. Thus

$$\theta \simeq \arctan(\epsilon/r) \simeq \epsilon/r = \frac{2\epsilon}{1-\epsilon^2}.$$

Thus $|a-b|/\epsilon \to 4$ as $\epsilon \to 0$.

Therefore $\rho_{\mathbb{R}^3_+}(R(z), R(z)) \leq (4+o(1))\tilde{\rho}_\Omega(z,w)$, so the retraction map is Lipschitz with respect to the quasi-hyperbolic metric on Ω. But the quasi-hyperbolic and hyperbolic metrics are boundedly equivalent by the Koebe $\frac{1}{4}$-theorem, so the retraction map is also Lipschitz with respect to the hyperbolic metric.

4. The gap-crescent decomposition for finitely bent domains

Now suppose Ω is a finitely bent domain. Then the dome S of Ω is a finite union of geodesic faces. On the interior of each face the retraction map has a well defined inverse and the images of the faces under R^{-1} are called the "gaps". The inverse images of the bending geodesics are crescents which separate the gaps. These are called "normal crescents" since their two boundary arcs are perpendicular to the two arcs of $\partial\Omega$ which meet at the common vertex. Therefore, we will call this decomposition of Ω the "normal crescent decomposition". Refer back to Figure **??**; that picture shows

a polygon, a finitely bent approximation, the normal crescent decomposition and the dome. See Figure 12 for more examples of gap/crescent decompositions.

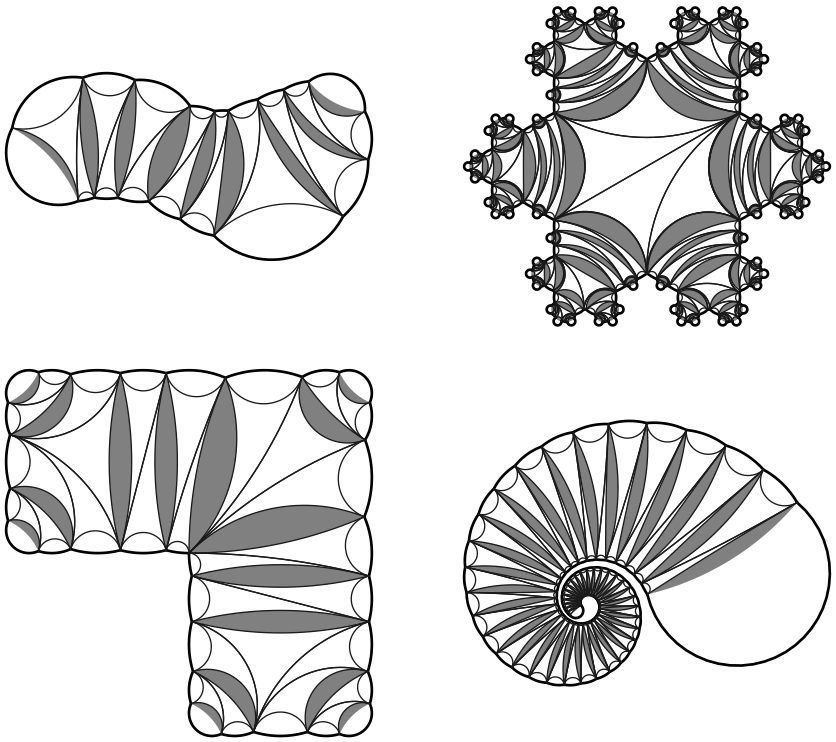


FIGURE 12. Normal crescent decompositions for some finitely bent domains. Also drawn are edges triangulating the gaps. These are added to make the bending lamination complete (see Section **??**).

If a gap G corresponds to a face $F \subset S$ then $G \subset D$, the disk in Ω which is the base of the hyperplane containing the face F. We will call D the "base disk" of G Moreover, G is the hyperbolic convex hull in D of the set where F meets $\partial\Omega$. The angle of a normal crescent C is the same as the angle made by the faces of the dome which meet at the corresponding bending geodesic. C is foliated by circular arcs which are orthogonal to both boundary arcs and each of these arcs is collapsed to single point by R. Thus for a finitely bent domain Ω, R will never be a homeomorphism (unless Ω is a disk).

The two vertices of each normal crescent are also the vertices of a crescent in the tangential crescent decomposition of Ω. Moreover, corresponding crescents from the

two decompositions have the same angle, and hence are simply images of each other by a $\pi/2$ elliptic rotation around the two common vertices. See Figure 13. Collapsing the two types of crescents simply gives the two different continuous extensions to the interior of the same map on the boundary (namely ι).

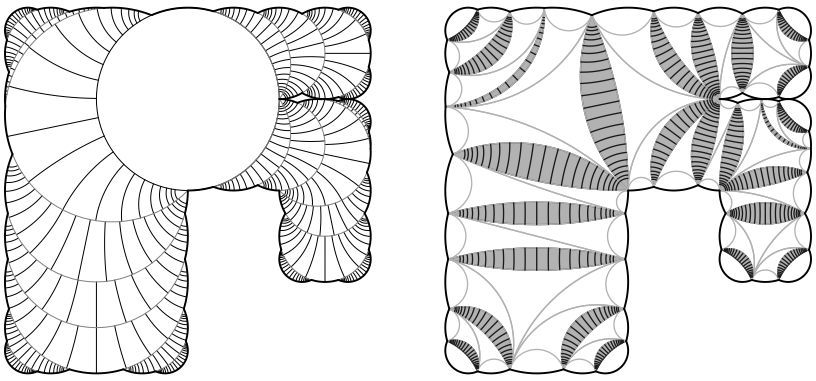


FIGURE 13. The tangential and normal crescent decomposition for a domain. There is a 1-to-1 correspondence between crescents in the two pictures; corresponding crescents have the same vertices and same angle, but are "rotated" by $\pi/2$.

The normal crescents can be defined without referring to the dome at all. Whenever we have a tree of overlaping disks we get a crescent corresponding to two adjacent disks by taking the hyperbolic geodesics in each disk which connect the two intesection points of the boundaries. After we remove the normal crescents, the remaining components are the gaps, one gap associated to each disk.

Both decompositions cut Ω into a "disk" and a union of crescents. In the tangential decomposition, it is a single connected disk, but in the normal decomposition the disk itself is broken into pieces called the gaps. The map $\varphi = \iota \circ R : \Omega \to S \to \mathbb{D}$ is Möbius on each gap and collapses every crescent to a hyperbolic geodesic in $\mathbb{D}$, thus the disk is written as a union of Möbius images of gaps. For example, see Figures 14. The picture on the left shows a normal crescent decomposition of a square and on the right are the φ images of the gaps in the disk. The images of the crescents is a finite union of geodesics which is called the "bending lamination" of Ω. If we record the angle of each crescent and assign it to the corresponding geodesic in the bending lamination, then we get a "measured lamination", and this data is enough to recover Ω, up to a Möbius image. We will discuss laminations further in Section **??**.

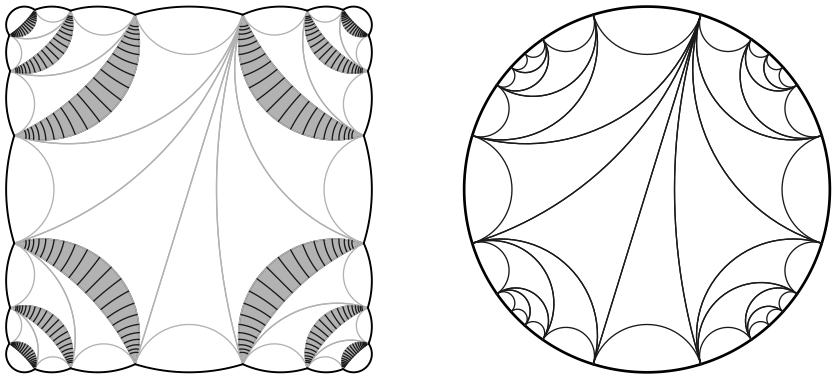


FIGURE 14. A normal crescent decomposition of a square and the corresponding bending lamination in the disk. We can recover the decomposition from the lamination by "thickening" each geodesics to a crescent of the correct angle.

5. Angle scaling

We can recover the normal crescent decomposition from the bending lamination by "thickening" each bending geodesic to a crescent of the correct angle, and moving the gaps by the corresponding elliptic transformations. If we do this continuously, we obtain a family of domains connecting the disk to Ω. For $0 \leq t \leq 1$, let Ω_t be the domain obtained by replacing a crescent of angle α in the normal decomposition by a crescent or angle $t\alpha$. (Note we would get the same domain if we replaced crescents in the tangential decomposition). See Figures 15 to 18 for some examples of these 1-parameter families. In general, the intermediate domains need not be planar, but we can think of them as Riemann surfaces which are constructed by gluing together crescents and gaps of given sizes along their edges. Figure 18 shows an example where the intermediate domains are not planar (one sees some small overlap for parameter value $t = .99$; bigger overlaps could be produced by other examples).

Given a pair of domains Ω_s, Ω_t with $0 \leq s < t \leq 1$, let $\iota_{s,t} : \partial\Omega_t \to \partial\Omega_s$ be the obvious boundary map obtained multiplying the angle of each crescent by s/t. We will extend this boundary map to the interiors by writing each crescent C in Ω_t of angle α as a union of crescents C_1, of angle $\alpha s/t$ and C_2, of angle $\alpha(1 - s/t)$. On C_1 we collapse each leaf of the E-foliation to a point (hence C_1 is maps to a circular arc) and we let our map be Möbius on C_2. By continuity, this Möbius transformation

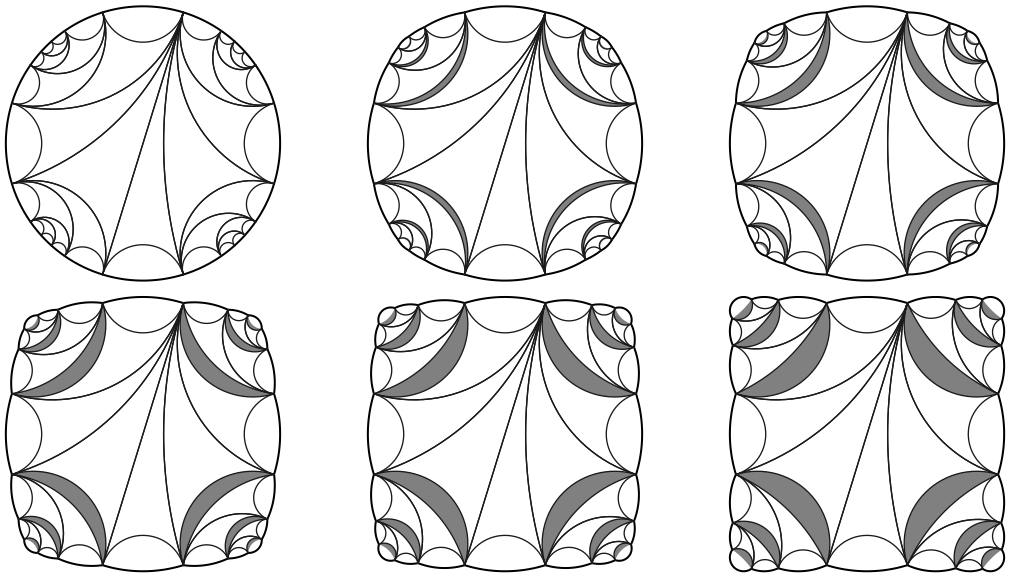


FIGURE 15. The one parameter family connecting the disk to a finitely bent approximation of the square. In each picture the angles have been multiplied by $t = 0, .2, .4, .6, .8, 1$

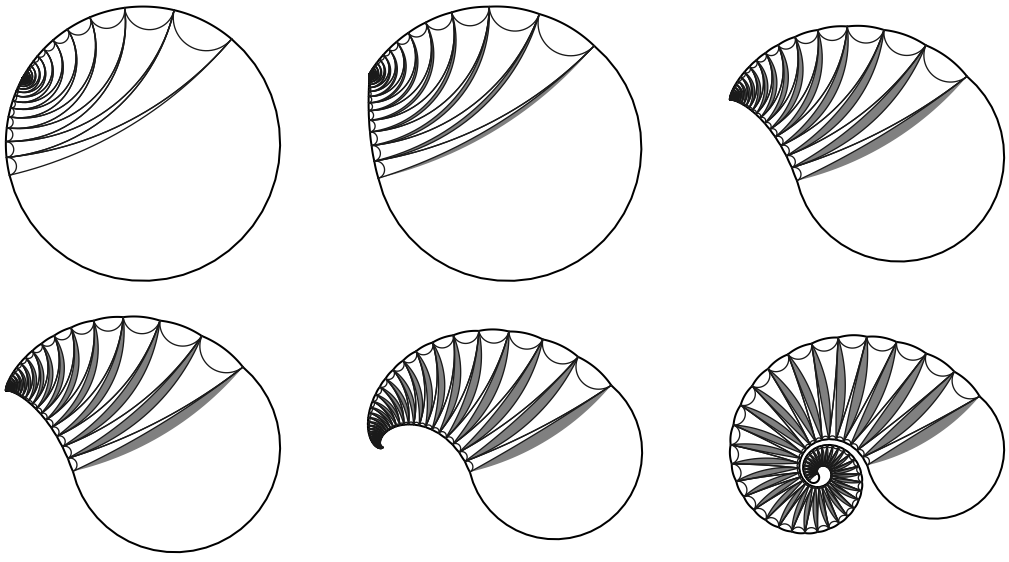


FIGURE 16. An approximate logarithmic spiral with $t = 0, .2, .4, .6, .8, 1$. Logarithmic spirals were used by Epstein and Markovic in [**?**] to disprove Thurston's $K = 2$ conjecture. The showed that (in a precise sense) certain spirals have too much gray.

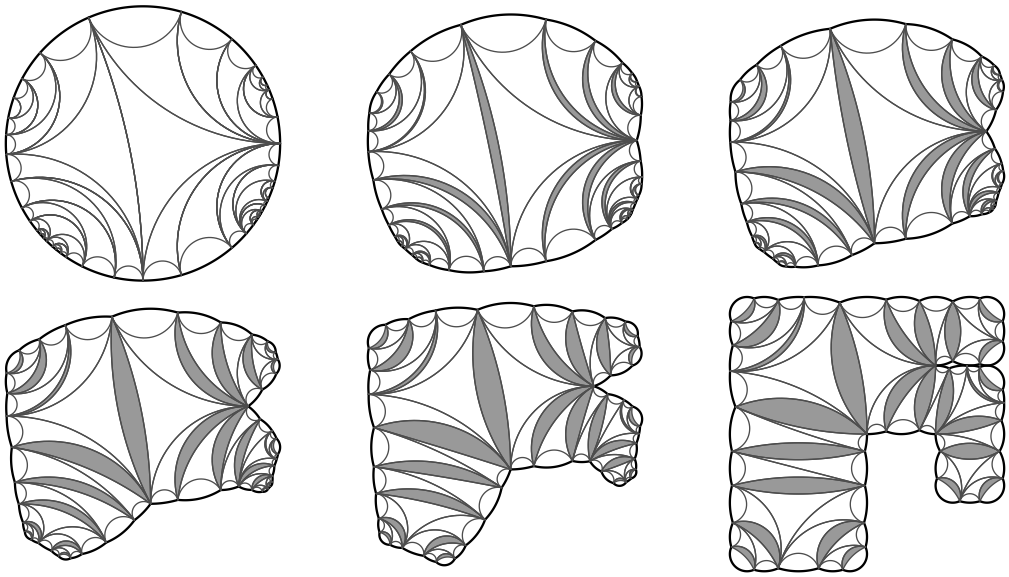


FIGURE 17. The domain from Figure 13 with $t = 0, .2, .4, .6, .8, 1$.

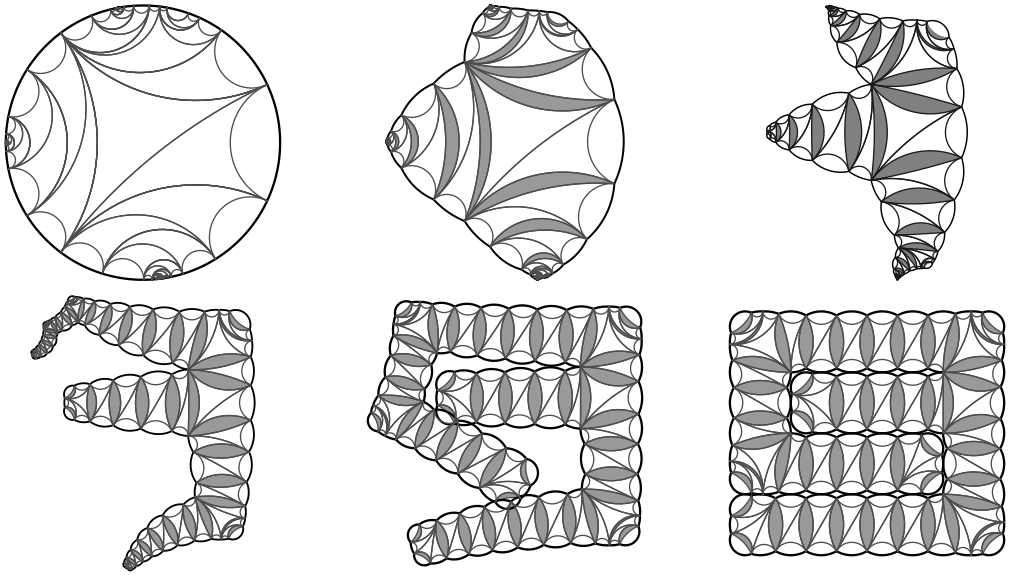


FIGURE 18. An example where intermediate domains need not be planar. The pictures correspond to multiplying the angles by $t = 0, .4, .8, .95, .99, 1$. Note that the parameter must be very close to 1 before we see the longer corridors clearly.

would have to agree with the map on the gap that is adjacent to C_2. We will let $\varphi_{s,t} : \Omega_t \to \Omega_s$ denote this map. Let $\rho_s = \rho_{\Omega_s}$ denote the hyperbolic metric on Ω_s. Suppose N is a large integer and choose points $t_0 = 0, t_1 = \frac{1}{N}, \ldots, t_N = 1$. Let $\Omega_k = \Omega_{t_k}$ for $k = 0, \ldots, N$. Let $\varphi_k : \Omega_{k+1} \to \Omega_k$ be defined by $\varphi_k = \varphi_{\frac{k}{n}, \frac{k+1}{n}}$.

6. Angle scaling is QC continuous

Given a pair of domains Ω_s, Ω_t with $0 \leq s < t \leq 1$, let $\iota_{s,t} : \partial\Omega_t \to \partial\Omega_s$ be the obvious boundary map obtained multiplying the angle of each crescent by s/t. We will extend this boundary map to the interiors by writing each crescent C in Ω_t of angle α as a union of crescents C_1, of angle $\alpha s/t$ and C_2, of angle $\alpha(1 - s/t)$. On C_1 we collapse each leaf of the E-foliation to a point (hence C_1 is maps to a circular arc) and we let our map be Möbius on C_2. By continuity, this Möbius transformation would have to agree with the map on the gap that is adjacent to C_2. We will let $\varphi_{s,t} : \Omega_t \to \Omega_s$ denote this map. Let $\rho_s = \rho_{\Omega_s}$ denote the hyperbolic metric on Ω_s.

We claim that angle scaling is continuous with respect to the quasiconformal distance. More precisely, angle scaling defines a map between the boundaries of Ω_s and Ω_t and we claim that this map has a $1 + O(|s - t|)$ quasiconformal extension to the interiors. As before it will be simplier to show it has a $1+o(|s-t|)$ quasi-isometric extension to the interiors.

LEMMA 111. *Suppose $\{tau_k\}$ is a collection of elliptic transformations acting on the hyperbolic ball $\mathbb{B}$, with τ_k a rotation of angle θ_k around points $\{a_k, b_k\}$. Assume that the Euclidean distance $|a_k - b_k| \geq 1$ for each k. Then no point on $\partial\mathbb{B}$ is moved more than Euclidean distance $O(\sum_k |\theta_k|$ by $\tau_1 \circ \cdots \circ \tau_n$.*

PROOF. This is obvious since the condition on the fixed points implies each individual map moves points by at most $O(|\theta_k|)$. □

Now suppose $z, w \in \Omega_s$ and normalize both Ω_s and Ω_t by Möbius transformations so that $0 = z = \varphi_{st}(z)$ and $\text{dist}(0, \partial\Omega_s) = \text{dist}(0, \Omega_t) = 1$. Then φ_{st} acts on w as a composition of elliptic transformations $\{\tau_k\}$ (one for each normal crescent separating w and 0) whose fixed points are well separated. Moreover, elliptic rotates by $|s-t|\alpha_k$ and $\sum_k \alpha_k \leq e^{3s}$ by Lemma 5 if 0 and w are within hyperbolic distance s. Thus φ_{st} moves w only Eucliean distance $O(|s - t|)$, which by our normalization, is also only

hyperbolic distance $O(|s-t|)$. Thus

$$\rho(z,w) \leq 1 \quad \Rightarrow \quad \rho(\varphi_{st}(z), \varphi_{st}(w)) \leq \rho(z,w) + O(|s-t|).$$

For general points (joining them by geodesics broken into unit length segments) this implies

$$\rho(\varphi_{st}(z), \varphi_{st}(w)) \leq \rho(z,w)(1 + O(|s-t|)) + O(|s-t|).$$

Note, however, that we have never specified whether $s < t$ or $t < s$, so this estimate applies equally well to the inverse function φ_{ts}, so

$$\rho(\varphi_{st}(z), \varphi_{st}(w)) \geq \frac{\rho(z,w)}{1 + O(|s-t|)} - O(|s-t|).$$

This gives both sides in the definition of quasi-isometry, so we are done.

As noted before, the retraction map $R : \Omega \to S$ is a quasi-isometry. Thus $\varphi = \iota \circ R : \Omega \to \mathbb{D}$ is also a quasi-isometry between the hyperbolic metrics. The same is true for the maps $\varphi_{s,t}$ for any $0 \leq s < t \leq 1$, with constant bounded by $O(|s-t|)$. We will prove this next, together with a variation we will use later to estimate the norm of a ϵ-representation when composed with $\varphi_{s,t}^{-1}$. First we need an estimate on the bending lamination.

Consider the bending lamination Γ associated to a finitely bent domain Ω. Suppose a hyperbolic s-ball hits geodesics in Γ with angles $\alpha_1, \ldots, \alpha_m$. We want to show that there is an upper bound $\sum_j \alpha_j \leq B(s)$ which only depends on s. See [**?**], [**?**] for some variations of this idea. Estimates of B are also closely tied to results of Bridgeman [**?**], [**?**] on bending of surfaces in hyperbolic spaces. Here we shall give a simple conceptual proof without an explicit estimate.

LEMMA 112. *There is a $C < \infty$ so that $B(s) \leq Ce^{3s}$.*

PROOF. Suppose Ω is normalized so $\infty \notin \Omega$. The normalization implies that if γ is a bending geodesic in $\mathbb{R}^3_+$ which hits the plane at 1 and -1, then the corresponding crescent is in the unit disk. Moreover, an easy estimate shows that a crescent with vertices ± 1 and angle α has area $\geq c\alpha$ for some fixed $c > 0$.

If $\tilde{\gamma}$ is a bending geodesic with angle β that passes within hyperbolic distance s of $(0,0,1)$ then the "highest" point of $\tilde{\gamma}$ has Euclidean height at least e^{-s} above the plane $\mathbb{R}^2$. Thus its two endpoints on the plane are at least $2 \cdot e^{-s}$ apart. Moreover at least one endpoint must be contained in the disk of diameter e^s around the origin

(if not then $\tilde{\gamma}$ lies outside the hemisphere with this disk as its base which means the hyperbolic distance to $(1,0,0)$ is $\geq s$).

Thus the part of the crescent corresponding to $\tilde{\gamma}$ inside the ball $B(0, e^s+1)$ has area at least $ce^{-s}\beta$. Consider the set of all bending geodesics that come within hyperbolic distance s of the point $(0,0,1) \in \mathbb{R}^3_+$ and let $\{\alpha_n\}$ be an enumeration of the bending angles. Since the crescents are disjoint we deduce $\sum_n \alpha_n \leq \frac{1}{c}\pi e^s(e^s+1)^2 \leq Ce^{3s}$, as desired. (Note that this argument is not sharp since the crescents can have small area only when then are close to the origin; also note that we must have $B(2s) \leq 2B(s)$ which implies at most linear growth for large s.) □

The following simple lemma quantifies the fact that a elliptic Möbius transformation with small rotation angle is close to the identity.

LEMMA 113. *Suppose σ is an elliptic Möbius transformation with fixed points a, b and rotation angle θ. If $r = |z - (a+b)/2| \leq \frac{1}{8}|b-a|/\theta$ and $|\theta| \leq \frac{1}{4}$, then we have*

$$|z - \sigma(z)| \leq C(1+r^2)\frac{|\theta|}{|b-a|}.$$

PROOF. This is an explicit computation. We may assume $a = 1$, $b = -1$ in which case σ has the form $\sigma(z) = \tau^{-1}(\lambda\tau(z))$ where $\lambda = e^{i\theta}$ and $\tau(z) = (z-1)/(z+1)$. Doing some arithmetic, and using $|1-\lambda| \leq |\theta|$, we get

$$|\sigma(z) - z| = |\frac{(1-\lambda) - (1-\lambda)z^2}{(1+\lambda) + (1-\lambda)z}| \leq |\theta|\frac{|1-z^2|}{1-|\theta|-|\theta||z|} \leq 4|\theta|(1+|z|^2),$$

if $|\theta| \leq \frac{1}{4}$ and $|z| \leq \frac{1}{4|\theta|}$. □

The following is the main result of this section.

LEMMA 114. *Suppose $r > 0$ is given. There is an $\epsilon > 0$, depending only on r, so that if $0 \leq s < t \leq 1$ and $|s-t| \leq \epsilon$ then the following holds. Suppose G_1 and G_2 are gaps in the normal crescent decomposition of the finitely bent domain Ω_s such that $\rho_S(R(G_1), R(G_2)) \leq r$. Suppose τ_j are Möbius transformations so that $\varphi^{-1}_{s,t}|_{G_j} = \tau_j$ for $j = 1, 2$. Then*

$$\rho_t(\tau_1(z), \tau_2(z)) \leq C_r|t-s|,$$

for every $z \in \Omega_s$ with $\rho_s(z, G_1) \leq r$.

PROOF. The statement is invariant under renormalizing by Möbius transformations so we may assume that G_1 has base disk $\mathbb{D}$, that $z_1 = 0 \in G_1$ is within $2r$ of G_2, and that τ_1 is the identity.

Then τ_2 is a composition of the elliptic transformations $\{\sigma_j\}$ that correspond to the normal crescents $\{C_j\}$ that separate G_1 and G_2. By Lemma 112, the measure of the bending geodesics separating G_1 and G_2 is at most $B(r)$.

Since $\rho_S(C_j, 0) \leq r$ for all j, C_j has diameter $\geq e^{-r}$ and one vertex is contained within $D(0, e^r)$ by the proof of Lemma 112. By Lemma 113 this means that σ_j moves points in $D(0, C)$ at most $C|\theta_j|$ with C depending only on r, assuming θ_j is small enough (depending only on r). Thus

$$|\tau_2(z) - z| \leq C_r|s-t| \sum_j |\theta_j| = O(|s-t|), \tag{31}$$

for $|z| \leq C$, assuming $|s-t|$ is small enough, depending only on r.

If $\rho_t(0, z) \leq r$, then $|z| \leq C_r$ and $\operatorname{dist}(z, \partial\Omega_s) \geq C_r > 0$ with estimates which only depend on r (see Lemma **??**, Appendix D). Thus for $|s-t|$ small enough, $\rho_t(z, 0) \leq r$ and $|z - w| \leq \epsilon$ imply $\rho_t(z, w) \leq C_r\epsilon$. Hence for a given r we can choose $|s-t|$ so small that (31) implies $\rho_t(\tau_2(z), z) \leq O(|s-t|)$, (with constant depending on r). □

LEMMA 115. *$\varphi_{s,t}$ is a quasi-isometry with constant $O(|s-t|)$.*

This follows immediately from the following technical result which will be proven in Appendix **??**. It also follows from a careful reading of [**?**] which gives an explicit construction of a quasiconformal map from $\mathbb{D}$ to a finitely bent domain Ω with boundary values φ^{-1}. The method can be adapted to give an explicit map $\Omega_s \to \Omega_t$ which is quasiconformal with constant $O(|s-t|)$.

THEOREM 116. *Suppose Ω_0, Ω_1 are simply connected and $\varphi : \Omega_0 \to \Omega_1$ has the following property: there is a $0 < C < \infty$ so that given any hyperbolic C-ball B in Ω_0, there is a Möbius transformation σ so that $\rho_{\Omega_0}(z, \sigma(\varphi(z)) \leq \epsilon$ for every $z \in B$. Then there is a hyperbolic $(1 + O(\epsilon))$-biLipschitz map $\psi : \Omega_0 \to \Omega_1$ so that $\sup_{z\in\Omega_0} \rho_{\Omega_1}(\varphi(z), \psi(z)) \leq O(\epsilon)$. In particular, φ is a quasi-isometry between the hyperbolic metrics with constant $O(\epsilon)$.*

7. Angle scaling and Davis' method

Suppose we have an iterative method of computing a conformal map that works if we start with a good enough guess. For example, suppose that given a a guess $\mathbf{w}$ for the Schwarz-Christoffel parameters $\mathbf{z}$, we have an algorithm that will converge to $\mathbf{z}$ assuming $d_{QC}(\mathbf{w}, \mathbf{z}) < \epsilon_0$ for some fixed ϵ_0. How can we turn this into a method that always finds $\mathbf{z}$? In Chapter 5 we saw that there are simple methods that give initial guesses that are guareented to be within some universal distance K of the correct answer, but this K may be larger than ϵ_0. How do we bridge the gap?

In this chapter, we will show that any finitely bent domain Ω can be associated to a finite chain of domains $\Omega_0, \Omega_1, \ldots, \Omega_N$ so that

(1) $\Omega_0 = \mathbb{D}$,
(2) $\Omega_N = \Omega$,
(3) $N = O(1/\epsilon_0)$,
(4) If $\mathbf{z}_n$ are the conformal preimages of the vertices of Ω_n, then $d_{QC}(\mathbf{z}_n, \mathbf{z}_{n+1}) < \frac{1}{2}\epsilon_0$.

We can easily find the Riemann map onto Ω_0 by taking the identity. This map is then within QC-distance ϵ_0 of the map onto Ω_1, so is a suitable starting point from which to iterate to the map onto Ω_1. In general, if we know the conformal map onto Ω_k to within QC -distance ϵ_2, then it is also within ϵ_0 of being a map onto Ω_{k+1}, so we can apply our iterative method until we approximate a map onto Ω_{k+1} to within $\epsilon_0/2$. Finally, when we get within $\epsilon_0/2$ of the for $\Omega_N = \Omega$ we simply apply the iterative method until we get the desired accuaracy.

CHAPTER 9

Linear methods

In this chapter we will consider some methods for computing conformal maps that require us to solve a linear system $Ax = b$. The first is Symm's method, which is based on the fact that finding the conformal map is equivalent to solving the Dirichlet problem. The next is a remmarkable formula of Kerzman and Stein which gives a formular for the Riemann map from Ω to the disk in terms of the Szegö kernel, a function that itself can be computed from the Cauchy kernel by inverting an explicit linear operator.

We will then give a very brief introduction to the vast field of of numerical linear algebra: how can we solve the necessary linear systems $Ax = b$? Naive methods are $O(n^2)$, but iterative methods which require only repeated applications of the matrix A offer two possible improvements. First, although such methods may require n iterations to reach an "exact" solution, they often reach an "acceptable" solution (e.g., within machine accuracy) much sooner. The conjugate gradient method is such a method for real symmetric matrices. Second, whereas the application of a general $n \times n$ matrix A to an n-vector x takes time $O(n^2)$, the matrices that arise in practice are highly structured and Ax can sometimes be computed approximately in time $O(n)$ or $O(n \log n)$ using the fast multipole method of Greengard and Rokhlin.

1. A linear algebra glossary

Here we review some general definitions and results that we will need in our discussion of linear methods. In a later section we will discuss a few specific techniques to solve linear systems more carefully.

If we have a $n \times m$ matrix $A = (a_{jk})$ (n rows and m columns), and x is a m-dimensional vector then Ax is the n dimensional vector whose jth component is $\sum_{k=1}^{m} a_{jk} x_j$. Entries may either be real or complex. The usual dot product between vectors of equal length will be denoted $\langle v, w \rangle = \sum_k v_k \overline{w_k}$.

Normed space: A norm on a vector space V is a function from V to the non-negative reals so that

(1) $\|x\| \geq 0$ with equality iff $x = 0$.
(2) $\|\lambda x\| = |\lambda|\|x\|$. item $\|x + y\| \leq \|x\| + \|y\|$.

A normed vector space is a Banach space if it is also complete, i.e., every Cauchy sequence converges.

Inner product: An inner product on a vector space V is a function from $V \times V$ into the scalars so that

(1) $\langle x, y\rangle = \overline{\langle y, x\rangle}$.
(2) $\langle x + y, x\rangle = \langle x, z\rangle + \langle y, z\rangle$.
(3) $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$.
(4) $\langle x, x\rangle \geq 0$ with equality iff $x = 0$.

An inner product defines a norm by $\|x\| = \langle x, x\rangle$. If V is complete with this norm, then it is called a Hilbert space.

Annihilator: Given a set $M \subset H$, the annihilator of M is $M \perp = \{x : \langle x, y\rangle$ for all $y \in M\}$.

%item [Kernel] **kernel:** The kernel of a linear map T is $T^{-1}(0)$.

Transpose: The tranpose of $A = (a_{jk})$ is $A^T = (a_{kj})$ (i.e., reflect elements over the main diagonal. The conjugate transpose is $A^* = (\overline{a_{kj}})$.

Adjoint: Most generally a linear map $A : X \to Y$ between normed vector spaces, the adjoint operator $A*$ maps Y^*, the dual of Y, to X^*, the dual of X and is defined by $\langle x, A^*y^*, \rangle = \langle Ax, y^*\rangle$. If H is a Hilbert space, then there is a cannoical identitfication between H and its dual (the element y is identified with the linear funtional $x \to \langle x, y\rangle$). So for a operator $T : H \to H$ we may consider the adjoint T^* as also mapping H to H. This adjoint has the properties

(1) $\|T^*\| = \|T\|$.
(2) $(T + S)^* = T^* + S^*$
(3) $(\lambda T)^* = \bar{\lambda} T^*$
(4) $(ST)^* = T^*S^*$
(5) $T^{**} = T$.

Self-Adjoint: The real matrix A is called self-adjoint if $A = A^*$.

Spectrum: For a general operator A on a Hilbert space, the spectrum is the set of complex number so that either $I - \lambda A$ is not onto or has non-trivial kernel. In the latter case we say λ is an eigen value of A. For finite dimensional spaces the two conditions are equivalent, i.e., the spectrum is the set of eigenvalues.

Normal: An operator is normal if $AA^* = A^*A$. The following facts hold:

(1) T is normal iff $\|Tx\| = \|T^*x\|$ for all x.
(2) If T is any operator then T^*T is normal.
(3) If T is normal then $\mathcal{N}(T) = \mathcal{N}(T^*) = \mathcal{R}(T)^\perp$.
(4) If T is normal and $Tx = \lambda x$ then $T^*x = \bar{\lambda}x$.
(5) If T is normal the eigenspace of the distinct eigenvalues are orthogonal.
(6) A normal operator is self-adjoint iff its spectrum lies on the real axis.
(7) A normal operator is unitary iff its spectrum lies on the unit circle.
(8) If T is normal, $\|T\| = \sup_{\|x\|\leq 1}\langle Tx, x\rangle$.

The spectral theorem: For a general Hilbert space, the spectral theorem says that a normal operator can be written as a integral over the spectrum of an operator valued measure. This requires some specialized definitions to state exactly (see e.g. Chapter 12 of Rudin's book [**?**]), but the finite dimensional case is easy to state: if T is a normal operator on a n dimensional space then we can choose n orthogonal eigenvectors $\{v_k\}$ with eigenvalues $\{\lambda_k\}$ so that

$$Tx = \sum_{k=1}^{n} \lambda_k \langle x, v_k\rangle v_k.$$

If U is the unitary operator defined by mapping the standard basis $\{e_k\}$ to $\{v_k\}$, then $U^{-1}TU$ is the diagonal matrix with entries $\{\lambda_k\}$.

Unitary: T is unitary if $T^*T = TT^*$. This occurs iff T is onto and either $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all x, y or $\|Ux\| = \|x\|$ for all x.

Projection: An operator is a projection if $T^2 = T$. If T is a projection, then it is normal iff it is self-adjoint iff $\langle Tx, x\rangle = \|Tx|^2$ for all x iff $\mathcal{R}(T) = \mathcal{N}(T)^\perp$.

Upper triangular: Just what you think, $A = (a_{jk})$ with $a_{jk} = 0$ if $j > k$.

Hessenberg: A is Hessenberg if it is almost upper triangular, i.e., $A = (a_{jk})$ with $a_{jk} = 0$ if $j > k + 1$.

Gaussian elimination: This is the standard direct method for making a matrix upper triangular. Start with the first row. If the first column is all zeros do nothing.

Otherwise, there is a non-zero entry. If the top entry is zero, (i.e., $a_{11} = 0$) then swap the first rows with some row the starts with a non-zero entry, otherwise leave the first row alone. Then multiply the first row by a_{21}/a_{11} and subtract it from the second row, creating a 0 in the $2, 1$ position. Then subtract a_{31}/a_{11} times the first row from the third row, making its leading elements zero as well. Proceeding in this way we can make the first column all zeros except possibly for the a_{11} position. Now repeat this procedure using the second column and all but the first rows. Repeating the procedure on smaller blocks of the matrix leads to an upper diagonal matrix. Since all the operations are linear, this means that there is a matrix Q so that $QA = T$ is upper triangular.

Positive operator: T is a positive operator on a Hilbert space if $\langle Tx, x\rangle \geq 0$ for every $x \in H$. If T is positive then there is a unique positive operator S so that $s^2 = T$. If T is any bounded operator, T^*T is positive (since $\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle \geq 0$) and its square root is the only positive operator that satisfies $\|Px\| = \|Tx\|$ for all x. Any invertible T can be written as $T = UP$ where U is unitary and P is positive (this is the polar decomposition of T). If, in addition, T is normal then U and P commute with each other.

Singular values: If A is a $n \times n$ matrix and A^* is its conjugate transpose, then A^*A is normal, so has n eigenvalues, and the singular values of A are the square roots of these eigenvalues. Any matrix A can be written $A = UDV$ where U, V are unitary and D is diagonal with the singular values of A as its entries.

Compact operator: An operator $T : X \to Y$ between two normed vector spaces is called compact of the image of the open unit ball has compact closure. If T is a normal operator from a Hilbert space to itself, then it is compact iff the spectrum has no limit point except zero and the eigenspace of every non-zero eigenvalue is finite dimensional. An important example of a compact operator on $L^2(\mathbb{T})$ is convolution with a smooth kernel.

Orthogonal: Q is orthogonal if $Q^TQ = I$. This means the columns of Q form an orthonormal basis.

QR decomposition: If A is an $m \times n$ matrix then we can write $A = QR$ where Q is a $m \times m$ orthogonal matrix and R is $m \times n$ upper triangular matrix. If $m > n$

this means the bottom $m-n$ rows of R will be all zeros. The QR decomposition can be computed uysing Gramm-Schmitt, Householde reflections or Givens rotations.

Gram-Schmidt: Given a list of vectors $\{a_k\}$ this produces a list of orthonormal vectors $\{q_k\}$ which span the same subspaces. The vector q_k is calculated by taking a_k and subtracting off its projections onto the subspaces corresponding to the perviously produced vectors, i.e., we write

$$\begin{aligned} b_1 &= a_1/\|a_1\| \\ q_1 &= b_1/\|b_1\| \\ b_2 &= a_2 - \langle a_2 q_1\rangle q_1 \\ q_2 &= b_2/\|b_2\| \\ &\vdots \\ b_n &= a_n - \langle a_n q_1\rangle q_1 \cdots - \langle a_n q_{n-1}\rangle q_{n-1} \\ q_n &= b_n/\|b_n\| \end{aligned}$$

which we can rewrite as

$$\begin{aligned} a_1 &= q_1\|b_1\| \\ a_2 &= \langle a_2 q_1\rangle q_1 + q_2\|b_2\| \\ &\vdots \\ a_n &= \langle a_n q_1\rangle q_1 + \cdots + \langle a_n q_{n-1}\rangle q_{n-1}\|b_{n-1}\| \end{aligned}$$

In matrix terms this system is $A = QR$ where a_k is the kth column of A, q_k is the kth column of Q and R is upper triangular. In actual computations, the vectors produced may not be orthogonal due to rounding errors. These can be lessened if, instead of writing,

$$b_k = a_k - \langle a_k q_1\rangle q_1 \cdots - \langle a_k q_{k-1}\rangle q_{k-1},$$

we write

$$\begin{aligned} b_k^1 &= a_k - \langle a_k q_1 \rangle q_1 \\ b_k^2 &= b_k^1 - \langle a_k q_2 \rangle q_2 \\ &\vdots \\ b_k = b_k^{k-1} &= b_k^{k-2} - \langle a_k q_{k-1} \rangle q_{k-1}. \end{aligned}$$

In exact arithmetic this gives the same result, but gives smaller errors in finite precision.

Householder reflections: These are another way to produce the QR decomposition. Let e_1 be the column vector starting with 1 and followed by all zeros. Let x be the first column vector of A and let v be the unit vector in direction$x - \|x\|e_1$ and define the matrix $Q_1 = I - 2vv^T$. Then $Q_1 v = v - 2vv^T v = v - 2v = -v$ and $Q_1 w = w - vv^T w = w$ for any vector orthogonal to v. Thus Q_1 defines a reflection through the hyperplane perpendicular to v. Also, $Q_1 x = \|x\|e_1$, so $Q_1 A$ has first column which is zeros except for the top entry. Applying the same procedure to the submatrix of $Q_1 A$ obtained by omitting the first row and column, we can define a matrix Q_2 so $Q_2 Q_1 A$ is upper triangular in the first two columns. After n steps we have $QA = Q_n \cdots Q_2 Q_1 A = R$ is upper triangular. Moreover, Q is orthogonal since it is the product of reflections.

Givens rotations: Yet another way to produce the QR decomposition by converting individual elements under the diagonal to zero. For example, suppose we want to make the bottom element of the first column of A a zero. We think of the bottom two elements of this column, $x = a_{1,n-1}, y = a_{1,n}$ as representing the point (x, y) in the plane and we want to rotate this point onto a the point $r = \sqrt{x^2 + y^2}$ on the real line. The 2 matrix that does this is

$$R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

where $\tan\theta = y/x$. So if we multiply A by the $n \times n$ matrix

$$Q = \begin{pmatrix} I_{n-2} & 0 \\ 0 & R_\theta \end{pmatrix}$$

we get $r, 0$ in the bottom two positions of the first column of QA. In a similar way we can convert other elements to zero.

Least squares: A least squares problem is to choose a vector x which minimizes the norm $\|Ax - y\|$, given a $m \times n$ matrix A and an n vector Y. Given the QR decomposition of A this can be solved as follows. Since Q is orthogonal,

$$\|Ax - y\| = \|QRx - y\| = \|Rx - Q^T y\|.$$

If R were invertible we then take $x = R^{-1}Q^T y$, which makes the norm 0 (an obviious minimum); since R is upper triangular this is easy to compute. If R is not invertible or not square, then it has $k < n$ non-zero rows. The minimal norm is attained by projecting y orthogonally onto span of first k columns of R. Let y^k be the first k coordinates of y and R_k the upper left $k \times k$ submatrix of R and let x solve $R_k x = y^k$. Extend x to an n vector by adding zeros in positions $k + 1$ to n.

2. Iterative metods for linear systems

We will see later that approximating the conformal map from Ω to the disk could be reduced to solving a linear equation $Ax = y$. It is very tempting to simply say "now apply A^{-1} to both sides and we arrive at the desired solution $x = A^{-1}y$." However, it is not quite this simple. Numerical linear algebra is devoted to describing how to find x in a reasonable amount of time and with reasonable accuracy. This is a vast area with numerous techniques adapted to various situations. We shall merely skim its surface by describing a couple of algorithms which are useful for the linear systems described earlier. Our treatment closely follows the book [**?**] of Bau and Trefethen. The reader is strongly encouraged to consult it directly.

Roughly speaking, given a linear system $Ax = y$ to solve, we can apply either direct or iterative methods. Direct methods include things like Gaussian elimination where we convert A into an upper triangular matrix by subtracting appropriate multiplies of each row from all the rows below it. This is solve the system exactly, but takes $O(n^3)$ operations to solve an $n \times n$ system and we don't even get an approximate answer until we are completely finished. Moreover, the phrase "solves the system exactly" must be treated suspiciously if the method is implemented on a computer with finite precision; in the presence of round-off errors no method is exact. There are other "exact methods" that will the $O(n^3)$ bound; Strassen discovered a direct

algorithm in 1969 that reduces the exponent 3 to $\log_2 7 \approx 2.81$ and Coppersmith and Winograd reduced this further to ≈ 2.376. However, almost nothing is known about the numerical stability of these methods and the improvement in the exponents is not large enough to have had a dramatic impact on practical computing. What has had a dramatic impact are iterative methods.

An iterative method for solving an $n \times n$ system $Ax = y$ consists of an algorithm that produces a sequence of vectors $\{x_k\}$ which converge to x (in the methods we consider, we get the exact answer $x_n = x$ after n steps) and which only involve applying A to various vectors. This offers two possible improvements over direct methods. First, if the sequence of approximate solutions converges to the actual solution fast enough, we may only have to apply a small number of iterations before we reach an acceptable solution. For example, if $\|x_k - x\|$ decays exponentially, then in $O(1)$ iterations we will have computed x to within machine precision. The second potential savings is that we only use the matrix A in order to apply it to a vector. For a general $n \times n$ matrix this takes $O(n^2)$ operations. However, the matrices that arise in practice are often highly stuctured, and we saw in Section 5 that some stuctured matrices could be applied exactly to vectors in time $O(n \log n)$. If we have certain structured matrices and require, not an exact answer, but only some degree of accuracy (say machine precision), then we can sometimes apply the matrix to a vector in time $O(n)$ (with a constant depending on the desired accuracy). If both senarios hold then we can hope to solve the system with small error in time only $O(n)$. When n is in the thousands, this is a significant improvement over $O(n^3)$ (even if the multiplicative constant is fairly large).

In this section we will discuss the first improvement: iterative methods that may converge quickly to the desired answer. In the next section we will discuss a method of applying certain analytically defined matrices to vector in time $O(n)$: the fast multipole method of Greengard and Rohklin.

Consider the m-dimensional subspace K_m generated by $Ay, A^2y, \ldots A^m y$, called the Krylov subspace. The two iterative methods we will describe choose the "best" approximate solution to $AX = y$ from this space, the main difference between them being how we define best. In the conjugate residual method we seek $x_m \in K_m$ which

minimizes

$$\|Ax_n - y\|_2,$$

and in the conjugate gradient method (which only works if A is self-adjoint and positive definite) we seek to minimize

$$\|x - x_m\|_A = \sqrt{\langle (x - x_m), A(x - x_m)},$$

where x denotes the true solution. (Since we don't know x ahead of time, we will have to explain how we can minimize a function we can't compute, but it turns out the same x_m minimizes another function we can compute).

We start with the more general conjugate residual method, a.k.a. GMRES for "generalized minimal residuals". This is really an iteration within an iteration. The outer iteration finds an orthonormal basis for the Krylov space and the inner loop finds the optimal solution within that space with the help of the basis.

Let K_n be the space spanned by the n vectors $y, Ay, \dots A^{n-1}y$,

$$AK_n = [Ab|A^2b| \dots |A^nb].$$

We want to find a linear combination $x_n = K_nc$ of these column vectors which minimizes the distance to y, i.e., find a column n vector to minimize $\|x_n - y\| = \|AK_nc - y\|$. We will find an orthonormal basis $\{q_n\}$ for K_n and then project y orthogonally onto each basis direction. Since $\{q_n\}$ also span K_n, the vector x_n can be written as a linear combination of these vectors, i.e.,we want to find $x_n = Q_nd$, (where the columns of Q_n are the $\{q_n\}$) for some column vector d which minimizes $\|Q_nd - y\|$.

Both vectors inside the norm are in the Krylov space K_{n+1} and hence in the span of the columns of Q_{n+1}. The conjugate transpose, Q^*_{n+1} has rows that are orthogonal vectors. Thus $Q^*_{n+1}v$ has the same norm as v. Thus applying Q^*_{n+1} to $Q_nd - y$ says we are trying to choose d to minimize where $H_n = Q^*_{n+1}AQ_n$ is a Hessenberg matrix. This means that it almost upper triangular; we have $h_{jk} = 0$ if $j > k + 1$. This happens because A maps q_k into the span of $q_1, \dots q_{k+1}$ and hence Aq_k is orthogonal to q_j for all $j > k + 1$, which means the corresponding entries of H_n are zero. Also note, that by construction, $Q^*_{n+1}y = \|y\|e_1$ where e_1 is the column vector with a 1 at the top and all 0's elsewhere. Thus we seek to minimze $\|H_nd - \|y\|e_1\|$.

To find the $\{q_k\}$ we start by setting $q_1 = y/\|y\|$. Then for each $k = 1, 2, \ldots,$ let $v = Aq_k$. For a fixed k, and $j = 1, \ldots, k$, set $v = v - q_j\langle q_k, v\rangle$ and finally $q_{n+1} = v/\|v\|$. This is just Gram-Schmit applies to Aq_k to make it othogonal to the previous basis vectors. If we reach a vector where $\|v\| = 0$, the last step is impossible, but in this case it means that v was not linearly independent of the previous basis elements and the exact solution of the linear system $Ax = y$ was already in the span of the previous basis elements (and so we already found it).

After we have found each q_k we can find the vector d that minimizes $\|H_n d - \|y\|e_1\|$. This is a least squares problem that can be solved by the QR factorization as described in Section 1.

How many iterations of GMRES are needed? At each stage we are minimizing the norm of $\|r_n\| = \|Ax_n - y\|$ over a nested, increasing set of spaces, so these norms decrease monotonically. Moreover, x_n can be written as a linear combination of powers of A applied to y, i.e.,

$$x_n = q_n(A)y,$$

so the residual can be written

$$r_n = y - Ax_n = (I - Aq_n(A))y = p(A)y,$$

where p_n has contant term 1. Thus

$$\|r_n\| = \|p_n(A)b\| \leq \|p_n(A)\|\|b\|,$$

Thus we expect good behavior when $\|p_n(A)\|$ goes to zero quickly.

If A is diagonalizable, say $A = VDV^{-1}$ where D is diagonal and V is invertible, then

$$\|p(A)\| = \|V\|\|p(D)\|\|V^{-1}\|.$$

The middle term is equal to the maximum of $|p_n(z)|$ over the entries of D (i.e., the eigenvalues of A) and the product of the two other terms is, by definition, the condition number $\kappa(V)$ of V.

If A is a normal matrix, i.e., it commutes with its conjugate transpose, then the spectral theorem says that we can take V to be unitary. This means its condition number is 1, and the behavior of GMRES depends only on the "distribution" of the eigenvalues of A. Let Λ be the set of eigenvalues. GMRES will behave well if there are kth degree polynomials that are 1 at 0 but small on Λ. This happens if Λ is

clustered around a non-zero point. For example, if $\Lambda \subset D(a,r)$ then the polynomial $p_k(z) = ((z-a)/|a|)^k$ satisfies $p_k(0) = 1$ and $\sup_\Lambda |p_k| \leq (r/|a|)^k$. On the other hand, if Λ consists of the nth roots of unity then any polynomial of degree $k < n$ which is 1 at the origin must be $\geq \epsilon \simeq k^{-1/2}$ at at least one point of Λ. (If $p(z) = \sum_j a_j z_j$ has modulus $\leq \epsilon$ at every root of unity then the discrete Fourier transform implies $\|a_j\|_2 \leq \epsilon$. Thus $|p'(z)| \leq \sum_j |a_j| j \leq \|\{a_j\|_2 k^{3/2} = O(\epsilon k^{3/2})$. The separation between kth roots of unity is $O(1/k)$ so this means that the maximum of p on the unit circle is at most $\epsilon + O(\epsilon k^{1/2})$. Since this polynomial is 1 at the origin, this violates the maximum prinicple if $\epsilon << k^{-1/2}$).

If the matrix A is not normal we can try to solve the system $A^*Ax = A^*y$, since A^*A is always normal (indeed, $(A^*A)^* = A^*A^{**} = A^*A$). In general, the eigenvalues of A^*A are the squares of the singular values of A (note that the eigenvalues of A do not influence the convergence in this problem), so at first glance it seems that this system might not be as good to work with. However, this need not be the case. An exteme example is given by the matrix A which cyclicly permutes the n basis elements,

$$A = \begin{pmatrix} 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

The eigenvalues of A are the nth roots of unity and GMRES will take the full n steps to find any better solution to $Ax = e_1$ than the zero vector. However A^*A is the identity, which can be solved trivially.

The second iterative method we will discuss is the well known conjugate gradient method. Our discussion here closely follows the presentation in [**?**]. There are variations for other types of matrices, but we will only discuss the simpliest version which deals with positive definite matrices A. For such a matrix

$$\|x\|_A = \langle x, Ax \rangle,$$

definies a norm and the conjugate gradient method seeks the vector $x_n \in K_n$ which minimizes $\|e_n\|_A$ where $e_n = x_* - x_n$ is the difference between x_n and the actual solution of the system $Ax = y$. How can we compute this without knowing the

solution x_*? The trick is to consider the function

$$\varphi(x) = \frac{1}{2}x^T Ax - x^T b.$$

Then using the self-adjointness,

$$\begin{aligned}
\|e_n\|_A &= (x_* - x_n)^T(Ax_* - Ax_n) \\
&= x_*^T Ax_* - x_n^T Ax_* - x_*^T Ax_n + x_n^T Ax_n \\
&= x_*^T Ax_* - 2x_n^T Ax_* + x_n^T Ax_n \\
&= x_*^T b - 2x_n^T b + x_n^T Ax_n \\
&= 2\varphi(x_n) + x_*^T b,
\end{aligned}$$

So minimzing φ is the same as minimizing $\|x_* - x_n\|_A$.

The method for doing this, due to Hestenes and Stiefel [] is surprisingly simple. Start by setting

$$x_0 = 0, r_0 = y, p_0 = r_0, \tag{32}$$

and for each $n = 1, 2, \ldots,$ set

$$\alpha_n = (r_{n-1}^T r_{n-1})/(p_{n-1}^T A p_{n-1}) = \frac{\|r_n\|}{\|p_{n-1}\|_A} \tag{33}$$

$$x_n = x_{n-1} + \alpha_n p_{n-1} \tag{34}$$

$$r_n = r_{n-1} - \alpha_n A p_{n-1} \tag{35}$$

$$\beta_n = (r_n^T r_n)/(r_{n-1}^T r_{n-1}) = \frac{\|r_n\|}{\|r_{n-1}\|} \tag{36}$$

$$p_n = r_n + \beta_n p_{n-1}. \tag{37}$$

To show that this has the claimed optimality, we need to verify two orthogonality properties:

$$r_n^T r_j = 0, \quad j < n \tag{38}$$

$$p_n^T A p_j, \quad j < n \tag{39}$$

We first claim that if $r_n \neq 0$, then the Krylov subspace K_n which, by definition, is spanned by the vectors

$$y, Ay, \ldots, A^{n-1}y,$$

is also spanned by each of the following sets:

$$x_1, x_2, \ldots, x_n,$$

$$r_0, r_1, \ldots r_{n-1},$$

$$p_0, p_1, \ldots p_{n-1}.$$

Let X_n, R_n, P_n be the spaces spanned by these vectors respectively. Then $X_1 = R_1 = P + 1 = K_n$ is clear from (32). Assuming it is true for $j < n$, we note that (**??**) implies $X_n \subset P_n$, (37) implies $P_n \subset R_n$, (**??** implies $R_n \subset K_n$. If $r_{n-1} \neq 0$, then x_n is linearly independent of X_{n-1} and hence has dimension n. Thus all for spaces are the same.

To prove (38) note that by (**??**),

$$r_n^T r_j = r_{n-1}^T r_j - \alpha_n p_{n-1}^T A r_j.$$

By our previous remarks r_j must be a linear combination of $p_0, \ldots p_j$, all of which are orthogonal to p_{n-1} by the induction hypothesis if $j < n-1$. If $j = n-1$ then it becomes

$$r_{n-1}^T r_{n-1} - \alpha_n p_{n-1}^T r_{n-1}.$$

But by (**??**),

$$\alpha_n = (r_{n-1}^T r_{n-1})/(p_{n-1}^T A p_{n-1}),$$

and by (**??**)

$$p_{n-1}^T A p_{n-1} = p_{n-1}^T A(r_{n-1} + \beta_{n-1} p_{n-2}) = p_{n-1}^T A r_{n-1},$$

so

$$\alpha_n = (r_{n-1}^T r_{n-1})/(p_{n-1}^T A r_{n-1}),$$

which implies the $r_n^T r_{n-1} = 0$.

To prove (39) we start with

$$p_n^T A p_j = r_n^T p_j + \beta_n p_{n-1}^T A p_j,$$

which is clearly 0 if $j < n-1$. If $j = n-1$ this becomes

$$p_n^T A p_{n-1} = r_n^T A p_{n-1} + \beta_n p_{n-1}^T A p_{n-1}.$$

This will be zero if

$$\beta_n = \frac{-r_n^T A p_{n-1}}{p_{n-1}^T A p_{n-1}} = \frac{-\alpha_n r_n^T A p_{n-1}}{\alpha_n p_{n-1}^T A p_{n-1}},$$

whereas by definition

$$\beta_n = \frac{r_n^T r_n}{r_{n-1}^T r_{n-1}}.$$

So if suffices to show the two numerators and two denominators are actually equal. First, by (35)

$$r_n^T r_n = r_n^T (r_{n-1} - \alpha_n A p_{n-1}) = -\alpha_n r_n^T A p_{n-1}.$$

Next, by (37) and (35),

$$\begin{aligned} r_{n-1}^T r_{n-1} &= (p_{n-1} - \beta_{n-1} p_{n-2})^T (r_n - \alpha_n A p_{n-1}) \\ &= 0 + 0 + 0 - \alpha_n p_{n-1}^T A p_{n-1}. \end{aligned}$$

This completes the proof of (39).

THEOREM 117. *If the CG iteration is applied to the system $Ax = y$ for a positive definiite matrix A, and the iteration has not already converged (i.e., $r_n \neq 0$) then x_n is the unique point in K_n that minimizes $\|x_* - x_n\|_A$. Thus this quantity decreases monotonically with n.*

PROOF. By construction $x_n \in K_n$ and since $K_n \subset K_{n+1}$ the monotonicty is obvious. Suppose $x \in K_n$. Let $e_n = x_* - x_n$ and $\Delta x = x_n - x$. Then

$$\begin{aligned} \|x_* - x\|_A &= (e_n + \Delta x)^T A (e_n + \Delta x) \\ &= e_n^T A e_n + \Delta x^T A e_n + e_n^T A \Delta x + \Delta x^T A \Delta x \\ &= e_n^T A e_n + \Delta x^T A e_n + \Delta x^T A e_n + \Delta x^T A \Delta x \\ &= \|e_n\|_A + \|\Delta x\|_A + 2 \Delta x^T A e_n. \end{aligned}$$

We claim that $Ae_n = (Ax_* - x_n) = y - Ax_n = r_n$. We can prove this by induction. For $n = 0$, we have $r_0 = y = y - A0 = y - Ax_0$, so it is true. Next suppose it is true up to $n - 1$. Then by (35), the induction hypothesis and (34),

$$r_n = y - Ax_{n-1} - \alpha_n A p_{n-1} = y - A(x_{n-1} - \alpha_n p_{n-1}) = y - Ax_n,$$

as desired

Thus by (38), $Ae_n = r_n$ is orthogonal to K_n, and so is orthogonal to Δx which is in K_n. Thus

$$\|x_* - x\|_A = \|e_n\|_A + \|\Delta x\|_A.$$

As a function of Δx this is clearly minimized by taking $\Delta x = 0$, or $x = x_n$. □

As in the GMRES iteration, we can approximate the size of e_n by considering polynomials on the spectrum of A. If the iteration has not already converged then $e_n = p(A)e_0$ for some polynomial of degree n with $p(0) = 1$. If $\sigma = \{\lambda_1, \ldots, \lambda_n\}$ is the spectrum of A (i.e., the union of eigenvalues) then by the spectral theorem we can write

$$e_0 = \sum_{k=1}^{n} a_k v_k,$$

$$e_n = P(A)e_0 = \sum_{k=1}^{n} a_k p(\lambda_k) v_k,$$

so

$$\|e_0\|_A^2 = \sum_{k=1}^{n} |a_k|^2 \lambda_k,$$

$$\|e_n\|_A^2 = \|P(A)e_0\|_A^2 = \sum_{k=1}^{n} |a_k|^2 \lambda_k |p(\lambda_k)|^2.$$

Hence

$$\|e_n\|_A \leq \|e_0\|_A \max_k |p(\lambda_k)|.$$

For example, if A has only k distinct eigenvalues, then we can choose a degree k polygnomial which vanishes at these points and is 1 at the origin, so $\|e_n\|_A = 0$, i.e., the iteration has converged.

If the spectrum lies in the interval $[1, \kappa]$ then we could set $a = \frac{1}{2}(1 + \kappa)$ and take the polynomial $p(z) = (1 - \frac{z}{a})^k$. Then $p(0) = 1$ and

$$|p(x)| \leq |\frac{z - a}{a}|^k = (\frac{\kappa - 1}{\kappa + 1})^k,$$

so we get exponential convergence.

We can do even better by using a trick. The rabbit we pull out of a hat is the Chebyshev polynomial

$$T_n(z) = \frac{1}{2}[(z + \sqrt{z^2 - 1})^n + (z - \sqrt{z^2 - 1})^n].$$

This really is a polynomial in z because expanding by the binomial theorem, every odd power of the square root on the left is canceled by a term from right, leaving

only even powers. If we set $z = \cos(\theta)$ and $w = \cos\theta + i\sin\theta$, then we get

$$\begin{aligned} T_n(\cos\theta) &= \frac{1}{2}[(\cos\theta + \sqrt{\cos^2 z - 1})^n + (\cos\theta - \sqrt{\cos^2 z - 1})^n] \\ &= \frac{1}{2}[(\cos\theta + i\sin\theta)^n + (\cos\theta - i\sin\theta)^n] \\ &= \frac{1}{2}[w^n + \bar{w}^n] \\ &= \Re(w) \\ &= \cos(n\theta). \end{aligned}$$

THus $|x| \leq 1$ implies $|T_n(x)| \leq 1$. Let $L(x)$ be the linear map that sends $[1, \kappa]$ to $[-1, 1]$ and set $p(x) = T_n(L(x))/T_n(L(0))$. Then $p(0) = 1$ and

$$\max_{[1,\kappa]} |p| = \frac{1}{T(L(0))}.$$

Since $L(x) = (x - \frac{1+\kappa}{2})\frac{2}{\kappa-1}$, we have $L(0) = \frac{\kappa+1}{\kappa-1}$. Thus

$$T_n(L(0)) = T_n(\frac{\kappa+1}{\kappa-1}).$$

A simple calculaton shows that if $z = (\frac{\kappa+1}{\kappa-1}$ then

$$z \pm \sqrt{z^2-1})^n = \frac{\kappa+1}{\kappa-1} \pm \frac{2\sqrt{\kappa}}{\kappa-1} = \frac{\kappa \pm 2\sqrt{\kappa}+1}{(\sqrt{\kappa}-1)(\sqrt{\kappa}+1)} == \frac{(\sqrt{\kappa}\pm 1)^2}{(\sqrt{\kappa}-1)(\sqrt{\kappa}+1)} =$$

Thus

$$T_n(z) = \frac{1}{2}[(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1})^n + (\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1})^n]$$

Hence

$$\max_{[1,\kappa]} |p| \leq \frac{2}{(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1})^n + (\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1})^n} \leq 2(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1})^n \leq 2(1 - \frac{1}{\sqrt{\kappa}})^n \leq 2e^{-n/\kappa}.$$

Thus any fixed degree of precision can be attained in about $O(\sqrt{\kappa})$ iterations.

3. Symm's method

If u is harmonic on $\mathbb{D}$ with continuous boundary values, then the mean value property (or the Poisson integral formula for $z = 0$) says

$$u(0) = \int_{\mathbb{T}} u(e^{i\theta})\frac{d\theta}{2\pi}.$$

If v is a harmonic funtion on a simply connected domain Ω and f is a conformal map from $\mathbb{D}$ to Ω taking 0 to z_0 then $v \circ f$ is harmonic on the disk so

$$v(z_0) = v(f(0)) = \int_{\mathbb{T}} v(f(e^{i\theta}))\frac{d\theta}{2\pi} = \int_{\partial\Omega} v(z)d\omega_{z_0},$$

where ω_{z_0} is the harmonic measure for Ω with respect to z_0.

We can discretize the integral by breaking $\partial\Omega$ into interval $\{I_j\}$ to form the linear equation

$$v(z_0) \;\; = \;\; \sum_j \int_{I_j} v(z)d\omega_{z_0}(z) \approx \sum_j \omega_j(I_j) \cdot \frac{1}{|I_j|}\int_{I_j} v(x)dx.$$

where $\omega_j = \omega_{z_0}(I_j)$ and we are assuming that the density for harmonic measure is approximately constant on each interval. If discretize into n intervals and choose n linearly indenpendent functions v, then we get n linear equations for the n unkown values $\{\omega_j\}$. Solve the system and multiply by $2pi$ to get the spacings for the Schwarz-Christoffel parameters.

What function can we choose? We discuss several possible choices of test functions.

Power functions: The most obvious collection of linearly independent harmonic functions might be $\{\Re(z^n)\}$. However, unless the domain is a circle centered at the origin, these functions will have huge oscillations over the bounadry (consider z^{100}) and we will not even attempt to use them.

Logarithmic poles: One example of harmonic function on Ω is

$$\log\frac{1}{|z-w|},$$

assuming $w \notin \Omega$, so we must have

$$\log\frac{1}{|z_0-w|} = \int_{\partial\Omega} \log\frac{1}{|z-w|}d\omega_{z_0}(z). \tag{40}$$

for any $w \notin \Omega$. If we take n such functions with n different poles, then they must be linearly independent (since a finite combination of finite values can't give a pole). See Figure 5, where we have used these functions to approximate the conformal map onto a C^1 domain. The matrix we get is almost, but not quite symmetric. If we were to replace the off diagonal elements

$$\frac{1}{|I_j|}\int_{I_j} \log\frac{1}{|a-x|}dx,$$

by the value $\log \frac{1}{|a-z_j|}$ at the center z_j of the interval I_j, then it would become symmetric.

The main concern with a method like Symms' is how diffiucult it is to solve the linear system. We will discuss this in greater detail later in the chapter, but the main idea is that the more "diagonal" the matrix looks, the easier it will be to solve. The logarithmic function does not decay very quickly, so that the matrix we get above does not decay very quickly away from the diagonal. There are several things we might try to improve this.

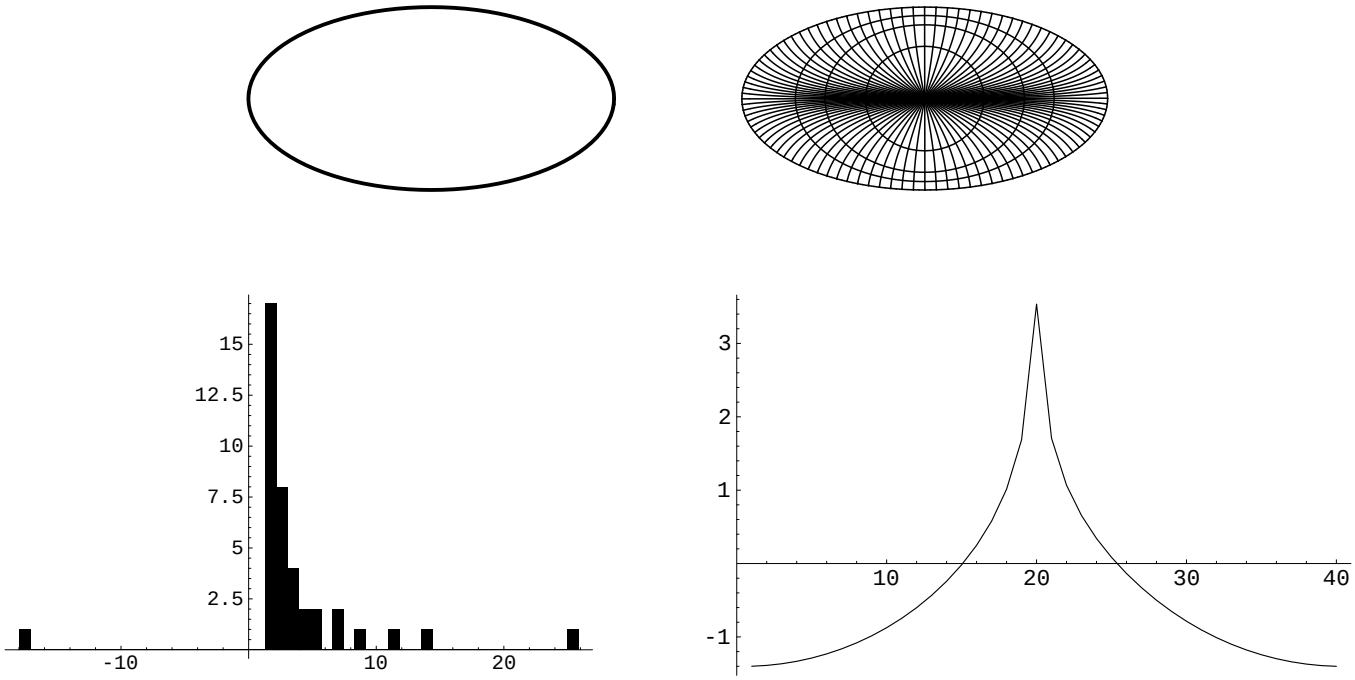


FIGURE 1. Symms method with logarithmic poles applied to an ellipse, discretized by 40 points. The top left shows the target domain and the top right is a reconstruction using a 100 term power series and the parameters given by Symms method. The second row shows a histogram of the distribution of the eigenvalues of the matrix and the graph of a singel row of the matrix. The full matrix for this example is shown in Figure 5. For this example the condition number is 18.3437. The quasiconformal distance (based on triangles all with one vertex at the origin) is 1.83272.

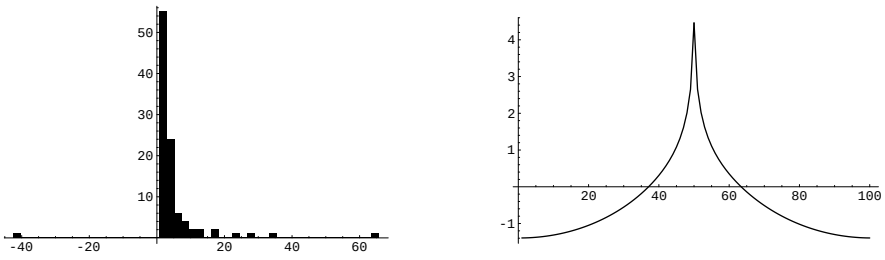


FIGURE 2. The same as the four bottom pictures in Figure 1, except that we have discritized the ellipse with 100 points. For this example the condition number is 46.9604. The quasiconformal distance (based on triangles all with one vertex at the origin) is 1.3427.

Normal dipoles: The most obvious option is to replace the single logarithmic pole by two nearby poles with opposite signs, e.g.,

$$\log \frac{1}{|z-w_1|} - \log \frac{1}{|z-w_2|} = \log \frac{|z-w_2|}{|z-w_1|},$$

which decays like $O(\epsilon|z-w|^{-1})$ away from the poles where $\epsilon = |w_1 - w_2|$. See Figure 3 for a contour plot of this function with $w_1, w_2 = \pm 1$. The function vanishes along the vertical axis; this is the set of points equidistant from both poles.

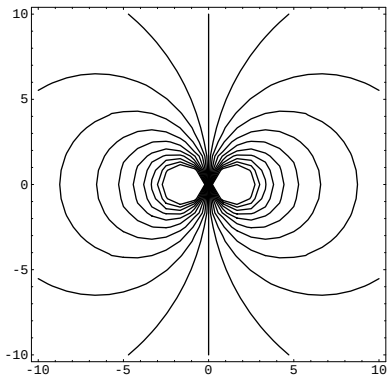


FIGURE 3. The level lines of the dipole kernel with poles at ± 1. The kernel vanishes on the perpendicular bisector of the segment connecting the poles.

So for each segment in the discretized boundary we place two logarithmic poles nearby. First consider the case when these points both lie on an outward normal to the curve. We will call this a normal dipole. The resulting kernel function vanishes along a curve parallel the boundary, as well as tending to zero far from the poles. Thus the matrix we get should be large along the diagonal and small everywhere else. See Figure **??**. This example uses poles located at the points $\frac{1}{2}(a+b)$, $\frac{1}{2}(a+b)+i\frac{1}{2}(b-a)\}$.

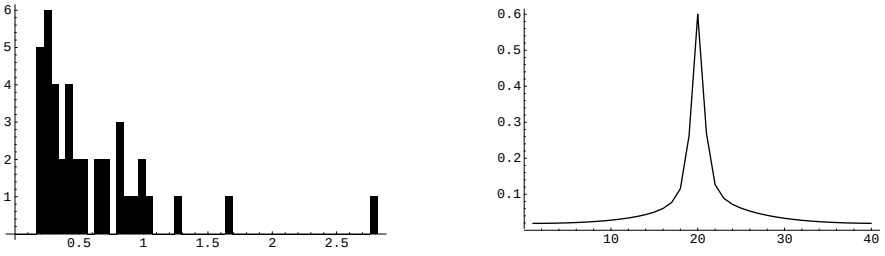


FIGURE 4. The same as the four bottom pictures in Figure 1, except that we are using a matrix derived from normal dipoles. For this example the condition number is 13.3005. The quasiconformal distance (based on triangles all with one vertex at the origin) is bounded by 1.0282.

Tangential dipoles: The other obvious arrangement of the two poles is to place them parallel to the boundary. For example, we could use the points $\{\frac{1}{3}a+\frac{2}{3}b, \frac{2}{3}a+\frac{1}{3}b\}$. Here the two poles are both on the boundary edge and the corresponding kernel function is zero on a line perpendicular to the boundary. The resultng matrix will be zero along the diagonal (since the integral of the two poles exactly cancel on the interval containing the poles) and be negative close to the diagonal on one side and positive near the diagonal on the other. Thus the matrix looks similar to an anti-symmetrix one. However, Figure **??** shows that the resulting matrix has eigenvalues very close to zero, and solving the system does not even lead to vector with positive components (which the vector of harmonic measures must be).

Subtended angles: A similar function is given by the "subtended angle", where we let

$$v_k(z) = \frac{1}{\pi}\arg(\frac{b_k - z}{a_k - z}),$$

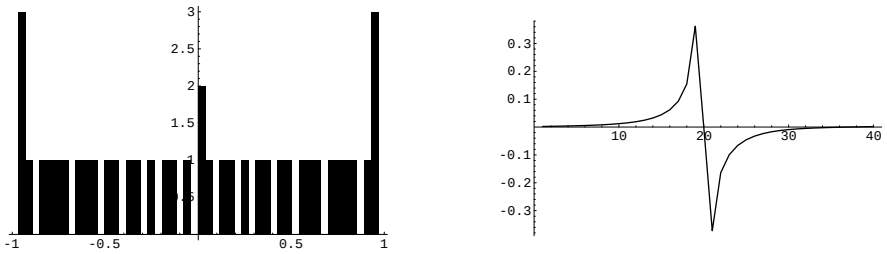


FIGURE 5. The same as the four bottom pictures in Figure 1, except that we are using a matrix derived from tangential dipoles. The resulting matrix has complex eigenvalues; there are two real eigenvalues, $\pm 7.17329 \cdot 10^{-7}$, four complex eigenvalues that have relatively large real parts, $\pm 0.00325618 \pm i 0.945374 I$, and the rest have real parts on the order of 10^{-17}. So in the plots of the eigenvalues, we have only shown the imaginary parts. The tiny eigenvalues make solving the system unreliable, and when we let *Mathematica* attempt to do so, it returns a vector with all negative components (whereas the solution should have positive components).

where $I_k = [a_k, b_k]$ (a_k comes first in the orientation of the curve). This vanishes on the line containing I_k. The general shape of these functions is simlar to the normal dipoles above, but the decay of the subtended angle functions seems a bit better and the eigenvalues of the resulting matrices seem to be more clustered, which should lead to faster solutions. Note that

$$v_k(z) = \frac{1}{\pi} \Im (\log \frac{b_k - z}{a_k - z}).$$

If we took the real part of the function instead, we would have the tangential dipole considered above (with the poles at the endpoints of the interval). Thus the subtended angle kernel is the harmonic conjugate of a tangential dipole kernel. so that this is almost the same as the tangential dipole kernels considered above.

Modified subtended angle: The subtended angle function is constant on circular arcs passing through the endpoints of I. Thus by taking the function $w_k(z) = (v_k(z) - c_k)/(1 - c_k)$, for some constant c_k, we can arrange for the our test function to still be 1 on I_k, but vanish on a circle which may be a good fit to

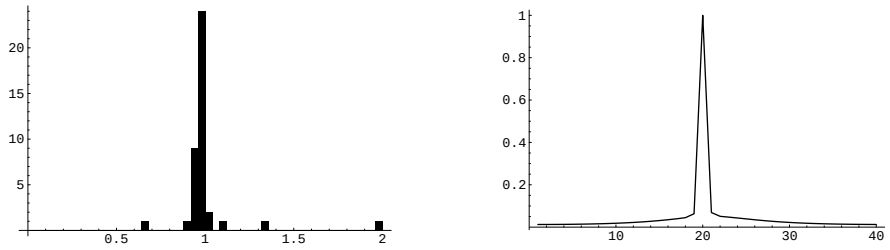


FIGURE 6. The same as the four bottom pictures in Figure 1, except that we are using the subtended angle kernel. For this example the condition number is 3.00654. However, this is mostly due to isolated outlying eigenvalues; if we remove the two largest and two smallest eigenvalues the ratio of largest to smallest of the remaining is only 1.19732. If we increase the number of points from 40 to 100, the bounds remain essentially the same. The quasiconformal distance is bounded by 1.0282.

the curve locally, if the radius of the circle is chosen to match the curvature of the boundary at each point. See Figure 5.

Quadrature: Further improvements can be made by assuming that the distribution function for harmonic measure is not constant on each interval, but is, for example, a linear function and set up two equations for each interval involving the parameters of this linear function. We will not expore this here. If the domain is a polygon, we can also make use of the information that harmonic measure behaves like a known power function near each vertex.

Multiply connected domains.

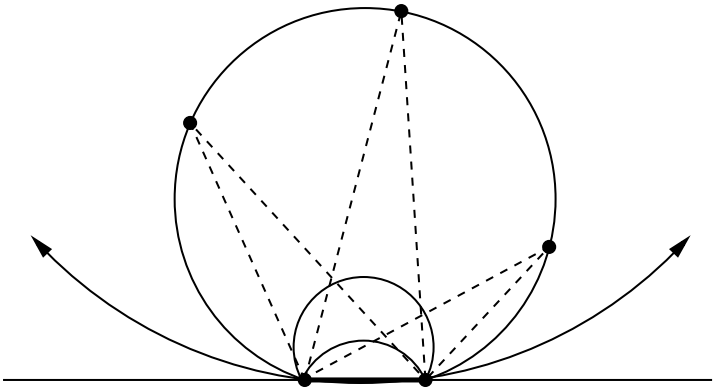


FIGURE 7. Given an interval I, that interval subtends the same angle from all points on a fixed circle through it endpoints. This means that the subtended angle function discussed in the text is constant on circles passing though the endpoints of the boundary interval. In particular, this function vanishes on the line containing this interval. If the boundary of the domain is soothly curved, then we can choose a value θ so that $v(z) - \theta$ vanishes on a circle which is a better fit to the boundary than the tangent line. This is what we called the "modified subtended angle".

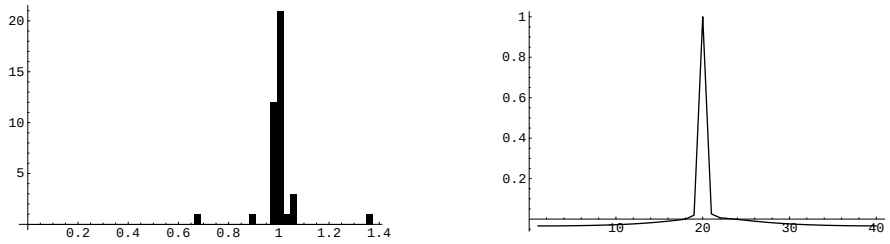


FIGURE 8. The same as the four bottom pictures in Figure 1, except that we are using the modified subtended angle kernel. For this example the condition number is 2.03454. The quasiconformal distance is bounded by 1.01931. If we increase the number of points from 40 to 100 the conditions number becomes 2.01365 and the QC error becomes 1.01049. The 40×40 matrix produced in this example has one pair of complex eigenvalues close to the real axis, $1.02879 \pm i0.0162981$, so in plotting the eigenvalues we have only used the real parts.

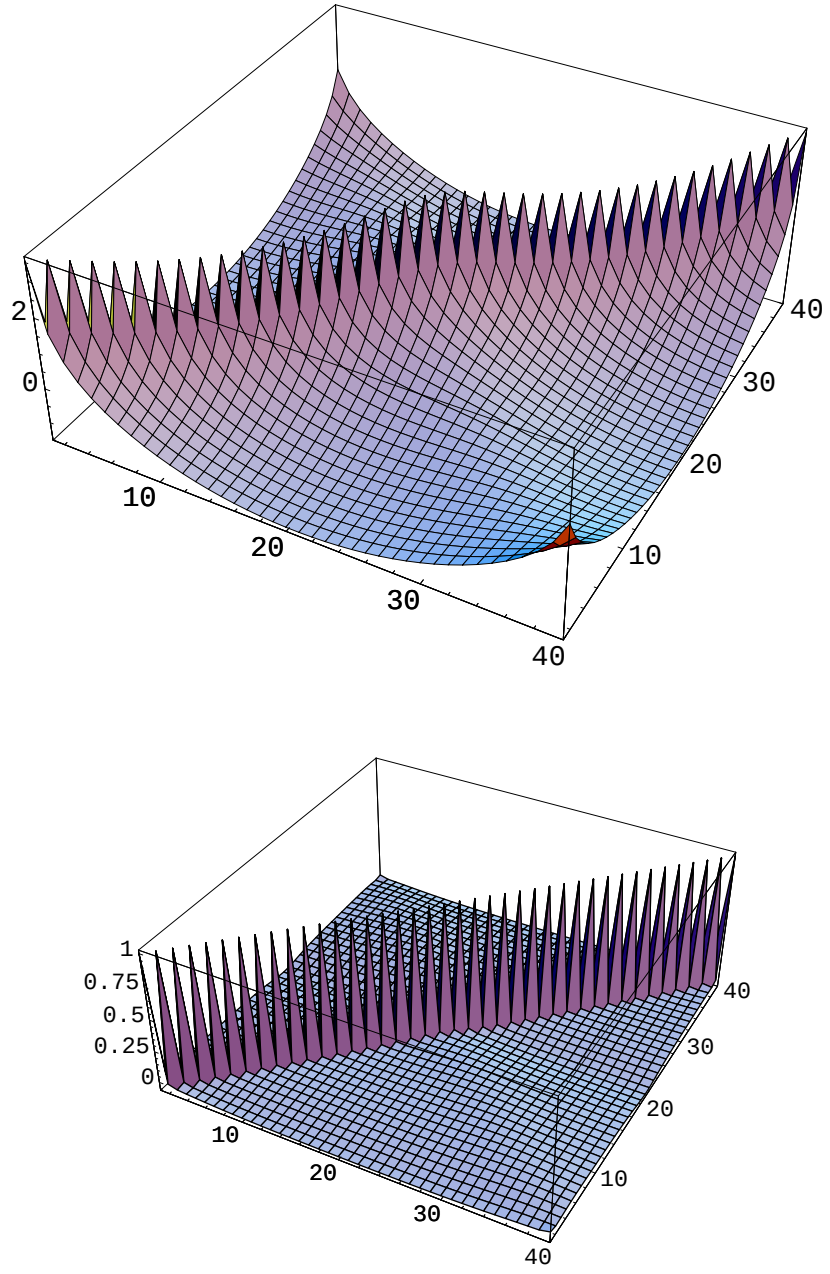


FIGURE 9. These show larger plots of the matrices that arise for the ellipse discretized with 40 points using single logarithmic poles versus the modified subtended angle functions. The second plot should look more like an identity matrix (and it does).

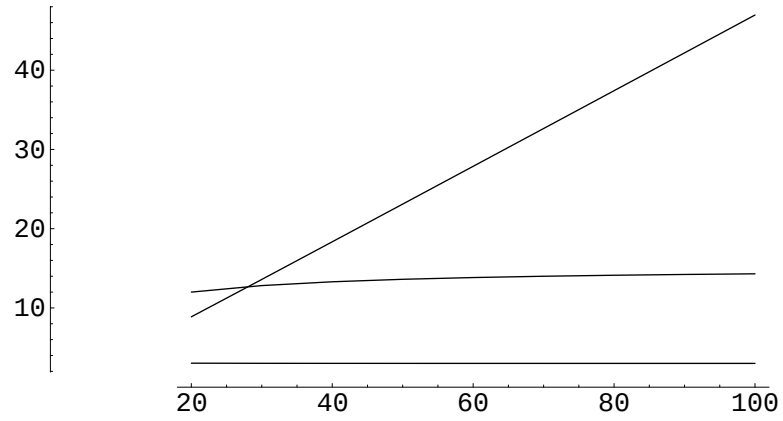
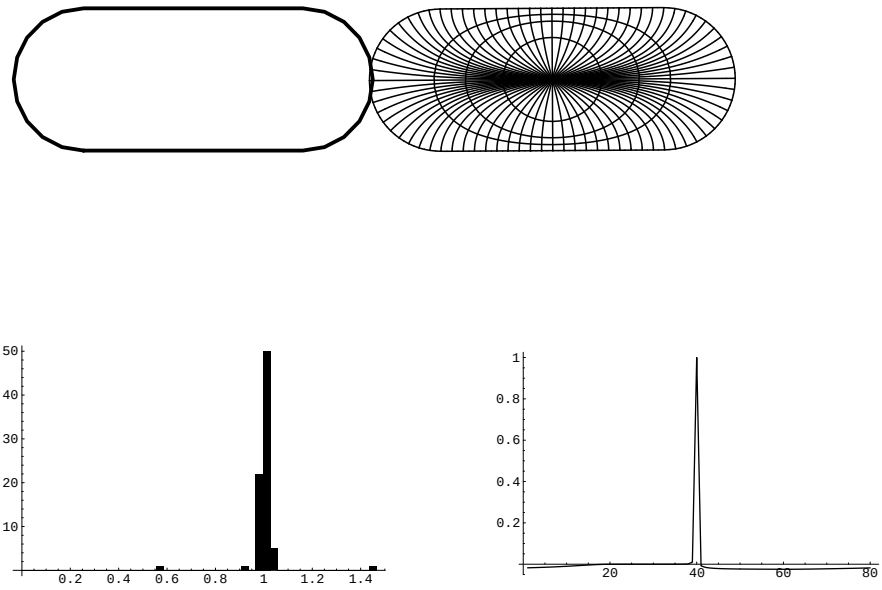


FIGURE 10. Plots of the condition numbers of the matrices corresponding to single logarithmic poles, normal dipoles and the subtended angle when the 2×1 ellipse is discretized using $20, 30, \ldots, 100$ points. For the single logarithmic pole the condition number seems to grow linearly with the size of the matrix, but for the other two methods it seems to approach a finite upper bound. In fact, for the subtended angle kernel, the condition numbers actually decrease slightly as the number of points increases.



arget domain and the top right is a reconstruction using a 100 term power series and the parameters given by Symms r

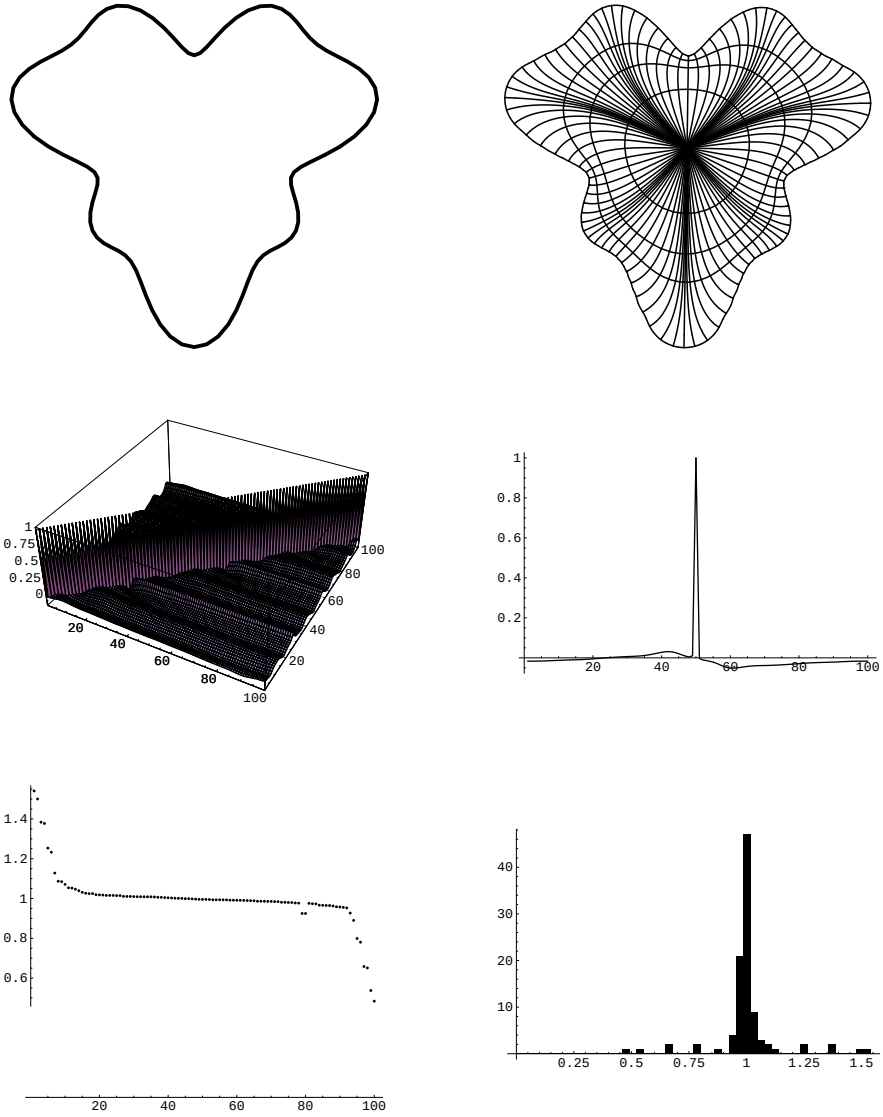


FIGURE 11. Symms method with modified subtended angle applied to the starshaped curve parameterized by

$$\theta \to \frac{1}{5}(\frac{1}{2}\sin(7\theta) + \sin(3\theta) + 4)e^{i\theta}).$$

discretized by 100 points. The top left shows the target domain and the top right is a reconstruction using a 100 term power series and the parameters given by Symms method. The middle row shows a graph of the full matrix on the left and a single row on the right. On the bottom is a plot of the eigenvalues and a histogram of these same points. For this example the condition number is 3.18457. The quasiconformal distance is bounded above by 1.18226.

4. The Kerzman-Stein formula

Suppose Γ is a smooth closed curve bounding a domain Ω and for a function f define on γ define the Cauchy integral

$$u(w) = \frac{1}{2\pi i} \int_\Gamma \frac{f(z)}{z-w} dz, \quad w \in \Omega,$$

or

$$u(z) = \int_\Gamma f(z) H(w,z) dz, \quad w \in \Omega,$$

where

$$H(w,z) = \frac{1}{2\pi i} \frac{1}{z-w} \gamma'(z),$$

where γ is the arclength parameterization of Γ and γ' denotes the tangent direction of Γ. Then u is a holomorphic funcion on Ω and if f was the boundary extension of some holomorphic function on Ω then the Cauchy integral recreates this function. This says that the map $\mathbf{H} : L^2(\Gamma) \to H^2(\Gamma)$ is a projection. (In this section, a boldface letter will indicate an operator and the non-bold version of the letter will denote the kernel which we integrate against to define the operator, i.e., H is the Cauchy kernel which is a function of two complex variables and $\mathbf{H}$ is the Cauchy projection from L^2 to H^2.)

Given a closed subspace of a Hilbert space there is an "obvious" projection: the orthogonal projections that maps each point to the closest point of the subspace. If Γ is a circle, then the Cauchy integral is the orthogonal projection, but in general (indeed, in all other cases) it is not. The orthogonal projection can also be written as in integral operator

$$Su(w) = \int_\gamma S(w,z) f(z) |dz|,$$

although in this case the formula for S is not obvious. The kernel $S(w,z)$ is called the Szegö kernel. However, in a remarakable paper [], Kerzman and Stein show that the kernel S is explicitly given in terms of the Riemann map of Ω to $\mathbb{D}$ and that it can also be written in terms of the Cauchy integral. Putting these two observations together gives a formula for the Riemann map in terms of the Cauchy integral.

First, let us explain the connection between the Szegö kernel and the Riemann map. Note that if $\{\phi_j\}$ is an orthonormal basis for $H^2(\Gamma)$, then

$$S(w,z) = \sum_{j=1}^{\infty} \phi_j(w)\bar{\phi}_j(z), w \in \Omega, z \in \Gamma,$$

since the orthogonal projection onto the span of $\{\phi_j\}$ is

$$f \to \sum_j \langle f, \phi_j \rangle \phi_j.$$

This expansion implies S can be defined on $\Omega \times \Omega$ and satisfie

$$S(z,w) = \overline{S(w,z)}.$$

If $\Omega = \mathbb{D}$ then we can take $\phi_j(z) = z^j/\sqrt{2\pi}$ and deduce that

$$S(w,z) = \frac{1}{2\pi} \sum_j w^j \bar{z}^j = \frac{1}{2\pi} \frac{1}{1 - w\bar{z}}, \tag{41}$$

by the neometric series formula.

Now fix a point $a \in \Omega$ and let $f : \Omega \to \mathbb{D}$ be a Riemann map which send a to 0 and assume $f'(a) > 0$. We claim that

$$f'(z) = \frac{2\pi}{S(a,a)} S^2(z,a), z \in \Omega. \tag{42}$$

To prove (42), note that there is a relation between the Szegö kernels S for Ω and $S_{\mathbb{D}}$ for $\mathbb{D}$.

$$S(w,z) = \sqrt{f'(w)} S_{\mathbb{D}}(f(w), f(z)) \overline{\sqrt{f'(z)}}.$$

This holds becuase if $\{\phi_j\}$ are orthonormal in $H^2(\mathbb{T}, ds)$ then $\phi_j \circ f$ are orthonormal in $H^2(\Gamma, |f'|ds)$ and hence $\{\phi_j \circ f \cdot \sqrt{f'(w)}\}$ are orthonormal in $H^2(\Gamma, ds)$. Note that we are using the fact that f' is non-vanishing and hence has a single valued square root on the simply connected domain Ω. Thus

$$\begin{aligned} S(w,z) &= \sum_j \phi_j \circ f \cdot \sqrt{f'(w)} \overline{\sqrt{f'(z)}} \\ &= \sqrt{f'(w)} [\sum_j \phi_j \circ f] \overline{\sqrt{f'(z)}} \\ &= \sqrt{f'(w)} S_{\mathbb{D}}(f(w), f(z)) \overline{\sqrt{f'(z)}}, \end{aligned}$$

as desired.

Now set $z = a$ and use $f(a) = 0$ and (41) to deduce

$$S(w, a) = \sqrt{f'(w)}\frac{1}{2\pi}\overline{\sqrt{f'(a)}}.$$

Setting $w = a$ gives

$$S(a, a) = \frac{1}{2\pi}|f'(a)|,$$

and squaring boths sides gives

$$S^2(w, a) = f'(w)\frac{1}{(2\pi)^2}\overline{f'(a)}.$$

replacing w by z, this implies (since $f'(a)$ is real),

$$f'(z) = (2\pi)^2 S^2(z, a)\frac{1}{f'(a)} = 2\pi\frac{S^2(z, a)}{S(a, a)},$$

which is the desired equality.

Therefore, if we can compute $S(a, a)$ and $S(z, a)$ explcitly, we can compute f' and then recover f by numerical integration for interior points. On the boundary, we can recover f by the formula

$$f(z) = \frac{\gamma'(z)f'(z)}{i|f'(z)|},$$

which is based on the fact that since $f(z)$ is on the unit circle, the tangent at $f(z)$ is $if(z)$.

Next we have to explain how to compute the Szegö kernel in terms of the Cauchy kernel. This will follow from writing the orthogonal projection $\mathbf{S} : L^2 \to H^2$ in terms of the Cauchy projection $\mathbf{H} : L^2 \to H^2$. Both $\mathbf{S}$ and $\mathbf{H}$ are projections onto the same subspace so

$$\mathbf{SH} = \mathbf{H}, \quad \mathbf{HS} = \mathbf{S}.$$

Since $\mathbf{S}$ is orthogonal it is also self-adjoint, i.e., $\mathbf{S} = \mathbf{S}^*$, and hence

$$\mathbf{H}^*\mathbf{S} = \mathbf{H}^*\mathbf{S}^* = (\mathbf{SH})^* = \mathbf{H}^*,$$

$$\mathbf{SH}^* = \mathbf{S}^*\mathbf{H}^* = (\mathbf{HS})^* = \mathbf{S}^*\mathbf{S}.$$

Subtracting gives

$$\mathbf{S}(\mathbf{H}^* - \mathbf{H}) = \mathbf{S} - \mathbf{H},$$

and setting $\mathbf{A} = \mathbf{H}^* - \mathbf{H}$ gives

$$\mathbf{S} = \mathbf{H}(I - \mathbf{A})^{-1}, \tag{43}$$

assuming $I - \mathbf{A}$ is invertible (we shall see below that it is). We also have the relation

$$\mathbf{H}^* = \mathbf{S} + \mathbf{H}^* - \mathbf{S} = \mathbf{S} + \mathbf{H}^*\mathbf{S} - \mathbf{H}\mathbf{S} = \mathbf{S} + \mathbf{A}\mathbf{S}. \tag{44}$$

This can be discretized and used to solve for the kernel S of $\mathbf{S}$ (since we know the kernels for $\mathbf{H}$ and $\mathbf{A}$ explicitly.

Both (43) and (44) offer a way to compute $\mathbf{S}$ from $\mathbf{H}$, although the first is only valid when $\mathbf{A}$ is sufficiently small. For example, if we define the kernel of $\mathbf{A}$ as $A(w,z)\bar{H}(z,w) - H(w,z)$, and assume that

$$\frac{1}{4\pi^2}\int_\Gamma\int_\Gamma |A(w,z)||dw||dz| < 1,$$

then $\mathbf{A}$ has operator norm < 1 and we have

$$\mathbf{S} = \mathbf{H}(I + \mathbf{A} + \mathbf{A}^2 + \dots).$$

The kernel for $\mathbf{A}$ is defined to be

$$A(w,z) = \overline{H(z,w)} - H(w,z),$$

a differece of two functions with sigularities. It turns our, however, that these singularities cancel and that $A(w,z)$ is a smooth function which vanishes on the diagonal $\{w = z\}$.

To see this first note that A is clearly C^∞ off of the diagonal. To check that it is C^∞ on the diagonal, we condsider $z = \gamma(s), w = \gamma(t)$ for t, s in some interval I and $\gamma : I \to \Gamma$ being a C^∞ parameterization. Think of $s = t + (s-t)$ and take a power series exapansion of γ around t to get

$$\gamma(s) = \gamma(t) + \gamma'(t)(s-t) + \frac{1}{2}\gamma''(t)(s-t)^2 + (s-t)^3\varphi(s,t),$$

where φ is some C^∞ function that will only appear in error terms and maybe different in different equations (note that out assumptions do not imply that the infinite power series for γ actually converges to γ). Then

$$z - w = \gamma(s) - \gamma(t) = \gamma'(t)(s-t)[1 + \frac{\gamma''(t)}{2\gamma'(t)}(s-t) + (s-t)^2\varphi,$$

and

$$\frac{1}{z-w} = \frac{1}{\gamma'(t)(s-t)}[1 - \frac{\gamma''(t)}{2\gamma'(t)}(s-t) + (s-t)^2\varphi,$$

Also

$$\gamma'(z) = \gamma'(s) = \gamma'(t) + \gamma''(t)(s-t) + (s-t)^2\varphi,$$

which gives

$$\frac{\gamma'(z)}{z-w} = \frac{1}{s-t} + \frac{\gamma''(t)}{2\gamma'(t)} + (s-t)\varphi.$$

Similarly,

$$\frac{\gamma'(w)}{z-w} = \frac{1}{s-t} - \frac{\gamma''(t)}{2\gamma'(t)} + (s-t)\varphi,$$

Since the leading terms are real and equal, they still cancel when we conjuagate and subtract, i.e.,

$$\begin{aligned} A(w,z) &= \overline{H(w,z)}H(z,w) \\ &= \frac{-1}{2pii}\overline{(\frac{\gamma'(w)}{w-z})} - \frac{1}{2\pi i}\frac{\gamma'(z)}{z-w} \\ &= \frac{1}{2pii}\overline{(\frac{\gamma'(w)}{z-w})} - \frac{1}{2\pi i}\frac{\gamma'(z)}{z-w} \\ &= \frac{-1}{2\pi i}[\Re\frac{\gamma''(t)}{\gamma'(t)} + (s-t)\varphi]. \end{aligned}$$

However, since γ is an arclength parameterization of a smooth curve, γ'' is perpendicular to γ' and hence the real part of the ratio is zero. Thus

$$A(z,w) = (s-t)\varphi(s,t),$$

for some smooth function and hence A is C^∞ and vanishes on the diagonal.

Thus integration against A is a smoothing operator, i.e., $\mathbf{A}f \in C^\infty$ for every $f \in L^2$ with an estimate $|D^\alpha f| \leq C_\alpha \|f\|_2$, for some constant depending only on Γ and α. This means $\mathbf{A}$ is a compact operator (since this estimate means the image of the unit ball is an equicontinuous family and hence every seqeuence in this image contains a uniformly convergent subsequence, whose limit is obviously bounded and hence in L^2). Moreover

$$(i\mathbf{A})^* = (i\mathbf{H}^* - i\mathbf{H})^* = -i\mathbf{H} + i\mathbf{H}^* = i\mathbf{A},$$

so $i\mathbf{A}$ is compact, self-adjoint and hence has real spectrum (e.g.,Theorem 12.26 of [**?**]). By the spectral theorem, this means that $I - \mathbf{A} = I - (-i)(iA)$ is invertible.

By the Cauchy integral formula

$$f(a) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-a}dz = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-a}\gamma'(z)|dz|.$$

This is exacly the same as writing

$$\langle f, \phi_a \rangle,$$

where

$$\phi_a(z) = \overline{\frac{1}{2\pi i} \frac{\gamma'(z)}{z-a}}.$$

Since $f \in H^2$, its inner product with any vector is the same as the inner product with the orthogonal projection of the vector into H^2. Thus

$$\langle f, \phi_a \rangle = \langle f, \mathbf{S}\phi_a \rangle.$$

Since this holds for all $f \in H^2$, and a vector is uniquely determined by its inner product with all other vectors, the right hand side of the inner product must agree with the Szegö kernel, i.e.,

$$S(z, a) = \mathbf{S}\phi_a.$$

Combined with our earlier remarks, this means

$$S(z, a) = \mathbf{H}(1 - \mathbf{A})^{-1}\phi_a,$$

and if $\|A\| < 1$, then

$$S(z, a) = \sum_{k=0}^{\infty} \mathbf{H}\mathbf{A}^k \phi_a,$$

where

$$\mathbf{H}\mathbf{A}^k \phi_a = \int_\Gamma \int_\Gamma \cdots \int_\Gamma H(z, w_1) A(w_1, w_2) \cdots A(w_k, w) \phi(w) |dw_1| d|w_2| \cdots d|w_k|.$$

The convergence is uniform on $\overline{\Omega}$.

Equation (second S eqn) offers a different way to compute the Szegö kernel. The equation

$$\mathbf{S} + \mathbf{A}\mathbf{S} = \mathbf{H}^*,$$

becomes

$$S(z, a) + \int_\Gamma A(z, w) S(w, a) |dw| = \overline{H(a, z)}.$$

We discretize this continuous equation by selecting n evenly spaced points $\{z_k\}_1^n$ along Γ and writing

$$S(z_j, a) + \frac{L}{n} \sum_k A(z_j, w_k) S(w_k, a) = \overline{H(a, z_j)},$$

where L is the length of the boundary. This is clearly a linear system of the form $Bx = y$ in which the solution x, represents values of the Szegö kernel at our chosen boundary points. We can solve this using the GMRES method discussed earlier.

It is also worthwhile taking a closer look at the matrix B that occurs here. If we use the fact that A vanishes on the diagonal and elsewhere replace A by its definition in terms of the Cauchy kernel, we get

$$S(z_j,a)+\frac{L}{n}\overline{(\frac{\gamma'(z_j)}{2\pi i})}\sum_{k\neq j}\frac{1}{\bar{z}_j-\bar{z}_k}S(w_k,a)+\frac{L}{n}\frac{1}{2\pi i}\sum_{k\neq j}\frac{1}{z_j-z_k}S(w_k,a)\gamma'(w_k)=\overline{H(a,z_j)},$$

so

$$Bx = x + aC_1x + bC_2y,$$

where a, b are the constants

$$a=-\frac{L}{n}\frac{\bar{\gamma}'(z_j)}{2\pi i},\quad b=\frac{L}{n}\frac{1}{2\pi i},$$

and C_1, C_2 are the matrices with zero diagonals and off-diagonal elements given by

$$C_1=(\frac{\bar{z}_j-\bar{z}_k}{}),\qquad C_2=(\frac{z_j-z_k}{}).$$

The matrix B is normal since it is the sum of the identity and a skew-Hermetician matrix. Moreover, it is known that this matrix has bounded condition number for large n with constants depending only on the smoothness of A, [**?**].

5. The fast multipole method

As was noted in Section 5, fast solution of an $n \times n$ linear system $Ax = y$ by iterative methods depends on two potential speedups: having to apply the matrix fewer than $O(n)$ times and being able to apply A to a vector in fewer than $O(n^2)$ operations. In this section we describe an algorithm for dealing with the latter problem; the fast multipole method of Greengard and Rokhlin. This method works best when we are given a collection of n points $\{z_k\}$ and a kernel function $K(z, w)$ and the matrix A is of the form $a_{jk} = K(z_j, z_k)$. The kernel K should have good analytic properties in the sense as a function of z it may have a pole at w, but that we can find series expansions that approximate $K(z, w)$ is balls $B(a, \frac{1}{2}|w - a|)$ or near infinity in regions $A = \{|x - w| > r\}$. The standard examples are the Cauchy kernel

$$K(z,w)=\frac{1}{z-w},$$

where we have

$$\frac{1}{z-w} = \frac{1}{(z-a)-(w-a)} = \frac{1}{z-a}\frac{1}{1-\frac{w-a}{z-a}} = \frac{1}{z-a}\sum_{k=0}^{\infty}(\frac{w-a}{z-a})^k,$$

if $|z-a| > |w-a|$ is the expansion near infinity and

$$\frac{1}{z-w} = \frac{-1}{w-a}\sum_{k=0}^{\infty}(\frac{z-a}{w-a})^k,$$

is the expansion near a. Other kernels that can be handled by the multipole method include

$$\log\frac{1}{|z-w|}, \quad \frac{1}{(z-w)^2}$$

in the plane and

$$e^{-|x-y|^2},$$

on the real line.

We want to compute $(Ax)_j = \sum_k K(z_j, z_k)x_k$, for n different indices j. Naively, this requrires $O(n^2)$ operations. However, if some of the z_j's are clustered close together, compared to their distance to z_k, then the value of $K(z_j, z_k)$ is about the same over the whole cluster. Thus the data can be grouped together and the combined effects computed simultaneously. By way of review, we first describe how the method works in an easier setting: binary trees.

Suppose T is a binary tree with vertex set V of size n, $f : V \to \mathbb{R}$ is given, and we want to evaluate the sum

$$F(v) = \sum_{w\in V\setminus\{v\}} K(v,w)f(w),$$

at every $v \in V$ where $K(v,w) = a^{\rho_T(v,w)}$ and ρ_T is the path distance in the tree. There are n inputs and n outputs and each input effects the evaluation of every output, so naively it seems that n^2 operations are required. However, all n values of F can be computed in $O(n)$ steps as follows. Choose a root $v_0 \in V$ and for any $v \in V$ let $D(v) \subset V$ be the vertices that are separated from the root by v, not including v. (i.e., its descendants). Let $\tilde{D}(v)$ be all other vertices other than v itself. Let

$$F_1(v) = \sum_{w\in D(v)} K(v,w)f(w), \quad F_2(v) = \sum_{w\in\tilde{D}(v)} K(v,w)f(w).$$

We can compute each of these functions in one pass through the tree. For F_1 start by setting $F_1(v) = 0$ for each leaf of T and proceed from the leaves to the root by setting

$$F_1(v) = a \sum_{w \in C(v)} F(w),$$

where the sum is over the children of v. This is called the "up-pass" since we start at the leaves and work towards the root. Next, we transfer values from F_1 to F_2 using an "across-pass". For each vertex where F_1 has already been evaluated by the up-pass, add $a^2 \cdot (f(v) + F_1(v))$ to $F_2(w)$, for each sibling $w \in S(v)$ of v ($w \neq v$ is a sibling of v if it has the same parent as v). Lastly, we compute F_2 using a "down-pass", by setting $F_2(v) = 0$ when v is the root (which has no sibling, so was not effected by the across-pass), and in general if $F_2(v)$ has already been computed, then for each of its children w we set

$$F_2(w) = F_2(w) + a(F_2(v) + f(v)).$$

We get the desired output by noting

$$F(v) = f(v) \cdot (F_1(v) + F_2(v)),$$

for every $v \in V$ (we can evaluate vertices in any order). Thus F has been evaluated at all n points in $O(n)$ steps.

The same method works more generally. If we are given a rooted tree T with vertex set V we turn it into a directed graph G by taking two copies V_1, V_2 of V, point all edges towards the root in V_1 and away from the root in V_2 and connect each vertex in V_1 to the copies of its siblings in V_2. Assume we have linear space X_v^1, X_v^2 for each vertex $v \in V$ and a linear map from each X_v^1 to its parent and from each X_v^2 to each of its children. Assume we also have an "across map" $A_v : X_v^1 \to X_v^2$. We define a linear map $L(w, v) : X_w \to X_v$ by composing maps along the path from w to v. Given the n values $x_v \in X_v^1$, $v \in V$ the method above evaluates all n values of

$$F(x_v) = \sum_{w \in V} L(w, v) x_w$$

in only $O(n)$ steps.

In the previous example, the linear spaces were one dimensional and the edge maps were multiplication by a. For our applications, the tree will be a tree of dyadic boxes that intersect the set $E = \{z_k\}$. To each box we will associate a finite set

of regions; either disks or disks complements. The linear spaces will be spaces of analytic functions on these regions. We will actually consider two situations: an idealized model and a finite dimensional approximation that we actually compute.

In the idealized version we consider the the space X_Q of all analytic functions on a region W (add ∞ for disk complements). If W is a disk then every such function has a power series $\sum_{k=0}^{\infty} a_n(z-a)^k$ converging in the disk and for the disk complements there is a Laurent series $\sum_{k=0}^{\infty} a_n(z-a)^{-k}$. The finite dimensional version of these spaces are the spaces X_Q^p. These consist of p term power series $\sum_{k=0}^{p} a_n(z-a)^k$ (for disks) or Laurent series $\sum_{k=0}^{p} a_n(z-a)^{-k}$ (for disk complements). There is an obvious truncation map $T : X_v \to X_v^p$ and the inclusion map $I : X_Q^p \to X_Q$.

Given any two boxes and associated regions, one of which is contained the other, we can restrict a function from the larger region to the smaller. This defines restriction maps R between the infinite dimensional spaces X_Q. For the finite dimensional analog, we define maps R^p between the spaces X_Q^p by restricting and then truncating the series expansion.

When we allow infinite expansions, then restricting an analytic function to a subdomain introduces no errors and the method described above allows us to compute series expansion for the Beurling transform in time $O(n)$ with no errors (except for filling in the initial values of the arrays). Similarly, if we restrict a power series to a smaller disk, there is no error introduced, since the restriction of a degree p polynomial is still a degree p polynomial. However, if we change the center of a Laurent expansion, then a finite expansion may become infinite and truncating to p terms causes an error (depending on p and the geometry of the regions). In this case, performing the restriction-truncation along a series of nested regions might not give the same result as restricting to the smallest domain is single step. Because of this, we have to estimate the errors at each step and show the total accumulated error along the whole path is still small.

Given a point set E with n points, assume they are contained in a unit square Q_0. Let $\mathcal{G}_n$ be the division of Q_0 into 4^n subquares of side length 2^{-n}. We refer to these as the nth generation dyadic subsquares of Q_0. The collection of all dyadic subsquares forms a tree. We let Q^* denote the parent of Q and $3Q$ the concentric square with 3 times the side length (which is the union of Q and the 8 squares of the

same generation that touch it). Given E, the set of squares which contain a point of E form an infinite subtree. We consider the finite subtree $\mathcal{T}_E$ formed by removing the squares which are smaller than the local feature distance (the distance of the square to the second closest point of E). For any kth generation square Q corresponding to a vertex of $\mathcal{T}_E$ we associate one disk complement, W_Q, concentric with Q and with boundary radius $\ell(Q)$. We also associate 27 disks $\{D_{Q'}\}$, one covering each kth generation square Q' of $\mathcal{T}_E$ that is in $3Q^* \setminus 3Q$. Note that each disk associated to Q is inside the disk complement assoicated to Q so that we can restrict from the disk complement to the disks. Also if $Q \subset Q'$ then $W_{Q'} \subset W_Q$ and $D_Q \subset D_{Q'}$.

For each leaf Q of $\mathcal{T}_E$ compute the Laurent expansion of

Example 1: The Cauchy kernel:

As an example, suppose we want to compute a discrete version of a Cauchy integral

$$\int_\gamma \frac{f(w)}{z-w} dw \simeq \sum_k \frac{f(w_k)\Delta_k}{z_k - w_k},$$

at n points.

Example 2: The logarithmic potential

Example 3: The Schwarz-Christoffel integral

6. Computing the Beurling transform

We will now introduce the elements needed to apply these general ideas to the specific problem of computing $\partial T\mu$.

We start with a partial ϵ-representation of a polygonal domain. As in Section **??**, we use this to construct a dilatation μ which is a sum $\sum \mu_k$ of terms which are supported in small squares which cover a neighborhood N_s of the boundary of our decomposition. We assume that μ is defined by reflection on the lower half-plane, so that solutions of the Beltrami equation will be real on the real line. Each μ_n is a polynomial in x and y of degree at most $O(n)$ restricted to a small square. The terms of this polynomial are of the form $z^k x^a y^b = (x+iy)^k x^a y^b$ with $0 \le k \le p$ and $0 \le a, b \le C$ where p grows depending of the desired accuracy, but C is fixed, depending only on the degrees of the piecewise polynomials used in our partition of unity associated to the decomposition $\mathcal{W}$ of our representation. Thus there are only

$O(n)$ terms to consider, not $O(n^2)$ as would be the case if all powers of x and y less than n had to be considered.

We take as our tree the collection of all Whitney squares in the upper half-plane which hit the support of μ, i.e., which hit N_s. There are $O(n)$ such, since there are $O(n)$ boundary components in our decomposition and each only hits a bounded number of Whitney boxes (for arches we only need to cover the edges of the arch, not the interior). The adjacency relation is the usual one; Q is a child of Q^* if the base of Q^* contains the base of Q and Q is maximal with this property. A given Whitney square can have zero, one or two children. Those with no children are called "leaves" of the tree. Often a child is half the size of its parent and the top edge of the child is half the bottom edge of the parent, but because of arches, there are some cases where a child is much smaller than its parent. A neighbor of a dyadic Whitney square is a distinct dyadic Whitney square of the same size which touches along the boundary. The terms "descendant" and "ancestor" have the usual meanings for a rooted tree.

For any Whitney box Q in $\mathbb{H}$ with base interval I (its vertical projection on $\mathbb{R}$) let c_Q denote the center of this base and let c_Q^j, $j = 1, \ldots, 8$ be 8 equally spaced points in I (including the right, but no the left endpoint of I). Let $A_Q = \{z : |z - c_Q| \geq \lambda |I|\}$ where we choose $\frac{1}{2}\sqrt{5} < \lambda < \frac{5}{4}$, and let $D_Q^j = \{z : |z - c_Q^j| \leq \frac{1}{4}|I|\}$. These will be called the type I and type II regions associated to Q respectively. See Figure 12. The number λ is chosen in this range so that the type I region does not intersect Q, but it does contain the type II regions of any Q' which is the same size as Q, but not adjacent to it. See Figure 13. A series expansion in terms of $(z - c_Q)^{-1}$ or $(z - c_Q^j)$ will be called type I and type II expansions respectively.

Given a Whitney box Q we can restrict μ to Q and compute type I and type II expansions for $\partial T\mu|_Q$. Since μ is a piecewise polynomial of degree $O(p)$, and there is an explicit formula for the expansion of each monomial, this can be done in time $O(p \log p)$ by the remarks in Appendix **??**.

Given a Whitney box Q and its parent Q^*, the type I region for Q contains that for Q^* and so we can take the analytic function f defined by the type I expansion for Q and compute its Laurent expansion in the type I region for Q^*. Then truncate this (infinite) series to get a type I expansion for Q^*. This is called a I-to-I conversion or

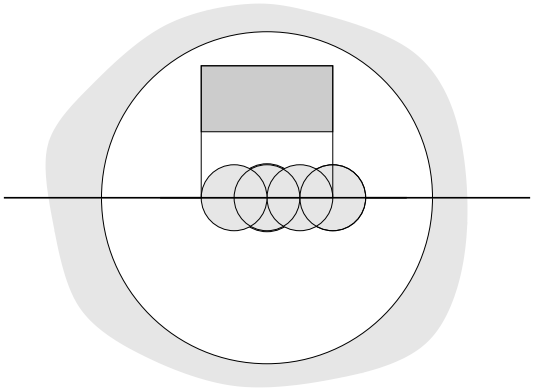


FIGURE 12. A Whitney box and its type I and type II regions.

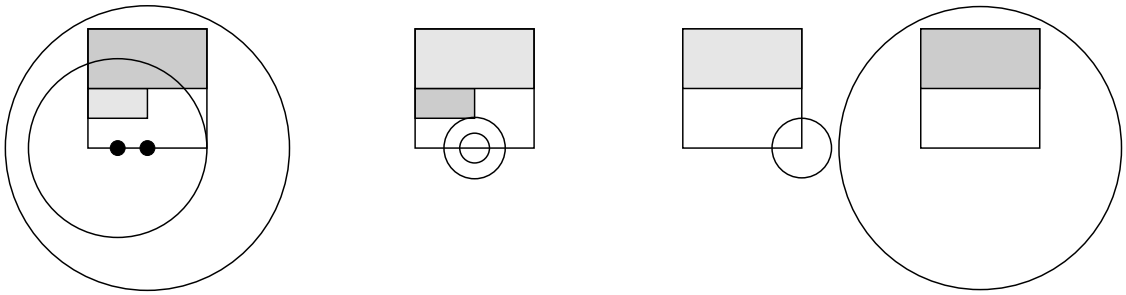


FIGURE 13. The three kinds of conversions: multipole-to-multipole, local-to-local and multipole-to-local. In each case the first Whitney square is shaded lighter than the second.

a multipole-to-multipole conversion. See Figure 13. This introduces an error of ϵM_f, where $\epsilon = \lambda^p$, where $\lambda < 1$ and M_f is the maximum of f on the type I region of Q.

Similarly, we can take a type II expansion for Q and restrict it to one of the two type II regions for a child of Q whose center agrees with the first center or is immediately to the left of it. Changing the center of the expansion just gives another degree p polynomial and there is no error introduced, i.e., $R = R^p$. This is a local-to-local conversion.

Finally, the type I region of a Whitney box Q contains the type II regions of a box Q' of the same size if Q' is in $3Q^*$ but not in $3Q$ (here Q^* denotes the parent of Q). Therefore, we can restrict the type I expansion of Q to a type II region of Q' and do a I-to-II conversion (or multipole-to-local conversion), with an error of ϵM_f, as above.

For a given box Q, the associated regions cover the whole upper half plane, except for a region of bounded hyperbolic diameter around Q.

We think of the type I and type II expansions associated to each Q as defining two arrays indexed by the Whitney boxes. We next describe how to initialize and update these arrays:

Initialize Type I array: For each Q compute the initial type I expansion.

Initialize Type II array: For each Q compute the initial type II expansion for each type II disk.

Modify Type II array: Compute type II expansions for two neighbors and add to their initial expansions. Every type II expansion now has contributions from at most three boxes (its parent and the parent's neighbors)

Perform the up-pass: Starting with leaves of the tree, do I-to-I conversion of the current type I expansion and add it to the type I expansion of the parent. Continue until we reach the root.

Perform the across-pass: For each square Q, do I-to-II conversions taking the current type I expansion and obtaining type II conversions for regions corresponding to centers in $3I^* \setminus 3I$ where I is the base of Q and I^* is the base of Q's parent.

Perform the down-pass: Starting with root square, do II-to-II conversions, taking each type II expansion and restricting it to the two type II expansions of the children. Continue downward until we reach the leaves of the tree.

This is clearly $O(n)$ steps and when we are finished, the type I expansion of a square Q contains the contribution of Q and every descendant of Q and the type II expansions of Q contain the contributions of every square which not a (strict) descendant of Q or its two neighbors.

The third step (Modify the type II array) is necessary because the tree structure on Whitney squares does not completely reflect their actual placement in $\mathbb{H}$; two squares that are far apart in the tree could be adjacent in $\mathbb{H}$. The type I region of a square does not contain the type II disks of its neighbors (they are too close), so we need this special step to pass the information to these regions (for other regions it is passed in the across-step).

Now suppose D is an empty piece of our decomposition and Q is the Whitney box containing W. If D is a disk, it is contained in a type II region of a grandparent of Q and is contained in the type I region of all grandchildren of Q and its two neighbors. Therefore we can do series conversions and compute the expansion in D due to these expansions. There are only a finite number of Whitney boxes whose contributions have not been accounted for and all these lie within a uniformly bounded distance of Q. For each piece of μ supported in one of these boxes, we compute the contribution to D directly.

Given any Whitney boxes Q' and Q there is a path in our directed graph that starts from an initial expansion for Q and goes to a terminal expansion for Q' or one of its neighbors. To see this we consider several cases.

(1) If $Q = Q'$ there is nothing to do.
(2) If $Q = Q_2$ is an ancestor of Q_1 then the all "down" path works.
(3) If Q_3 is a neighbor of square Q_2 in (1), then start with the special "Modify type II" step and follow by all downs.
(4) If Q_4 is a descendant of a case (2) square, then follow "up" paths until we hit a child of Q_3 and then use a "across" step to bring us to an ancestor of Q' (this works because by the definition of the across step).
(5) If Q is a descendant of Q', then use all "up"'s.
(6) The only remaining case is that Q is a neighbor of Q' or a descendant of a neighbor. Using an all "up" path works.

If D is an empty piece of our decomposition, we want to show that the desired expansion for $\partial T\mu$ can be computed using a bounded number or type I and type II expansions, plus a bounded number of direct expansions of nearby squares. Fix such a D and suppose Q is any Whitney box, then D is a subset of one the type I or II regions associated to Q or one of its neighbors, unless Q is within a uniformly bounded hyperbolic distance M of D. If D is a subset of one of these regions, then we can convert the series expansion on the region to one on D, the conversion being one of three types. First, we might have to convert a power series in $(z-a)$ to one in $(z-b)$; this happens when D is a disk or arch contained in a disk and involves no loss of accuracy. Second, converting an expansion in $(z-a)^{-1}$ to one in $(z-b)^{-1}$; this happens when D is an arch which contains Q in its bounded complementary

component and there is a loss of accuracy (described in Lemma 118). Finally, in all other cases we must convert an expansion in $(z-a)^{-1}$ to one in $(z-b)$; this also involves a small loss of accuracy.

LEMMA 118. *Suppose $|a| \le \frac{1}{4}$, $|b| \le \frac{1}{2}$ and f is analytic on $\{z : |z-a| > 1\}$ and $|f| \le 1$ there. Also assume $f(z)z^3$ is bounded as $z \to \infty$. Suppose $A = \{z : |z-b| > 2\}$ and let $f(z) = \sum_{j=0}^{\infty} a_j (z-b)^{-j}$ be the Laurent expansion for f in A and let $g(z) = \sum_{j=0}^{p} b_j (z-b)^{-j}$. Then there is $0 < \lambda < 1$ so that for $\epsilon = \lambda^p$,*

(1) $a_0 = a_1 = a_2 = 0$,
(2) $|g(z)| \le (1+\epsilon)|z|^{-3} \le |z|^{5/2}$, *if p is large enough.*
(3) $|f(z) - g(z)| \le \epsilon |z|^{-3}$.

The proof is just the standard estimates for Taylor series and left to the reader. If we start with an expansion f_0 on the type I region of a box Q of size 1 and then restrict the expansion to get an expansion f_1 for the type I region of its parent, then repeat this over and over, we accumulate an error each time. Suppose f_k is kth expansion on the kth region A_k. Then $\sup_{A_k} |f_k| \le |z|^{-5/2}$ which means the maximum error between $f_k(z)$ and $f_{k+1}(z)$ on A_{k+1} is bounded by $\epsilon |z|^{-5/2}$. So the total error that is ever possible inside A_N is

$$\begin{aligned} &\le \sum_{k=1}^{N} |f_k(z) - f_{k+1}(z)| \\ &\le \sum_{k=1}^{N} \epsilon \mathrm{diam}(\partial A_k)^{-5/2} (\mathrm{diam}(\partial A_n)/\mathrm{diam}(\partial A_k))^{-3} \\ &\le O(\epsilon)\mathrm{diam}(\partial A_N)^{-3} \sum_{k=1}^{N} \mathrm{diam}(\partial A_k)^{1/2}. \end{aligned}$$

Since the regions grow by at least a factor of two at each stage, the final sum is dominated by its final term, and so the total error is less than $O(\epsilon \mathrm{diam}(\partial A_N)^{-2.5}$.

Summarizing this argument gives:

LEMMA 119. *Suppose $f_0(z) = \sum_{k=3}^{p} a_k (z - c_Q)^{-k}$ is a type I expansion associated to a dyadic Whitney square Q_0 and $|f_0|$ is bounded by M on the type I region of Q_0. Suppose $Q_1, \ldots, Q_N$ are ancestors of Q and f_k is the result of applying a I-to-I*

conversion to f_{k-1} for $k = 1, \dots N$. Then on the type I region of Q_N, $|f_0 - f_N| \leq O(\epsilon M(\text{diam}(Q_0)/\text{diam}(Q_N))^{2.5})$ with $\epsilon = \lambda^p$ and a constant that is independent of N.

COROLLARY 120. *Suppose Q_0 is a Whitney square and for each Q which is a descendant of Q_0 let f_Q be the type I expansion of $\partial T\mu$ for μ restricted to Q. Let f_0 be the type I expansion for Q_0 obtained by running the up-pass overall descendants of Q with initial data $\{f_Q\}$. Let $F_0 = \sum_{Q \in D(Q_0)} f_Q$ be the exact sum of these initial expansion restricted to the type I region of Q_0. Then*

$$|f_0(z) - F_0(z)| = O(\epsilon \frac{\text{diam}(Q_0)^2}{|z - c_{Q_0}|^3} \|\mu\|_\infty).$$

PROOF. The maximum of f_Q on the type I region of Q is clearly $O(\|\mu\|_\infty/\text{diam}(Q))$. Therefore, by the Lemma 119 the error of applying I-to-I conversions until we reach Q_0 is

$$O(\epsilon \|\mu\|_\infty \text{diam}(Q)^{1.5}/\text{diam}(Q_0)^{2.5}).$$

There are at most 2^k descendants of Q_0 with $\text{diam}(Q) = 2^{-k}\text{diam}(Q_0)$, so the total error of all of these is

$$O(\epsilon \|\mu\|_\infty \text{diam}(Q)^{.5}/\text{diam}(Q_0)^{1.5}) = O(\epsilon \|\mu\|_\infty 2^{-k/2}/\text{diam}(Q_0)).$$

We now sum $k = 1, 2, \dots$ and see the total error over all descendants of Q_0 is at most $O(\epsilon \|\mu\|_\infty/\text{diam}(Q_0))$. The error is an analytic function on the type I region of Q_0 which decays like $|z - c_{Q_0}|$ near infinity (since it is a difference of functions which do), and this gives the estimate in the corollary. □

COROLLARY 121. *Let f be a type II expansion for a Whitney square Q which is obtained by starting with expansions of $\partial T\mu$ and applying the up-pass, across-pass and down-pass. Then the total error between f and simply adding all the initial expansions which contribute to f is $O(\epsilon \|\mu\|_\infty/\text{diam}(Q))$.*

PROOF. The contribution of the ancestors of Q are through II-to-II conversions, which introduce no error. The neighbors of ancestors contribute are direct computation of a type II expansion, followed by II-to-II expansions, so also contribute no error. Every other contribution comes from a sequence of I-to-I conversions (the up-pass), followed by a I-to-II conversion (the across-pass) and then II-to-II conversions (the down-pass). Fix a square Q_j to which the across-pass is applied. The errors due to all

the descendants of Q_j is bounded by $O(\epsilon\|\mu\|_\infty/\text{diam}(Q_j)$. If $\text{diam}(Q_j) = 2^j\text{diam}(Q)$, then this is $O(2^{-j}\epsilon\|\mu\|_\infty/\text{diam}(Q))$. The across-pass adds an error with the same bound and the following down-pass adds no new error. Thus the total contribution of Q_j to the error is $O(2^{-j}\epsilon\|\mu\|_\infty/\text{diam}(Q))$. There are only a bounded number (at most 4) of such Q_j's of a given size, so summing over all possible j's shows the total error is at most $O(\epsilon\|\mu\|_\infty/\text{diam}(Q))$. □

Now every expansion that contributes to the final expansion on an empty region is either directly computed from the data, or comes from an expansion created by the up-pass, across-pass and down-pass. By the corollary, the up-pass creates a small error, and we already known the across-pass creates a small error and the down-pass creates no errors. Thus the total error comes from a uniformly bounded number of terms, each of which has error bounded by $O(\epsilon\|\mu\|_\infty/\text{diam}(Q))$.

This completes the proof of Lemma **??** and hence of the theorem.

CHAPTER 10

The conjugation operator

After the Fourier trnasform, one of the most important operators in analysis is harmonic conjugation: start with a function f on the unit circle, extend it harmonically to the disk, take its harmonic conjugate and take the boundary values of this function to get $\tilde{f}$. The map $f \to \tilde{f}$ (which we will sometimes also denote by $f \to K[f]$) can be expressed in terms of Fourier series or in terms of convolution with a singular kernel. We will consider both versions and prove that this operator is an isometry on L^2 and preserves the Hölder classes of continuous functions.

The conjugation operator is the prototypical example of a singular integral operator and study of these operators and their generalizations make up a large part of modern analysis. This particular operator is important to us becuase it gives a test for a function $f = u+iv$ to be the boundary values of a holomorphic function, namely, f has a holomorphic extension iff u and v satisfy the Cauchy-Riemann equations iff

$$\tilde{f} = \tilde{u} + i\tilde{v} = i(\tilde{v} - i\tilde{u}) = i(u + iv) = if.$$

This observation can be applied to computing Riemann maps by considering methods with take an arbitary mapping f and trying to improve it by imposing the condition $\tilde{f} = if$. In this chapter we will consider two such methods; one due to Theodorsen which works for starshaped domains and another of Wegman which works for a larger class of domains and which is closely related to Fornberg's method (see Section 5).

1. Harmonic conjugates

If u is harmonic on the unit disk, a harmonic conjugate for u is a harmonic funtion v so that $f = u + iv$ is holomorphic. This means than u and v satisfy the Cauchy-Riemann equations,

$$u_x = v_y, \quad v_x = -u_y,$$

hence

$$|\nabla u|^2 = u_x^2 + u_y^2 = v_x^2 + v_y^2 = |\nabla v|^2 = 2|f'|^2.$$

Therefore, any property of u which is determined by $|\nabla u|$ alone is also true for v. Two such properties involve the L^2 norm of u on the boundary and Hölder continuity. We deal with the L^2 norm first.

LEMMA 122. *Suppose u is harmonic on $\mathbb{D}$ and continuous up to the boundary. Then*

$$\int_{\mathbb{D}} |\nabla u|^2 \log \frac{1}{|z|} dxdy = \frac{1}{2} \int_0^{2\pi} |u(e^{i\theta}) - u(0)|^2 d\theta.$$

In particular, if u and v are harmonic conjugates which both extend continuously to the unit circle, then

$$\|u - u(0)\|_2 = \|v - v(0)\|_2,$$

where the norm is taken on integrating over the unit circle.

PROOF. We may assume that $u(0) = 0$. Recall Green's theorem

$$\iint_\Omega (u\Delta v + v\Delta u) dxdy = \int_{\partial\Omega} u\frac{\partial v}{\partial n} + v\frac{\partial u}{\partial n} ds, \tag{45}$$

Here Δu denotes the Laplacian, $\Delta u = u_{xx} + u_{yy}$. We will use this with the functions $|u|^2$ and $\log \frac{1}{|z|}$ and the domain $\Omega = \mathbb{D} \setminus D(0, \epsilon)$ for some small $\epsilon > 0$. Since $\log \frac{1}{|z|}$ is harmonic over this region, left hand integral in Green's theorem is just

$$\iint \Delta |u|^2 \log \frac{1}{|z|} dxdy = \iint 2|\nabla u|^2 \log \frac{1}{|z|} dxdy.$$

There are two boundary integrals; one over the unit circle and one over the circle of radius ϵ. Since $\log |z|^{-1} = 0$ on the unit circle, and has normal derivative 1 there, the first boundary integral is just

$$\int |u(e^{i\theta})|^2 d\theta.$$

On the other boundary component, $\log |z|^{-1}$ is equal to $\log \frac{1}{\epsilon}$ and has normal derivative $-\frac{1}{\epsilon}$. On this boundary, the $|u^2| = O(\epsilon^2)$ and its normal derivative is $O(|u||\nabla u|) = O(\epsilon)$ (since u is a smooth function vanishing at zero). Moreover, the length of this boundary is $2\pi\epsilon$). Thus the boundary integral over the ϵ circle is $O(\epsilon^2 \log \frac{1}{\epsilon}) +$

$O(\epsilon^2 \frac{1}{\epsilon}) = o(1)$. Thus the contribution of this boundary component is zero as $\epsilon \to 0$. Putting all these together we get

$$2 \iint |\nabla u|^2 \log \frac{1}{|z|} dxdy = \int |u(e^{i\theta})|^2 d\theta,$$

which is the desired estimate. □

The equality above remains true even if the boundary values of u are not continuous; they only need to be in $L^2(\mathbb{T}, d\theta)$; one uses the case above on a slightly smaller disk and then takes a limit in an appropriate way. For our purposes however, the continuous case will be enough. A more subtle points concerns the final claim of the lemma, where we assumed that both u and v have continuous extensions to the boundary. It would be more convenient to assume that u has such an extension, and deduce the continuity of v from that. However, this is not true. Consider the conformal map $f = u_i v$ from the disk onto the region $\Omega = \{(x, y) : -\infty < x < \infty, 0 < y < (1+x^2)^{-1}$. The imaginary part of f if clearly bounded, and in fact is continuous (there are only two quetionable points and $\Im(f)$ tends to zero at both of them). The real part of f is unbounded however, and hence not continuous. Thus u is a harmonic conjuagte of a continous function which is not itself continuous. See Figure 5

FIGURE 1 FIGURE 2

One might object that u really is continuous if we allow the values $\pm\infty$ and give $[-\infty, \infty]$ the obvious topology. However, one can build a conformal map onto a bounded domain with non-locally connected boundary for which the the real part of the conformal map is continuous, but the imaginary part is not. See Figure 5. One can even build a conformal map where the real part is continuous and the imaginary part does not have a continuous extension to any boundary point! Unfortunately, taking harmonic conjugation simply does not preserve the property of having continuous boundary values.

However, conjugation does preserve cetain types of continuity. We say that a function f is α-Hölder if there is a $C < \infty$ so that

$$|f(x) - f(y)| \leq C|x - y|^\alpha.$$

More generally, the modulus of continuity of a function f is an increasing function η on $[0, \infty)$ so that

$$|x - y| \leq t \quad \Rightarrow \quad |f(x) - f(t)| \leq \eta(t).$$

Thus Hölder functions have modulus of continuity which is $O(t^\alpha)$. It turns out that if the boundary values of a harmonic function have a "good enough" modulus of continuity then the harmonic conjugate also has continuous boundary values. The sharp result is that the modulus of continuity satisfies a Dini condition

$$\int_0^1 \eta(t)\frac{dt}{t} < infty.$$

For our purposes, we will deal with a stronger requirements that the function is Hölder.

LEMMA 123. *Suppose u is a harmonic function on $\mathbb{D}$ and $0 < \alpha < 1$. Then u extends continuously to a α-Hölder funtion on the boundary iff*

$$\sup_{z\mathbb{D}} |\nabla u|(1-|z|)^{1-\alpha} < \infty. \tag{46}$$

This holds for u iff it also holds for any harmonic conjugate.

PROOF. First suppose (46) holds. Then if z, w lie on the same radius with $|z| < |w|$,

$$|u(z) - u(w)| \le \int_{|z|}^{|w|} |\nabla u(tz/|z|)|dt = O(|z|^\alpha),$$

and if $|z| = |w|$ then

$$|u(z) - u(w)| \le O(|w - z|(1-|z|)^{\alpha-1}).$$

Together these imply that the image of Carleson square with base I has diameter at most $O(|I|^\alpha)$, which implies u extends to be α-Hölder on the closed disk (and in particular on the boundary).

For the other direction, suppose u has α-Hölder boundary values f. Suppose $z, w \in \mathbb{D}$ are hyperbolic distance ≤ 1 apart. Then there is a Möbius transformation τ that maps z to w and moves a point $e^{i\theta} \in \mathbb{T}$ by at most $O(|z - e^{i\theta}|)$ (in the Eucludean metric). Let P_z denote the Poission kernel, $I \subset \mathbb{T}$ the arc centered at $z/|z|$ of length

$1-|z|$ and λI the centric arc of length $\lambda|I|$ (or the whole circle if $\lambda|I| \geq 2\pi$). Then

$$\begin{aligned}
|u(z)-u(w)| &= |\int_{\mathbb{T}} f(e^{i\theta})P_z(e^{i\theta})d\theta - \int_{\mathbb{T}} f(e^{i\theta})P_z(e^{i\theta})d\theta| \\
&= |\int_{\mathbb{T}} [f(e^{i\theta}) - f(\tau(e^{i\theta})]P_z(e^{i\theta})d\theta| \\
&\leq \int_I |f(e^{i\theta}) - f(\tau(e^{i\theta})|P_z(e^{i\theta})d\theta \\
&\qquad + \sum_{n=1}^{\infty} \int_{2^n I \setminus 2^{n-1} I} |f(e^{i\theta}) - f(\tau(e^{i\theta})|P_z(e^{i\theta})d\theta \\
&\leq O((1-|z|)^\alpha) + \sum_{n=1}^{\infty} O(2^n(1-|z|))^\alpha 2^{-n}) \\
&= O((1-|z|)^\alpha) \sum_{n=1}^{\infty} 2^{\alpha-1} \\
&= O((1-|z|)^\alpha)
\end{aligned}$$

Thus u is a harmonic function which oscilates by less than $O((1-|z|)^\alpha)$ on a Euclidean ball of radius $\approx (1-|z|)$ around z. Thus

$$|\nabla u(z)| = O(\frac{(1-|z|)^\alpha}{1-|z|}) = O((1-|z|)^{\alpha-1}).$$

□

The same proof works for Dini continuous. Note that the proof breaks down for $\alpha = 1$ because the sum $\sum_{n=1}^{\infty} 2^{\alpha-1}$ diverges. In this case, we should note that the infinite series really only has $O(|\log(1-|z|)|)$ terms in, each of which is $O(1)$. Thus if u is Lipschitz (i.e., 1-Hölder), the conjugate funtion need only have modulus of continuity $\eta(t) = t \log \frac{1}{t}$, and examples such as $u(e^{i\theta}) = |\theta|$ show this is sharp.

Every continuous function on the unit circle, $\mathbb{T}$, can be extended to a harmonic function on the unit disk using the Poission integral formula

$$u(z) = \int_{\mathbb{T}} f(e^{i\theta})P_z(\theta)d\theta,$$

where

$$P_z(\theta) = \frac{1}{2\pi}\frac{1-|z|^2}{|e^{i\theta}-z|^2} = \frac{1-|z|^2}{1-2|z|\cos(\theta - \arg(z)) + |z|^2} = \Re(\frac{e^{i\theta}+z}{e^{i\theta}-z}).$$

The last formula is the real part of a holomorphic function and hence the harmonic conjugate of the Poission kernel is the imaginary part of this function,

$$Q_z(\theta) = \Im(\frac{e^{i\theta}+z}{e^{i\theta}-z}) = \frac{2|z|\sin(\arg(z)-\theta)}{1-2|z|\cos(\arg(z)-\theta)+|z|^2}.$$

If $z = r \nearrow 1$ along the positive real axis then this becomes

$$\lim_{r\to 1} \frac{2r\sin\theta}{1-2r\cos\theta+r^2} = \frac{\sin\theta}{1-\cos\theta} = \cot\frac{1}{2}\theta.$$

Using the double angle formulas $\sin 2\phi = 2\sin\phi\cos\phi$ and $\cos 2\phi = \cos^2\phi - \sin^2\phi$ with $\phi = \theta/2$ converts the right hand side to

$$= \frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{1-\cos^2\frac{\theta}{2}+\sin\frac{\theta}{2}} = \frac{\cos\theta/2}{\sin\theta/2} = \cot\theta/2.$$

Thus the harmonic conjugation operator on the circle should given by

$$\tilde{f}(e^{i\theta}) = \frac{1}{2\pi}\int \cot(\frac{\theta-\phi}{2}) f(\phi) d\phi. \tag{47}$$

However, the integral in this formula is not defined; $\int_0^\pi |\cot\theta| d\theta = \infty$ since $\cot\theta$ has a $1/\theta$ singularity at the origin. Thus the integrand is not integrable in general. We resolve this in the usual way, by interpreting the integral as a principle value

$$\tilde{f}(e^{i\theta}) = \lim_{\epsilon} \frac{1}{2\pi}\int_{|\phi-\theta|>\epsilon} \cot(\frac{\theta-\phi}{2}) f(\phi) d\phi,$$

when this limit exists. If f is merely continuous, then this limit need not exist for every θ (although it is a theorem that it exists almost everywhere, even if f is only integrable). However, if f is Hölder continuous then the principle value exists everywhere and agrees with the boundary values of a harmonic conjugate of f's harmonic extension. More precisely,

LEMMA 124. *Suppose $0 < \alpha < 1$ and f is α-Hölder continuous on $\mathbb{T}$. Let u be its Poisson extension to the disk and let v the harmonic conjugate of u which vanishes at the origin. Then*

$$\tilde{f}(e^{i\theta}) = v(e^{i\theta}),$$

in the sense that the left hand side exists as a principle value at every point and the right hand side is the continuous extension to the boundary given by Lemma 123.

PROOF. Since $\cot\theta$ is odd,

$$\int_{\epsilon<|\phi-\theta|<\delta} \cot(\frac{\theta-\phi}{2})d\phi = 0,$$

and so by multiplying this by $f(\theta)$ and subtracting gives

$$|\int_{\epsilon<|\phi-\theta|<\delta} \cot(\frac{\theta-\phi}{2})f(\phi)d\phi| = |\int_{\epsilon<|\phi-\theta|<\delta} \cot(\frac{\theta-\phi}{2})[f(\phi)-f(\theta)]d\phi|$$

$$= O(\int_{\epsilon}^{\delta} t^{\alpha-1}dt = O(\delta^{\alpha}).$$

This implies the principle value exists. Moreover,

$$|\tilde{f}(e^{i\theta}) - \frac{1}{2\pi}\int_{|\phi-\theta|>\epsilon} \cot(\frac{\theta-\phi}{2})f(\phi)d\phi| = O(\epsilon^{\alpha}).$$

Since $Q_z(e^{\imath\phi}) = O(|\theta-\phi|^{-1})$ we also have

$$|v(re^{i\theta}) - \int_{|\phi-\theta|>\epsilon} Q_z(e^{i\phi})f(e^{i\phi})d\phi| = O(\epsilon^{\alpha},$$

with a constant indpendent of r. Finally, since f is bounded on $\mathbb{T}$ and $Q_r(\theta) \to \cot\theta/2$ uniformly away from 1, we have

$$|\int_{|\phi-\theta|>\epsilon} [\cot((\theta-\phi)/2) - Q_z(e^{i\phi})]f(e^{i\phi})d\phi| \to 0.$$

Putting together these inequalities, we get $|\tilde{f}(e^{i\theta}) - v(re^{i\theta})| \to 0$ as $r \to 0$. □

It is worth noting that we have placed extremely strong restrictions on the boundary functions in order to have both the Poission extension and its harmonic conjugate have continuous boundary values. If we start merely with an integrable function f on the boundary, then there is no difficulty defining a harmonic extension u on the disk using the Poission integal formula. If f is continuous on $\mathbb{T}$, then u will extend continuously to $\overline{\mathbb{D}}$ and agree with f on $\mathbb{T}$. For a general $f \in L^1$, u need not have a continuous extension anywhere on the boundary, but it does have radial limits that agree with f almost everywhere, i.e.,

$$\lim_{r\nearrow 1} u(re^{i\theta}) = f(e^{i\theta}),$$

except on a set of Lebesgue measure zero. Similarly, if $f \in L^1$, then the harmonic conjugare v of u has radial limits almost everywhere and these agree with the principle value in (47), which we can also show exists almost everywhere. In partciular, if $f \in L^1$, then $\tilde{f}$ is a well defined function almost everywhere on the circle. These

results are a little more difficult and require the Hardy-Littlewood maximal theorem. Since $L^p(\mathbb{T}) \subset L^1(\mathbb{T})$, we see that we can define $\tilde{f}$ for any $f \in L^p$. It is a famous result of M. Riesz that

$$\|\tilde{f}\|_p \leq C_p \|f\|_p, 1 < p\infty,$$

although the result fails for $p = 1, \infty$ (we have aready seen a continuous function f so that $\tilde{f}$ is unbounded. There are substitute results at the endpoints. If $f \in L^1$, then $\tilde{f}$ is in a space called weak L^1, i.e., it satisfies the inequality

$$|\{\theta : |\tilde{f}(e^{i\theta})| > \lambda\}| = O(\frac{\|f\|_1}{\lambda}).$$

if $f \in L^\infty$, then $\tilde{f}$ is in a space called BMO (for Bounded Mean Oscillation) defined by

$$\|f\|_{BMO} = \sup_I m_I(f - m_I(f)) < \infty,$$

where $m_I(f)$ is the mean value of f over I,

$$m_I(f) - \frac{1}{|I|}\int_I f(x)dx.$$

This space is of fundatmental importance in the study of singular integrals, such as the conjugate operater, althoug we shall not say much about it here.

The conjugate function has a simple expersssion in terms of Fourier series. Note that

$$P_r(\theta) + iQ_r(\theta) = \frac{1 + re^{i\theta}}{1 - re^{i\theta}} = 1 + 2\sum_{n=1}^{\infty} r^n e^{in\theta}.$$

Since P_r and Q_r are real valued

$$P_r = 1 + 2\sum_{n=1}^{\infty} r^n \Re(e^{in\theta}) = 1 + 2\sum_{n=1}^{\infty} r^n \frac{1}{2}[e^{in\theta} + e^{in\theta}],$$

and hence

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}.$$

Similarly,

$$Q_r = 2\sum_{n=1}^{\infty} r^n \Im(e^{in\theta}) = 1 + 2\sum_{n=1}^{\infty} r^n \frac{1}{2i}[e^{in\theta} - e^{in\theta}],$$

so,

$$Q_r(\theta) = \sum_{n\neq 0}(-i)\frac{n}{|n|} r^{|n|} e^{in\theta}.$$

Suppose that f on the unit circle has Fourier series

$$\sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

If f is real valued, then $a_{-n} = \overline{a_n}$. Thus the Poisson extension u of f is given by

$$u(re^{i\theta}) = P_r * f(\theta) = \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta},$$

and the harmonic conjugate v which vanishes at 0 is

$$v(re^{i\theta}) = Q_r * f(\theta) = \sum_{n\neq 0} \frac{n}{|n|} a_n r^{|n|} e^{in\theta}.$$

Thus the Fourier series of $\tilde{f}$ is

$$\sum_{n\neq 0} \frac{n}{|n|} a_n e^{in\theta}.$$

Even if f and $\tilde{f}$ are continuous, the Fourier series need not converge everywhere (one of the great theorems of the twentieth century is Carleson's theorem that the Fourier series of a continuous function converges pointwise to that function almost everywhere). However, the theory of Fourier series gives several methods to recover a continuous function from its Fourier series, such as Cesaro summation.

2. Theodorsen's method

We assume that the domain Ω is starshaped with respect to the origin. This means that the boundary can be parameterized by a function in polar coordinates, i.e.,

$$\gamma(z) = \rho(\arg(z))z,$$

for some positive, 2π-periodic function ρ on $[0, 2\pi$. The mapping $\gamma : \mathbb{T} \to \partial\Omega$ is not necessarily holomorphic there is a homeomorhism $h : \mathbb{T} \to \mathbb{T}$ so that $\gamma \circ h$ are the boundary values conformal map (just take $h = \gamma^{-1} \circ f$). But how do we compute h without knowing the conformal map f?

Think of the homeomorphism h as a perturbation of the identity, i.e.,

$$h(e^{i\theta}) = \exp(i\theta + i\delta(\theta)).$$

If $\gamma \circ h$ has a holomorphic extgension f to the the disk, and this extension vanishes at the orgin then $f(z)/z$ is non vanishing. Thus $g(z) = \log(f(z)/z)$ is a well defined

holomorphic function on the disk and its boundary values satisfy $\Im g = K(Reg)$. Here we are using K to denote the conjugation operator, $Kf = \tilde{f}$. But unwinding the definitions this becomes

$$\delta(\theta) = K(\log|\rho(\theta + \delta(\theta)|),$$

which is Theodorsen's equation. Thus the desired h is $\theta + \delta(\theta)$ where δ is a fixed point of the map

$$\delta \to K(\log|\rho(\theta + \delta(\theta)|),$$

which we find by starting with the identity $\delta_0(\theta) = \theta$) and iterating

$$\delta_{k+1} = K(\log|\rho(\theta + \delta_k(\theta)|).$$

This iteration will converge to the fixed point if, for example, the map is a contraction map. The conjugation operator is an isometry on L^2, so this occurs if

$$\|\rho(\theta + \delta_1(\theta) - \rho(\theta + \delta_2(\theta)\|_2 < \lambda\|\delta_1 - \delta_2\|_2,$$

for some $\lambda < 1$. If ρ is smooth then by the mean value theorem

$$\rho(\theta + \delta(\theta)) = \rho(\theta) + \rho'(\phi)\delta(\theta),$$

for some ϕ and hence

$$\|\rho(\theta + \delta_1(\theta) - \rho(\theta + \delta_2(\theta)\|_2 \leq 2(\max\rho')\|\delta_1 - \delta_2\|_2.$$

In particular, if ρ' is small (and so if ρ is close to constant), the iteration will converge. I do not know if anything is known about convergence in general.

The iteration can be performed by discreitizing $[0, 2\pi]$ into n points and using a discrete aproximation to the convolution in 5) or using a FFT to find the Fouier expansion of $\log\rho(\theta + \delta_k(\theta))$ and then using (5) and an inverse FFT to compute δ_k (which is faster).

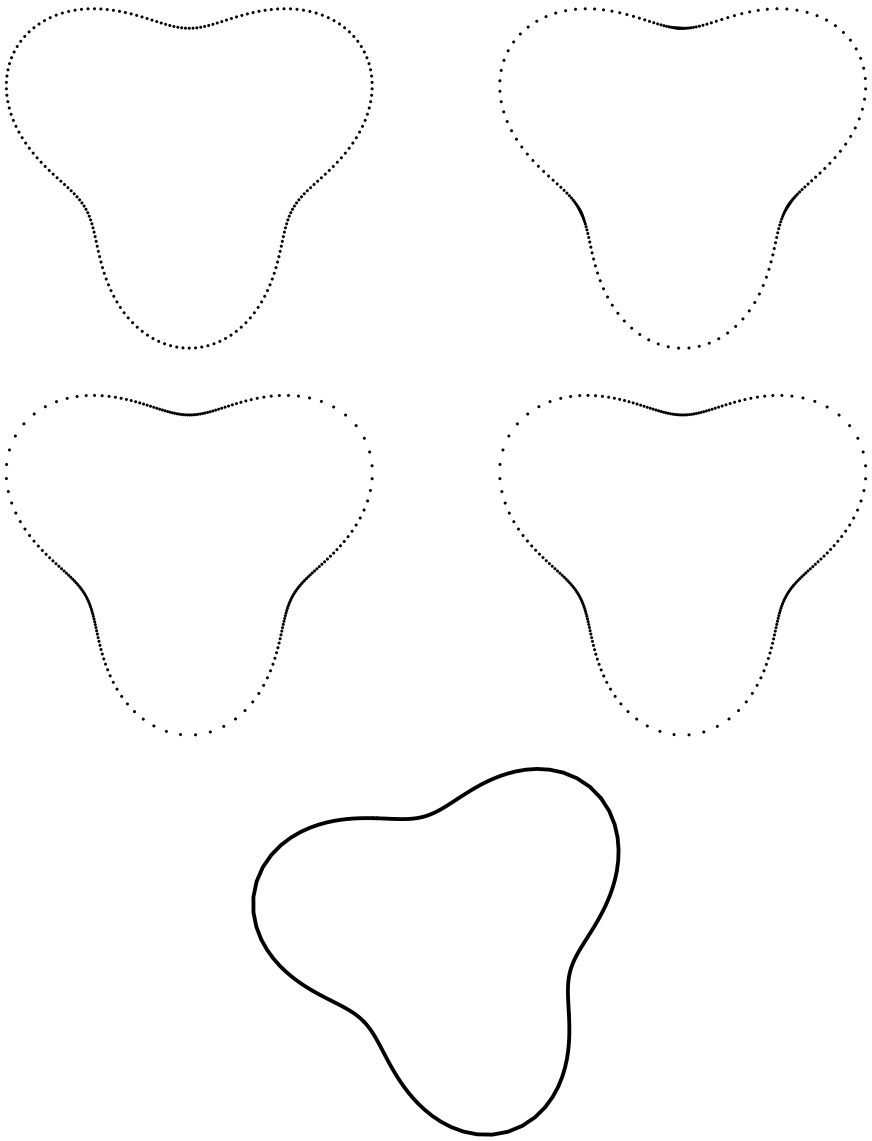


FIGURE 1. The top left shows the intial guess in Theodorsen's method; their are 200 points spread along the curve γ with equally spaced arguments. The top right figure shows the result after 1 step of the iteration. The next two figures show the results after 10 and 100 steps. The bottom figure shows the Schwarz-Christoffel image where we used the 200 equally spaced parameters on the unit circle and angles derived from the 100th iteration. If we have computed the images of these points correctly, the SC image should agree closely with the Theodorsen solution (which it does).

3. Fornberg's method

Suppose Ω is bounded by a smooth curve Γ which is parametreized by a map $\gamma : \mathbb{T} \to \Gamma$. Given N, Fornberg's method is a way to approximate $(N/2)$ terms of the power series of the conformal map onto Ω, by approximating the points on Γ where the Nth roots of unity are mapped. This is done by a iterative procedure, where each step of the iteration requires solving an $N \times N$ linear system. Since the method approximates the power series, it can only be applied for domains and values of N where we expect this to give a good approximation to the map. However, in these cases Fornberg's algorithm has several nice properties:

(1) Once we are near a solution, the convergence of the iteration is quadratic.
(2) The linear system one needs to solve has well clustered eigenvalues, so an iterative method such as the conjugate gradient method works well.
(3) The matrix in the linear system is a combination of diagonal matrices and the discrete Fourier matrix, so can be applied using FFT's in $N \log N$ time.

As in Theodorsen's method, we know there is a homeomorphism $h : \mathbb{T} \to \mathbb{T}$ so that $\gamma \circ h$ are the boundary values of a conformal map $f : \mathbb{D} \to \Omega$ and we wish to find such an h. More specifically, we seek to evaluate h at the Nth roots of unity. Fornberg's method is to start with an inital guess for the values of h at the roots of unity and then set up a system of equations to find correction terms. The initial system is replaced by a linear approximation which is then solved. The process can be iterated and the will give quadratic convergence to the correct values in favorable cases. add corrections terms image points.

What is this system of equations. The basic idea is to take the Discrete Fourier Transform of the current guess. This produces a polynomial of degree $N-1$ so that

$$z_k = \sum_{j=0}^{N-1} c_j (w_k)^j, k = 0, \ldots N-1,$$

where $\{w_k\}$ are the Nth roots of unity and $\{z_k\} = \{\gamma(h(w_k))\} \subset \Gamma$ are the current guesses for the conformal images of the roots of unity. By definition, an Nth root of unity satisfies $w_k^N = 1$, so $w_k^j = w_k^{j-N}$. Using this, we will convert our polynomial

into a Laurent series by converting the $N/2$ highest degree terms to negative terms

$$z_k = \sum_{j=-(N/2)+1}^{0} d_j w_k^j + \sum_{j=1}^{N/2} c_j (w_k)^j,$$

where $d_j = c_{j+N}$ for $j < 1$ and $d_0 = c_0$. If this Laurent series is supposed to represent a homolomorphic map, then all the negative powers should have coefficient zero. We also want to set the constant term, $d_0 = c_0$ equal to zero in order for the map to send 0 to 0.

How do we modify h so that the negative coefficients become zero? We can use an inverse discrete Fourier transform to compute these coefficients as

$$d_k = \frac{1}{N} \sum_{j=0}^{N-1} z_j w_j^{-k}.$$

We want to move the z_k along Γ so as to make all these sums zero for non-positive indices k. Instead of moving z_k along Γ, think of moving it along the tangent line to Γ at z_k. Thus our new guess for z_k will be of the form $z_k + \gamma'(z_k)t_k$, where t_k is a real number. We want to choose the N values $\{t_k\}_0^{N-1}$ so that

$$0 = \frac{1}{N} \sum_{j=0}^{N-1} (z_j + \gamma'(z_j)t_j) w_j^{-k},$$

or

$$d_k = -\frac{1}{N} \sum_{j=0}^{N-1} \gamma'(z_j) t_j w_j^{-k}.$$

This is a linear system of the form

$$y = At,$$

where y is a $N/2$ complex vector, A is a $(N/2) \times N$ matrix and t is a N real vector. More precisely we have

$$\begin{pmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{1-N/2} \end{pmatrix} = -\frac{1}{N} \begin{pmatrix} \gamma_0' & \gamma_1' & \gamma_2' & \cdots & \gamma_{N-1}' \\ \gamma_0' & \gamma_1' w & \gamma_2' w^2 & \cdots & \gamma_{N-1}' w^{N-1} \\ \gamma_0' & \gamma_1' w^2 & \gamma_2' w^4 & \cdots & \gamma_{N-1}' w^{2(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_0' & \gamma_1' w^{N/2-1} & \gamma_2' w^{N-2} & \cdots & \gamma_{N-1}' w^{(N/2-1)(N-1)} \end{pmatrix} \begin{pmatrix} t_0 \\ t_1 \\ t_2 \\ \vdots \\ t_{N-1} \end{pmatrix}. \tag{48}$$

This is not a square system because the vector on the left hand side is comple and the vector on the right is real. Thus each row really represents two real linear equations; the real and imaginary parts. Thus we can rewrite the system as a real $N \times N$ system.

We then solve this system for $\{t_k\}$ and compute the points $\{z_k + \gamma'(z_k)t_k\}$. These points are not on the curve Γ, but since Γ is smooth, the distance of these points to Γ is $O(t_k^2)$ and we project them onto the curve to find the updated $\{z_k\}$'s. One way to do this is update the value of $h(w_k)$ by adding t_k to argument of the current value. In Fornberg's paper [] he considers curves given by an equation of the form $\Gamma = \{(x, y) : f(x, y) = 0\}$ and he uses Newton's method starting at $z_k + \gamma'(z_k)t_k$ to find a point on Γ. Regardless of how the point on the tangent lines is mapped to a point on the curve, this completes the iteration. The whole procedure can then be repeated until it converges to a collection of points $\{z_k\}$ on Γ which are the images of the N roots of unity under a degree $\frac{N}{2} - 1$ polynomial. This polynomial if Fornberg's approximation to the power series for f, the conformal map from $\mathbb{D}$ to Ω.

The linear system described above may seem a bit unwieldy, but in fact it is quite highly stuctured and can be quickly applied to a vector using the Fast Fourier Transform. Note that (48) can be written

$$-d = \frac{1}{N} F D_0 s_0 + \frac{1}{N} W F D_1 s_1, \tag{49}$$

where

$$d = \begin{pmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{1-N/2} \end{pmatrix}, \qquad s_0 = \begin{pmatrix} t_0 \\ t_2 \\ t_4 \\ \vdots \\ t_{N-2} \end{pmatrix}, \qquad s_1 = \begin{pmatrix} t_1 \\ t_3 \\ t_5 \\ \vdots \\ t_{n-1} \end{pmatrix},$$

$$W = \begin{pmatrix} w & & & & \\ & w^2 & & & \\ & & w^3 & & \\ & & & \ddots & \\ & & & & w^{N/2-1} \end{pmatrix},$$

$$D_0 = \begin{pmatrix} \gamma_0' & & & & \\ & \gamma_2' & & & \\ & & \gamma_4' & & \\ & & & \ddots & \\ & & & & \gamma_{N-2}') \end{pmatrix},$$

$$D_1 = \begin{pmatrix} \gamma_1' & & & & \\ & \gamma_3' & & & \\ & & \gamma'(z_5) & & \\ & & & \ddots & \\ & & & & \gamma_{N-1}' \end{pmatrix}$$

and F is the discrete Fourier transform matrix corresponding to the $N/2$nd root of unity w^2,

$$F = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w^2 & w^4 & \cdots & w^{N-2} \\ 1 & w^4 & w^8 & \cdots & w^{2(N-2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & w^{N-2} & w^{2(N-2)} & \cdots & w^{2(N/2-1)^2} \end{pmatrix}$$

However, (49) is not quite a standard linear equation, since the left hand side is a complex vector and the desired solution is a pair of real vectors. If we solve this system for s_1 in terms of s_0 we get

$$s_1 = D_1^{-1} F^{-1} W^{-1} (F D_0 s_0 - d),$$

where the inverses of the diagonal matrices are easy to compute and the inverse of the Fourier matrix can be applied in $N \log N$ time. Moreover, if there is a real solution (s_0, s_1) of (49), then we must have

$$s_1 = \Re(D_1^{-1} F^{-1} W^{-1} (F D_0 s_0 - d)),$$

Similarly,

$$s_0 = \Re(D_0^{-1} F^{-1} (W F D_1 s_1 - d)).$$

Thus we must have

$$s_0 = \Re(D_0^{-1} F^{-1} (W F D_1 (\Re(D_1^{-1} F^{-1} W^{-1} (F D_0 s_0 - d))) - d)).$$

This looks complicated, but if we set

$$R = \Re(\frac{2}{N} D_0^{-1} F^{-1} W F D_1),$$

$$A = R R^T,$$

$$G = 1 - A,$$

$$b = \Re(D_0^{-1} F^{-1} (W F D_1 (\Re(D_1^{-1} F^{-1} W^{-1} (d))) + d)).$$

then it becomes

$$G s_0 = b.$$

Moreover $A = RR^T$ and hence G are symmetric, positive semi-definite, so that the conjugate gradient method can be applied.

In his paper, Fornberg reports that in all the examples he tested, the matrix G had one eigenvalue close to zero and all other eigenvalues clustered close to one. By throwing away one row and column of G is thus obtained a $N-1 \times N-1$ system with all eigenvalues clustered near 1. In a later paper Wegman , explained the Fornberg's empirical observations using his solution of the numerical mapping problem via the Riemann-Hilbert equation. We will discuss this later.

4. Wegman's method

The starting point of Wegman's method is the same as Fornberg's method; we have a parameterization $\gamma : \mathbb{T} \to \Gamma = \partial\Omega$ and we seek a homeomorphism $h : \mathbb{T} \to \mathbb{T}$ so that $\gamma \circ h$ has a holomorphic extension to the disk. In Fornberg's method we test this by computing the Laurent expansion of our current guess for $\gamma \circ h$ and then modifying h to make the negative coefficients zero.

In Wegman's method, if $\gamma \circ h_k$ does not have a holomorphic extension, then we let $e^{i\beta(t)}$ denote the tangent of Γ at $\zeta(t) = \gamma(h_k(t))$ and choose a real valued function $u(t)$ so that

$$\Psi(t) = \zeta(t) + u(t)e^{i\beta(t)}, \tag{50}$$

is the boundary value of a holomorphic function. If we can find such a u, then set

$$h_{k+1} = h_k + u. \tag{51}$$

The main point is then whether such a u exists and how to compute it. We can rewrite (50) as

$$e^{-i\beta(t)}\Psi(t) = e^{-i\beta(t)}\zeta(t) + u(t),$$

and since u is real valued,

$$\Im(e^{-i\beta(t)}\Psi(t)) = \Im(e^{-i\beta(t)}\zeta(t))$$

or

$$\Re(e^{-i\beta(t)}(-i)\Psi(t)) = \Im(e^{-i\beta(t)}\zeta(t))$$

or

$$\Re(e^{-i\beta(t)}\Phi(t)) = g(t), \tag{52}$$

where $\Phi(t) = -i\Psi(t)$ is in H^2 and $g(t) = \Im(e^{-i\beta(t)}\zeta(t))$ is real. Thus the desired u exists if, given g and β, there is an analytic Φ satisfying (52). This is an example of a Riemann-Hilbert problem. Moreover, if we can compute Ψ, then we can compute u by rewriting (50) as

$$u = e^{-i\beta}(\Psi - g). \tag{53}$$

Suppose Φ were a holomorphic solution of (52). Let $\alpha(t) = K[\beta(t) - t]$, so that $(\beta + i\alpha)/z$ has a holomorphic extension to the disk. Thus $(\beta + \imath\alpha)$ has a holomorphic extension which vanishes at the origin. Hence $e^{-i\beta+\alpha}$ is holomorphic and so $e^{-i\beta+\alpha}\Phi$ is holomorphic, so

$$\Im(e^{-i\beta+\alpha})\Phi) = K[\Re(e^{-i\beta+\alpha}\Phi)].$$

Since e^{α} is real valued, this is the same as

$$e^{\alpha}\Im(e^{-i\beta}\Phi) = K[e^{\alpha}\Re(e^{-i\beta}\Phi)] = K[e^{\alpha}g],$$

or

$$\Im(e^{-i\beta})\Phi) = e^{-\alpha}K[e^{\alpha}g].$$

Thus

$$e^{-i\beta}\Phi = \Re(e^{-i\beta}\Phi) + i\Im(e^{-i\beta}\Phi) = g + ie^{-\alpha}K[e^{\alpha}g],$$

or

$$\Phi = e^{i\beta}(g + ie^{-\alpha}K[e^{\alpha}g]) = e^{i\beta-\alpha}(e^{\alpha}g + iK[e^{\alpha}g]). \tag{54}$$

The right hand side is clearly a product of holomorphic functions, so this defines a holomorphic solution of (52). Together, equations (51), (53) and (54) give a formula for computing h_{k+1} from γ and h_k, and since, by smoothness, the curve Γ differs only quadratically from its tangent line, the sequence $\{\gamma \circ h_k\}$ converges quadratically to the desired conformal map (once we are close enough to the correct answer). If we use FFT's to compute the necessary conjugate functions, then clearly we only need $o(n \log n)$ to compute each iteration.

5. Comparing wegman's and Fornberg's methods

Wegman's method is to solve a Riemann-Hilbert problem directly using the conjugate function to define the solution. However, he also shows that same solution

occurs for a linear system

$$(I + R)u = g,$$

where

$$g = -\Re(e^{-i\beta}(I - iK + J)\zeta),$$

and

$$Ru = \Re(e^{-i\beta}(J - iK)[e^{i\beta}u],$$

where K is the conjugate operator and J maps a function to its mean value. The operator R can also be written as an integral operator with kernel

$$R(t, s) = -\frac{1}{2\pi}\frac{\sin(\beta(t) - \beta(s) - (t - s)/2)}{\sin((t - s)/2)}.$$

Wegman shows that $\|R\| \leq 1$, that -1 is a simple eigenvalue of R, and 1 is not an eigenvalues and that the remaining eigenvalues cluster rapidly around 0 if Γ is smooth.

Wegman shows that Fornberg's G matrix is a discretization of the operator $I - R^2$. By our previous remarks, $I - R^2$ has one zero eigenavalue and the rest cluster quickly around 1. Thus Fornberg's emperical observations about the eigenvalues of his G matrix are explained by Wegman's analysis of his own method.

CHAPTER 11

Higher dimensions

1. Liouville's theorem

2. Hamilton's theorem

3. Spectral geometry

CHAPTER 12

Higher conenctivity

1. The uniformization theorem

2. Koebe's theorem

3. Koebe's conjecture

4. Slit mappings

CHAPTER 13

Circle packings

No method of computing conformal mappings is as appealing to the eye or intuition as the circle packing method, introduced by Thurston and implemented and studied by Stephenson and his students. Here we take the idea that conformal maps preserve infimitesimal circles and consider maps that preserve actual circles of positive radius. Given a domain we pack it with disjoint circles and prove that there is an essentially unique packing of the unit disk by circles that have the same tangentcy relations as the given packing. This map between circle packings induces a quasiconformal map between the domains, which converges to a conformal map when the original packing consists of "small enough" circles.

We do not have the space here to recreate the whole theory of circle packings, but we shall sketch the proofs circle packing maps exist, converge to conformal maps and describe an algorithm for computing these maps.

1. Definitions

2. The Perron method

3. The hexagonal packing is rigid

4. Packing maps converge to conformal maps

CHAPTER 14

Conformal Welding

1. The fundamental theorem

2. Koebe's theorem and conformal welding

3. Marshall's Zipper algorithm

4. SLE

We saw above that Brownian motion has a close conenction to harmonic measure, conformal invariants and conformal maps. One of the most important and intriguing applications of conformal maps has been to the theory of random simple curves. This is currently rather technical topic to discuss in detail, but it is so important that we should at least draw a few pictures.

Suppose $\kappa > 0$ and let $a = \frac{1}{2}(1 - \sqrt{\kappa/(16+\kappa)})$. Let

$$f_1(z) = (z-a)^a (z-(a-1))^{1-a},$$

$$f_2(z) = (z+a)^a (z+(a-1))^{1-a}.$$

Each of these maps sents the upper half-plane into itself minus a slit at 0. In each case, two points are mapped to the the origin. For f_1 these points are $a, 1-a$ and for the second they are $-a$, and $a-1$. Now iterate these maps at random.

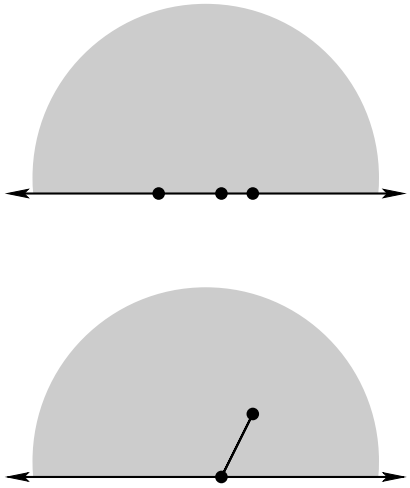


FIGURE 1. given three points on the real line there is an explicit map of the upper half-plane to a slit half-plane which sens the center point to the tip of the slit and maps both the outer points to the origin, as shown in this figure. By randomly selecting the points to be identified, and composing the corresponding maps, we create a random path in the upper half plane.

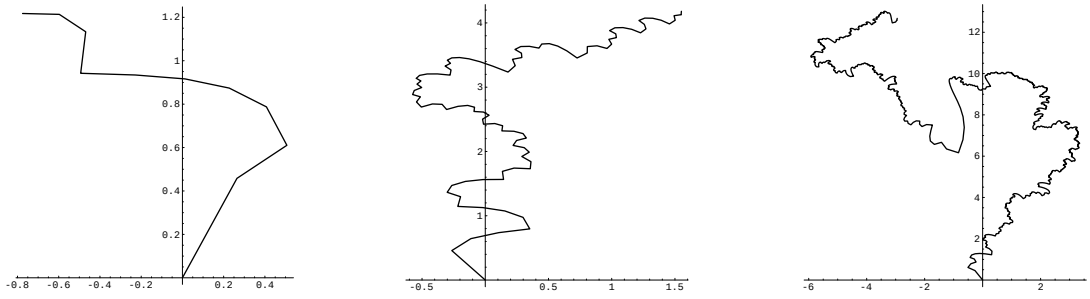


FIGURE 2. SLE paths for $\kappa = 2$, steps = 10,100, 1000.

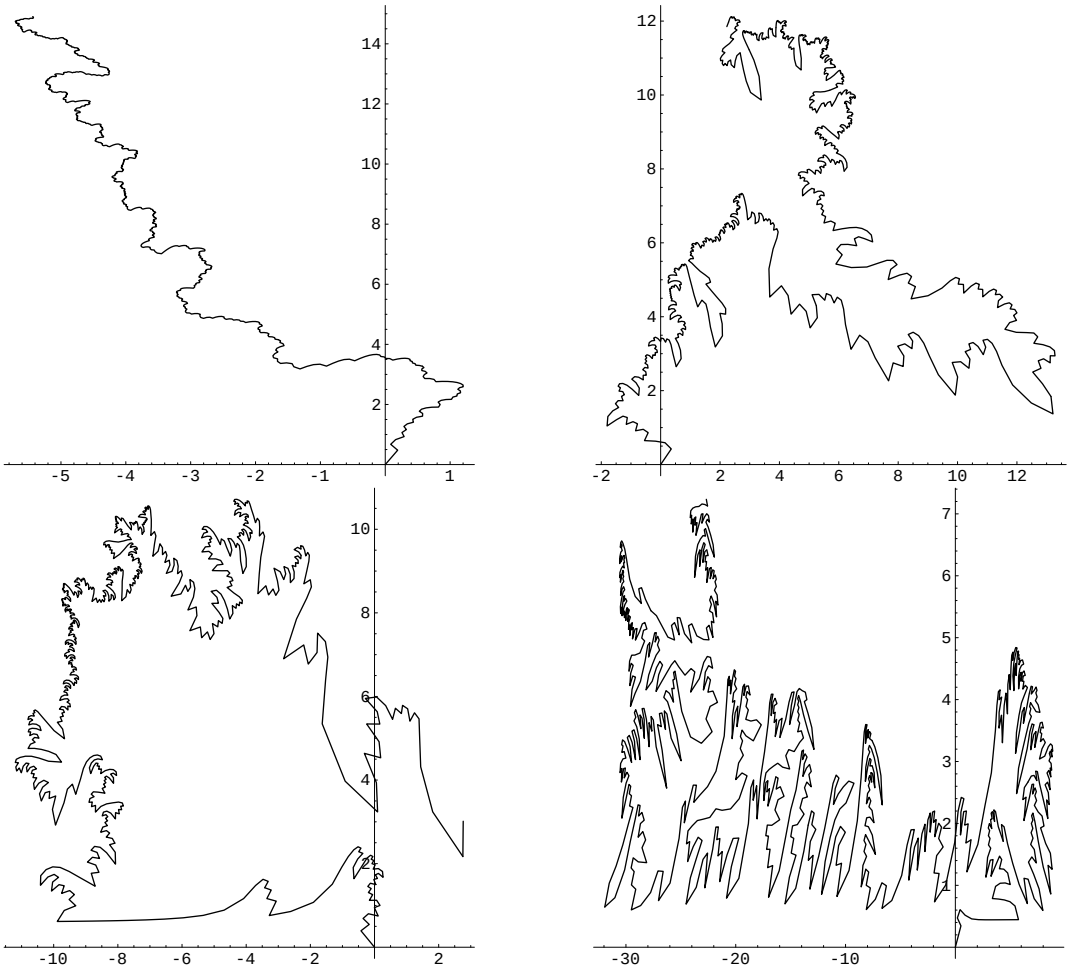


FIGURE 3. SLE paths for $\kappa = 2, 4, 6, 10$, steps $= 1000$.

CHAPTER 15

The Schwarz-Christoffel formula (again)

1. Circular-arc polygons
2. Multiple connected domains
3. Black box solvers

CHAPTER 16

Conformal mapping in linear time

1. The idea

2. Thick and thin parts of a polygon

3. Arches

4. Building approximate bending laminations

5. Angle scaling is continuous

6. The algorithm

CHAPTER 17

Conformal maps and martigales

1. The Bloch space and Nehari's theorem

2. Bloch functions and Bloch martingales

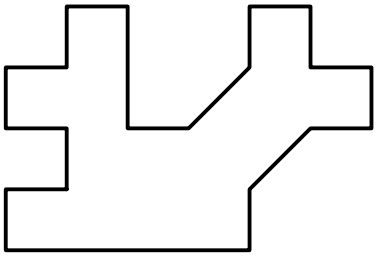
3. Radial limits of conformal maps

4. Makarov's upper bound

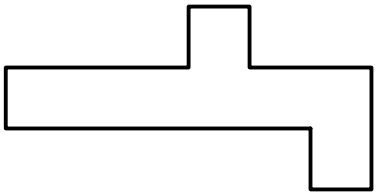
5. The law of the iterated logarithm

APPENDIX A

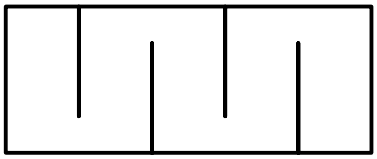
Some domains used in the text



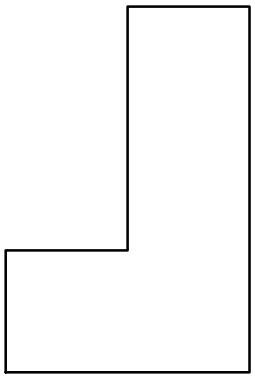
$\{2 - I, 2, 3 + I, 4 + I, 4 + 2I, 3 + 2I, 3 + 3I, 2 + 3I, 2 + 2I, 1 + I, I, 3I, -1 + 3I, -1 + 2I, -2 + 2I, -2 + I, -1 + I, -1, -2, -2 - I\}$
Figures **??**, **??**, 8, 4, **??**, 4



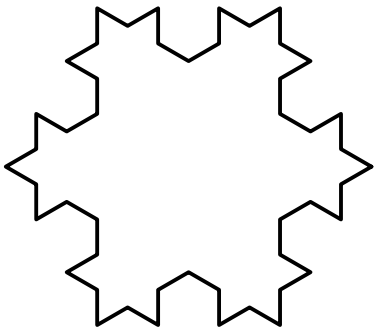
$\{10 + 2I, 10, 12, 12 + 4I, 8 + 4I, 8 + 6I, 6 + 6I, 6 + 4I, 4I, 2I\}$
Figures **??**, 22, 25, 27



$\{3-2I, 3+I, 3-2I, 5-2I, 5+2I, 1+2I, 1-I, 1 + 2I, -3 + 2I, -3 - I, -3 + 2I, -5 + 2I, -5 - 2I, -1 - 2I, -1 + I, -1 - 2I\}$
$\{\{7, 15, 16\}, \{1, 7, 16\}, \{1, 2, 7\}, \{2, 6, 7\}, \{2, 5, 6\}, \{2, 4, 5\},$
Figures 26 29 .

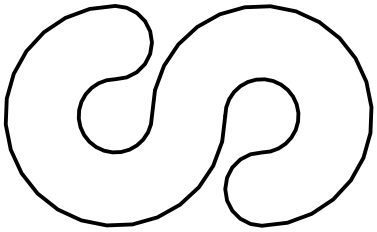


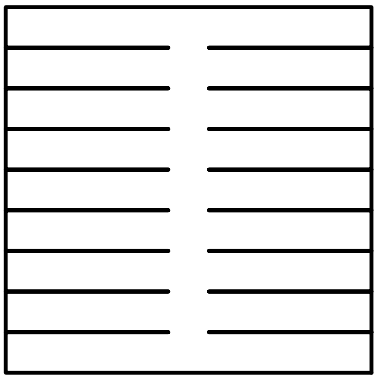
$\{0, 2, 2+3I, 1+3I, 1+I, I\}$
$\{\{1,2,5\},\{2,3,5\},\{3,4,5\},\{5,6,1\}\}$
Figures 33, 34, 25, 27

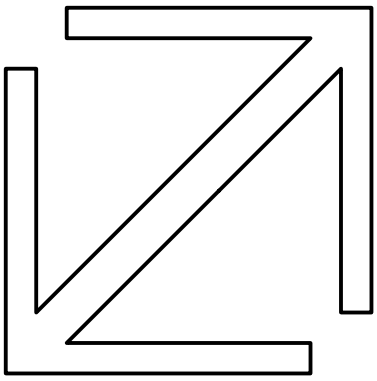


Program for compuing vertices is given in Appendix **??**.
triVK= 5, 21, 29, 5, 13, 21, 5, 29, 45, 29, 37, 45, 1, 2, 48, 2, 3, 4, 46, 47, 48, 46, 48, 2, 2, 4, 46, 4, 5, 46, 5, 45, 46, 9, 10, 8, 10, 11, 12, 6, 7, 8, 6, 8, 10, 10, 12, 6, 12, 13, 6, 13, 5, 6, 17, 18, 16, 18, 19, 20, 14, 15, 16, 14, 16, 18, 18, 20, 14, 20, 21, 14, 21, 13, 14, 25, 26, 24, 26, 27, 28, 22, 23, 24, 22, 24, 26, 26, 28, 22, 28, 29, 22, 29, 21, 22, 33, 34, 32, 34, 35, 36, 30, 31, 32, 30, 32, 34, 34, 36, 30, 36, 37, 30, 37, 29, 30, 41, 42, 40, 42, 43, 44, 38, 39, 40, 38, 40, 42, 42, 44, 38, 44, 45, 38, 45, 37, 38;
Figures **??**, **??**, **??**, 10, refVK2offcenter

This is a $1 \times \pi/2$ rectangle with half-disks attached to each shorter side. In most examples this is approximated by a 40-gon with 10 edges approximating each circular arc and straight edge.
Figures 26, **??**,







```
makevert34=(
   vert34={1}; n1=20; n2=20; n3=10;
   Do[ AppendTo[vert34,2+1 Exp[I  (Pi - (3/2)Pi
   Do[ AppendTo[vert34,2-2 I + Exp[ I ((Pi/2) +P
   Do[ AppendTo[vert34,2+3 Exp[I  (Pi - (3/2) Pi
   Do[ AppendTo[vert34,-2+ Exp[I  (3/2) Pi(-k/n1
   Do[ AppendTo[vert34,-2+2 I + Exp[ I ((-Pi/2) +
   Do[ AppendTo[vert34,-2+ 3 Exp[I  (3/2) Pi(-k/n
   vert34=N[vert34];
)
```

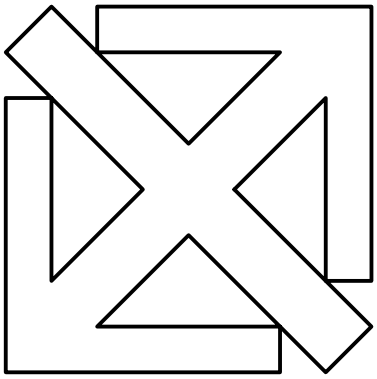
Figures 27, **??**,

$\{9-7I, 9-5I, 1-5I, 9-5I, 9-3I, 1-3I, 9-3I, 9-I, 1-I, 9-I, 9+I, 1+I, 9+I, 9+3I, 1+3I, 9+3I, 9+5I, 1+5I, 9+5I, 9+7I, 1+7I, 9+7I, 9+9I, -9+9I, -9+7I, -1+7I, -9+7I, -9+5I, -1+5I, -9+5I, -9+3I, -1+3I, -9+3I, -9+I, -1+I, -9+I, -9-I, -1-I, -9-I, -9-3I, -1-3I, -9-3I, -9-5I, -1-5I, -9-5I, -9-7I, -1-7I, -9-7I, -9-9I, 9-9I, 9-7I, 1-7I\}$

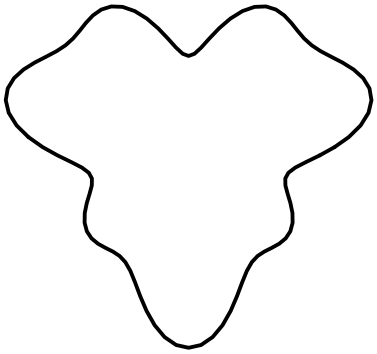
Figures 33, **??**,

{1, 5 + 4 I, 5 - 4 I, 6 - 4 I, 6 + 6 I, -4 + 6 I, -4 + 5 I, 4 + 5 I, I, -1, -5 - 4 I, -5 + 4 I, -6 + 4 I, -6 - 6 I, 4 - 6 I, 4 - 5 I, -4 - 5 I, -I}

Figures 5



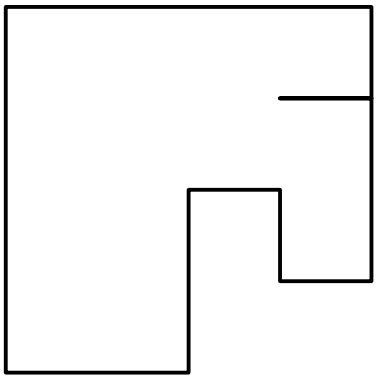
{1, 3 + 2 I, 3 - 2 I, 4 - 2 I, 4 + 4 I, -2 + 4 I, -2 + 3 I, 2 + 3 I, I, -3 + 4 I, -4 + 3 I, -1, -3 - 2 I, -3 + 2 I, -4 + 2 I, -4 - 4 I, 2 - 4 I, 2 - 3 I, -2 - 3 I, -I, 3 - 4 I, 4 - 3 I}
Figures 5



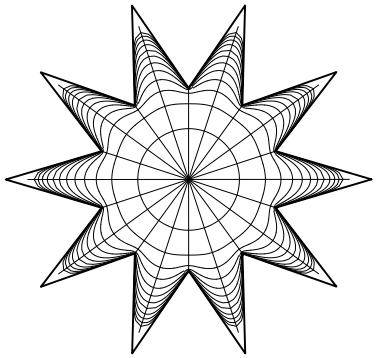
Asmooth starshaped region with radius given by

$$\theta \to \frac{1}{5}(\frac{1}{2}\sin(7\theta) + \sin(3\theta) + 4)e^{i\theta}).$$

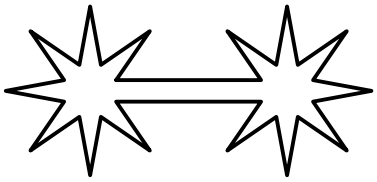
Figures 11



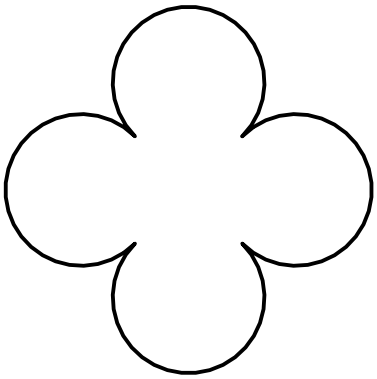
{2, 4, 4 + 2 I, -4 + 2 I, -4 - 6 I, -6 I, -2 I, 2 - 2 I, 2 - 4 I, 4 - 4 I, 4, 2}
Figures 23 , 13, 17



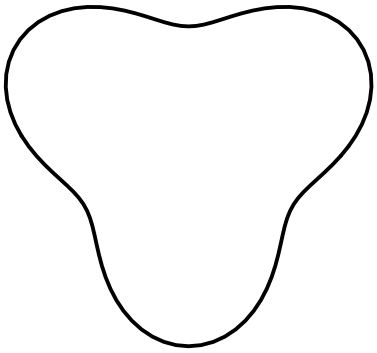
This has ten equally spaced vertices oneach of the cirles of radius 1 and 2 around the origin. Figures 35



{0.92388 + 0.382683 I, 2.12132 + 2.12132 I, 0.382683 + 0.92388 I, 0. + 3. I, -0.382683 + 0.92388 I, -2.12132 + 2.12132 I, -0.92388 + 0.382683 I, -3., -0.92388 - 0.382683 I, -2.12132 - 2.12132 I, -0.382683 - 0.92388 I, 0. - 3. I, 0.382683 - 0.92388 I, 2.12132 - 2.12132 I, 0.92388 - 0.382683 I, 6.07612 - 0.382683 I, 4.87868 - 2.12132 I, 6.61732 - 0.92388 I, 7. - 3. I, 7.38268 - 0.92388 I, 9.12132 - 2.12132 I, 7.92388 - 0.382683 I, 10., 7.92388 + 0.382683 I, 9.12132 + 2.12132 I, 7.38268 + 0.92388 I, 7. + 3. I, 6.61732 + 0.92388 I, 4.87868 + 2.12132 I, 6.07612 + 0.382683 I}
Figures 5



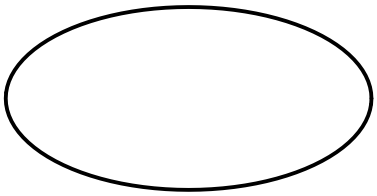
This is four circular arcs of radius $\sqrt{2}$ centered at the points ± 2, $\pm 2i$. each arc is discrtized by 25 points
Figures **??**



This is a smooth starshaped domain where the radius is given by

$$r(\theta) = \frac{1}{2}(\sin(3x) + 4).$$

Figures 1



A 2×1 ellipse. Figures 1

APPENDIX B

Some *Mathematica* code

This prints a polygon when the vertices are given as a list of pairs of real coordinates.

```
plotpoly[vertlist_]:=
         Show[Graphics[
            {
              GrayLevel[1],Thickness[.01], {Polygon[vertlist] } ,
              GrayLevel[0],Thickness[.01], {Line[vertlist] }
             },
            {PlotRange -> All, Axes -> None, AspectRatio -> Automatic}]];
```

This plots a polygon when the vertices are given as a list of complex numbers.

```
ctor[z_]:=(
   output={};
   Do[ AppendTo[output,{Re[z[[k]]],Im[z[[k]]]}] ,{k,1,Length[z]}];
   output
);


plotpolyC[vert_]:=(
  vert1=ctor[vert];
  vert2=AppendTo[vert1,vert1[[1]] ];
  plotpoly[vert2];
 );
```

This prints a list of ploygons given as lists of complex numbers

```
plotpolylistC[vert_]:=(
  vert3={};
  Do[
```

```
      vert1=ctor[vert[[k]]];
      vert2=AppendTo[vert1,vert1[[1]] ];
      AppendTo[vert3,vert2];
     ,{k,1,Length[vert]}];
    plotpolylist[vert3];
   );
```

This takes two triples of complex numbers and computes the Beltrami coefficient of the affine map which sends the three vertices of the first to the three vertices of the second. It is assumed that the orientation of the triangles is the same in both cases.

```
comparetri[data1_,data2_]:=(
   (* map the first two points of each triangle to 0,1 by conformal linear
      map and compute image of third point *)
     z=data1[[1]];
     w=data1[[2]];
     x=data1[[3]];
     a= (x-z)/(w-z);
     z=data2[[1]];
     w=data2[[2]];
     x=data2[[3]];
     b= (x-z)/(w-z);
     If[ Im[a] Im[b] < 0, output =-1];
     If[ Im[a] Im[b] >0,(
         mu=Abs[(b-a)/(Conjugate[a]-b)];
         output = (mu+1)/(1-mu);
         (* {a,b,mu,output}*)
     )];
     N[output]
);
```

Takes two (equal length) lists of triangles and computes the maximum value from `comparetri`.

```
findmaxcompare[data1_,data2_]:=(
    max= -1;
    comparelist={};
    Do[
     current=comparetri[data1[[k]],data2[[k]] ];
     AppendTo[comparelist,current];
     If[current > max, {max=current;maxk=k;}];
    ,{k,1,Length[data1]}];
   max
);
```

Given two lists of vertices and one list of triples of vertex indices, compute two lists of triangles in plane and compute maximum distortion

```
comparepolys[vert1_,vert2_,triangles_]:=(
   tri1=Table[{
               vert1[[triangles[[k,1]] ]],
               vert1[[triangles[[k,2]] ]],
               vert1[[triangles[[k,3]] ]]
              }, {k,1,Length[triangles]}];
   tri2=Table[{
               vert2[[triangles[[k,1]] ]],
               vert2[[triangles[[k,2]] ]],
               vert2[[triangles[[k,3]] ]]
              }, {k,1,Length[triangles]}];
  findmaxcompare[tri1,tri2]
);
```

`compute_iota` takes a tree of disks and computes the Mobius maps from each disk to its parent. These are stored as 2 by 2 matrices. The tree of disks data is in data1 which is a list of 4-tuples of the form z,r,p,q where z is a complex number giving the center of the disk, r is a non-negative real giving the radius of the disk, p is an integer giving the label of the parent disk and q is in 1,2,3,4 giving the type of edge (1 is

vertex-vertex bisector, 2 is an edge-edge bisector, 3 is a vertex-edge bisector with the vertex on the right and 4 is vertex-edge with vertex on the left.

After computing disk to parent maps, we compute the disk to root maps.

The we take the vertex list in data2 and each disk is mapped to the closest boundary point of the nearest disk (for concave vertices this a disk whose boundary contains v and for convex vertices it is the disk at the other end of the medial axis edge ending at v).

```
computeiota[data1_,data2_]:=(
  z=data1[[1,1]];
  r=data1[[1,2]];
  mat={{1,-z},{0,r}};
  iota={};
  tau={};
  test={};
  AppendTo[iota,mat];
  AppendTo[tau,{{1,0},{0,1}}];
  Do[(
     z1=data1[[k,1]];
     r1=data1[[k,2]];
     p =data1[[k,3]];
     t =data1[[k,4]];
     z2=data1[[p,1]];
     r2=data1[[p,2]];
     r=Abs[z1-z2];
     (* ------    case 1 ---------------*)
     If[t==1,(
        cosalpha=(r^2 -r1^2-r2^2)/(-2 r1 r2);
        sinalpha=Sqrt[1-cosalpha^2];
        eialpha=cosalpha + I sinalpha;
        Ralpha={{eialpha,0},{0,1}};
        costheta=(r1^2 -r^2-r2^2)/(-2 r r2);
        sintheta=Sqrt[1-costheta^2];
```

```
    eitheta=costheta + I sintheta;
    a=z2+r2 eitheta (z1-z2)/r;
    b=z2+r2 Conjugate[eitheta] (z1-z2)/r;
    sigma1={{1,-b},{1,-a}};
    mat= Inverse[sigma1] . Ralpha . sigma1;
)];
 (* ------    case 2,  r1=r2  ---------------*)
 If[(t==2)&&(r1==r2) ,(
    eta={{1,-z1},{0,(z2-z1) r1 /r}};
    A=r/r1;
    w=(1-Exp[-A])/(1+Exp[-A]);
    sigma1={{1,A},{0,1}} . {{1,-w},{-w,1}};
    mat=Inverse[eta] . sigma1 . eta;
 )];
(* ------    case 2, r1 < r2   ---------------*)
 If[(t==2)&&(r1<r2)&&(r1>0) ,(
    s=r1 r /(r2-r1);
    z3=z1+s(z1-z2)/r;
    eta={{1,-z3},{0,z1-z3}};
    B=r1/Abs[z1-z3];
    y=Abs[z2-z3]/Abs[z1-z3];
    A=Log[y]/B;
    w=(1-Exp[-A])/(1+Exp[-A]);
    sigma1= {{y B ,y},{0,1}} . {{1,-w},{-w,1}} . {{1,-1},{0,B}};
    mat=Inverse[eta] . sigma1 . eta;
 )];
 (* ------    case 2, r1 > r2   ---------------*)
 If[(t==2)&&(r1>r2) ,(
    s=r1 r /(r2-r1);
    z3=z1+s(z1-z2)/r;
    eta={{1,-z3},{0,z1-z3}};
    B=r1/Abs[z1-z3];
```

```
   y=Abs[z2-z3]/Abs[z1-z3];
   A=Log[y]/B;
   w=-(1-Exp[A])/(1+Exp[A]);
   sigma1= {{y B ,y},{0,1}} . {{1,-w},{-w,1}} . {{1,-1},{0,B}};
   mat=Inverse[eta] . sigma1 . eta;
)];
 (* ------------------- case 2 vertex ---------------------*)
If[(t==2)&&(r1==0),(
  mat={{1,(z2-z1)(r-r2)/r},{0,1}};
)];
(* ---------------------- case 4  vertex on right--------------*)
If[t==4,(
   costheta=(r1^2 -r^2-r2^2)/(-2 r r2);
   sintheta=Sqrt[1-costheta^2];
   eitheta=costheta + I sintheta;
   a=z2+r2 eitheta (z1-z2)/r;
   s=r1-r2;
   L=Sqrt[r^2-s^2];
   costheta=s/r;
   sintheta=Sqrt[1-costheta^2];
   eitheta=costheta + I sintheta;
   c= z1+ r1 eitheta (z2-z1)/r ;
   d= z2+ r2 eitheta (z2-z1)/r ;
   astar=reflect[a,c,d];
   eta= {{-I(c-a)/(c-astar) ,0},{0,1}} . {{1,-astar},{1,-a}};
   cosalpha=(r^2 -r1^2-r2^2)/(-2 r1 r2);
   sinalpha=Sqrt[1-cosalpha^2];
   eialpha=cosalpha + I sinalpha;
   alpha=ArcCos[cosalpha];
   sigma1={{eialpha, -eialpha alpha},{0,1}};
   mat= Inverse[eta] . sigma1 . eta;
)];
```

```
   (*---------------- case 3  vertex on left ---------------------*)
   If[t==3,(
      costheta=(r1^2 -r^2-r2^2)/(-2 r r2);
      sintheta=Sqrt[1-costheta^2];
      eitheta=costheta - I sintheta;
      a=z2+r2 eitheta (z1-z2)/r;
      s=r1-r2;
      L=Sqrt[r^2-s^2];
      cospsi=s/r;
      sinpsi=Sqrt[1-cospsi^2];
      eipsi=cospsi - I sinpsi;
      c= z1+ r1 eipsi (z2-z1)/r ;
      d= z2+ r2 eipsi (z2-z1)/r ;
      astar=reflect[a,c,d];
      eta= {{-I(c-a)/(c-astar) ,0},{0,1}} . {{1,-astar},{1,-a}};
      cosalpha=(r^2 -r1^2-r2^2)/(-2 r1 r2);
      sinalpha=Sqrt[1-cosalpha^2];
      eialpha=cosalpha + I sinalpha;
      alpha=ArcCos[cosalpha];
      sigma1={{Conjugate[eialpha], Conjugate[eialpha] alpha},{0,1}};
      mat= Inverse[eta] . sigma1 . eta;
   )];
  (* test that tau map disk to parent*)
   AppendTo[test,{k,t,alpha}];
   AppendTo[test,Abs[z2-eval[tau[[k]],z1+r1]]-r2];
   AppendTo[test,Abs[z2-eval[tau[[k]],z1+I r1]]-r2];
   AppendTo[test,Abs[z2-eval[tau[[k]],z1-I r1]]-r2];
),{k,2,Length[data1]}]   ;
  (* we have now created the iota map. Now we apply
     it to the vertices. First associated each vertex
     to a medial axis vertex disk and then apply the
     iota map for that disk to the vertex *)
```

```
  data3=vert2data2[data1,data2];
  out={};
   Do[(
     p=data3[[k]];
     v=data2[[k]];
     AppendTo[out,eval[iota[[p]],v]];
     ),{k,1,Length[data2]}];
    out
);
```

`vert2data` this takes in a tree-of-disks array and a list of vertices as complex numbers and and returns an equal length list with these points listed as disk number,angle. This output can then be used in the `drawtraj` routines .

```
vert2data[data1_,data2_]:=(
  output={};
  L1=Length[data1];
  L2=Length[data2];
  Do[
     v=data2[[k]];
     mindist=10000;
     nearest=-1;
     Do[
        w=data1[[L1-j+1,1]];
        r=data1[[L1-j+1,2]];
        dist=Abs[v-w];
        flag=verifyangle[data2,k,w];
        If[(flag==1)&&(dist-r<=mindist+.0000001),(mindist=dist-r;nearest=L1+1-j;)];
        (* If[(flag==1)&&((dist-r)==0),(mindist=dist;nearest=L1+1-j;)];  *)
     ,{j,1,L1}];
     w=data1[[nearest,1]];
     r=data1[[nearest,2]];
     If[mindist+r>0,AppendTo[output,{nearest,Arg[v-w],mindist}]];
     If[mindist+r==0,AppendTo[output,{nearest,0}]];
```

```
  ,{k,1,L2}];
  N[output]
);
```

vert2data2 also checks to see if the any of the vertices are zero radius medial axis disks and if so assigns them to their parent

```
vert2data2[data1_,data2_]:=(
  output={};
  L1=Length[data1];
  L2=Length[data2];
  Do[
     v=data2[[k]];
     (* if v is a medial axis disk find correct   index in data1
        and set out to this index *)
     flag1=0;
     Do[If[v==data1[[j,1]], (
           zparent=data1[[data1[[j,3]],1]];
           flag2=verifyangle[data2,k,zparent];
           If[flag2==1,{AppendTo[output,j];flag1=1;}]
        )];
     ,{j,1,L1}];
    (* if v is not a medial axis disk, then flag1 =0. In this
       case, seach all medial axis disk and assign v to the
       one  for which v is closest to the boundary *)
     If[flag1==0,(
        mindist=10000;
        nearest=-1;
        Do[
           w=data1[[L1-j+1,1]];
           r=data1[[L1-j+1,2]];
           If[Abs[v-w]==r,nearest=L1-j+1];
        ,{j,1,L1}];
```

```
        AppendTo[output,nearest];
      )];
  ,{k,1,L2}];
  output
);
```

Perform one iteration of the Koebe method.

```
koebe[data_]:=(
   L = Length[data];
   minabs=10000;
   Do[
       If[Abs[data[[k]] ]<minabs,(minabs=Abs[data[[k]] ];loc=k;)];
   ,{k,1,L}];
   data1=shift[data,loc];
   min=data1[[1]];
   data2=N[diskmap[data1,min]];
   zeroimage=diskmap[0,min];
   data3=findsquareroot[data2];
   zeroimage=findsquareroot2[data3,zeroimage];
   data4=diskmap[data3,zeroimage];
   shift[data4,2-loc]
);
```

Do n iterations of Koebe iteration and return the list of vertices.

```
iterkoebe[data_,n_]:=(
  data6=data;
  Do[ data6=koebe[data6]; ,{k,1,n}];
  data6
);
```

Perform n iterations of Koebe iteration and store intermediate results. Plot the resukts as an arrary.

```
ik[data_,n_]:=(
 data8=Table[0,{k,1,n+1}];
 data8[[1]]=data;
```

```
 data9=data;
 closestpoint={};
 Do[
    data9=koebe[data9];
    cols=4; skip=5;
    If[Mod[k,skip]==0,(
       j=(k)/skip;
       translate= 2.5 (Mod[(j),cols] + -I Floor[(j)/cols]);
       data8[[j+1]]=data9+ translate  ;
       AppendTo[closestpoint,Min[Abs[data9]]];
    )];
   ,{k,1,n}];
 plotpolylistC[data8]
) ;
```

Apply a circular shift to a list.

```
shift[data_,n_]:=(
    L=Length[data];
    data2={};
    Do[
     AppendTo[data2,data[[1+Mod[k-2+n,L] ]]];
    ,{k,1,L}];
    data2
)
```

Find square root of a list of complexes which starts with 0 and is continuous branch

```
findsquareroot[data_]:=(
  data6={0};
  AppendTo[data6,Sqrt[data[[2]] ] ];
  Do[
    z=data6[[k-1]];
    w=Sqrt[data[[k]]];
    If[Abs[Arg[w]-Arg[z]]< Pi/2, AppendTo[data6,w], AppendTo[data6,-w]];
```

```
  ,{k,3,Length[data]}];
  data6
)
findsquareroot2[data_, point_]:=(
 argsum=0;
 point2=Sqrt[point];
 Do[
   argsum=argsum+Arg[(data[[k]]-point2)/(data[[k+1]]-point2)];
   ,{k,1,Length[data]-1}];
 If[Abs[argsum]<Pi, - point2, point2]
)
```

Find vertices of nth generation of von Koch snowflake.

```
makeKSvert[n_Integer?NonNegative]:=(
   out=      FixedPoint[
             (#1/.Line[{start_,finish_}]
              :>doline[start,finish])&,
              {Line[{{Sqrt[3],0},{-Sqrt[3]/2,-3/2}}],
              Line[{{-Sqrt[3]/2, -3/2},{-Sqrt[3]/2, 3/2}}],
              Line[{{-Sqrt[3]/2, 3/2},{Sqrt[3],0}}]},
              (*{Line[{{0,0},{1/2,Sqrt[3]/2}}],
              Line[{{1/2, Sqrt[3]/2},{1,0}}],
              Line[{{1,0},{0,0}}]},*)
           n];
   out=N[out];
   out=Flatten[out];
    vertVK={};
   Do[ AppendTo[vertVK, out[[k,1,1,1]]+I out[[k,1,1,2]]]
        ,{k,1,Length[out]}];
    vertVK = Conjugate[vertVK]
  );
```

APPENDIX C

Bits and pieces

In this appendix we present some related results that may have been used or mention in the text, but not fully discussed there.

1. Alternative definitions of quasiconformality

The metric definition: An orientation preserving homeomorphism $\Omega_1 \to \Omega_2$ is K-quasiconformal if for all $x \in \Omega_1$,

$$\limsup_{r\to 0} \frac{\max_{y:|x-y|=r} |f(x)-f(y)|}{\min_{y:|x-y|=r} |f(x)-f(y)|} \le K.$$

We shall see later that this is equivalent to only requiring the the liminf to be less than K for all x.

Geometric definition: An orientation preserving homeomorphism $\Omega_1 \to \Omega_2$ is K-quasiconformal if for every generalized quadrilateral Q in Ω_1 , $M(f(Q)) \le KM(Q)$.

Our first step is to show these definitions are equivalent. To prove "Metric $\Rightarrow$" Geometric" we need the following well known result.

THEOREM 125 (Besicovitch covering lemma). *Let E be a bounded set in $\mathbb{R}^n$ and a covering by balls centered at points of E and such that each point of E is the center of some disk. Then there is constant $C = C(n)$ and a subcover $\{B_j\}$ of E so that each point of E is an at most C elements of the subcover.*

PROOF. We may assume the balls have bounded diameters, for otherwise simply choose one large ball containing E. We essentially use a greedy algorithm, taking one of largest balls not yet used. Let α_1 be the supremum of the diameters of the balls and choose B_1 so that $|B_1| \ge \frac{3}{4}\alpha_1$. Let $E_1 = E \setminus B_1$ and let α_2 be the supremum of the diameters of balls centered at points of E_1. Choose B_2 so that $|B_2| \ge \frac{3}{4}\alpha_2$. Continue in this way.

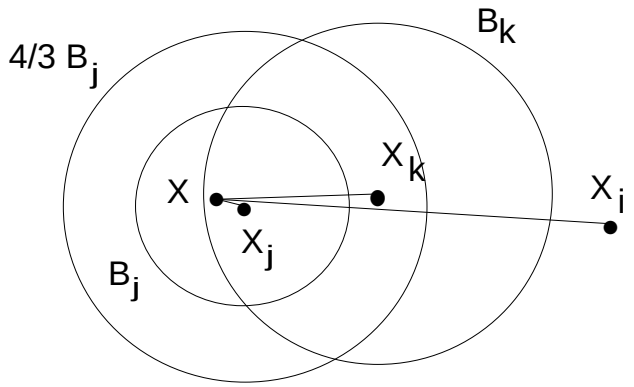


FIGURE 1. Proof that there cannot be three balls containing x, all with centers in same direction.

To check that the resulting collection has bounded overlap, take $x \in E$ and suppose x lies in B_j and B_k with $k > j$. Then $|B_k| \leq \frac{4}{3}|B_j|$ since it was chosen later. Let x_j, x_k be the centers of B_j, B_k. Since $x_k \notin B_j$ and

$$|x - x_k| \leq |B_k| \leq \frac{4}{3}|B_j| \leq \frac{4}{3}|x - x_j|.$$

If we are given three points x_i, x_j, x_k so that the three segments $[x, x_i], [x, x_j]$ and $[x, x_k]$ form three small angles at x, then since x_i must be in $\frac{4}{3}B_j \setminus B_j$ we see that it must also lie in B_k, a contradiction. Thus if we connect x to the center of each chosen ball containing it, no more than two segments can be within angle ϵ of each other. The compactness of the unit sphere now proves there can only be a bounded number of such balls containing x.

To see that $\{B_j\}$ is a covering suppose $x \in E$ is to covered. Then there is a ball B_x centered at x which was never chosen. This implies $\alpha_n \geq \frac{4}{3}|B_x|$ for all n. But then we have chosen an infinite number of bounded overlap balls, all with diameters bounded away from zero and lying in a bounded set of $\mathbb{R}^n$. This is contradiction, so E must be covered. □

2. The Hardy-Littlewood maximal theorem

We say that a collection of balls $\{B_j\}$ covers a set E in the *sense of Vitali* if each point of E is contained in balls of the collection of arbitrarily small diameter.

Now suppose $f : \Omega_1 \to \Omega_2$ satisifies the metric definition of K quasi-conformal. Suppose Q is a generalized quadrialteral. Suppose it can be conformal mapped to

a $1 \times r_1$ rectangle R_1 and its image, $f(Q)$ can be mapped to a $1 \times r_2$ rectangle R_2. Then f induces a map F between these rectangles which also satisfies the metric defintion of K-quasiconformality (with the same K). Choose some $H > K$ and for every $x \in R_1$ choose a ball, $B(x, r)$ centered at x so that

$$\max_{y:|x-y|=r} |f(x) - f(y)| \leq K \min_{y:|x-y|=r} |f(x) - f(y)|.$$

The conseqeunce that we shall use is that

$$\pi \text{diam}(F(B_k)) \leq H^2 \text{area}(f(B_k)). \tag{55}$$

By the Besicovitch lemma we can extra a covering $\{B_k\}$ of the interior of R_1 so that no point is in more than a bounded number of balls. Define a metric ρ on R_1 by taking

$$\rho(z) = \frac{1}{r_2} \sum_k \frac{\text{diam}(F(B_k))}{\text{diam}(B_k)} \chi_{2B_k}(z).$$

Then any curve $\gamma \in \Gamma$, the family of paths which connect the two sides of length one, and hits a ball B_k spends at least length $\text{diam}(B_k)$ inside $2B_k$. Thus

$$\int_\gamma \rho ds \geq \frac{1}{r_2} \sum_{k:\gamma \cap B_k \neq \emptyset} \text{diam}(B_k) \frac{\text{diam}(F(B_k))}{\text{diam}(B_k)} \geq \frac{1}{r_2} \sum_{k:\gamma \cap B_k \neq \emptyset} \text{diam}(F(B_k)),$$

which is bigger than 1 since the image $f(\gamma)$ connects the two sides of length 1 in R_2 and so the sets $f(B_k)$ that it hits must have diameters that sum to at least r_2. Thus ρ is admissible for Γ. Then

$$\begin{aligned}
\int_{R_1} \rho^2 dxdy &\leq \sum_k \frac{1}{r_2^2} (\frac{\text{diam}(F(B_k))}{\text{diam}(B_k)})^2 \text{area}(2B_k) \\
&\leq \sum_k \frac{1}{r_2^2} (\frac{\text{diam}(F(B_k))}{\text{diam}(B_k)})^2 \pi \text{diam}(B_k)^2 \\
&\leq \sum_k \frac{\pi}{r_2^2} (\text{diam}(F(B_k))^2 \\
&\leq \sum_k \frac{1}{r_2^2} H^2 \text{area}(f(B_k)) \\
&\leq \sum_k \frac{1}{r_2^2} C H^2 \text{area}(R_2) \\
&\leq \sum_k \frac{C H^2}{r_2}
\end{aligned}$$

Thus $r_2 \leq CH^2 r_1$, which means the f is CH^2-quasiconformal for the geometric definiion.

THEOREM 126 (Vitali covering lemma). *Suppose $E \subset \mathbb{R}^n$ is bounded and covered in the sense of Vitali by a collection of balls. Then there is a disjoint subcollection $\{B_j\}$ so that $E \setminus \cup_j B_j$ has n-dimensional measure zero,*

PROOF. The proof similar to the one above. Choose an open set $E \subset U$ with $\mathcal{H}^n(U \setminus E) < \epsilon$ and only consider the covering of E by balls inside U. Let α_1 be the supremum of the diameters of these balls and choose B_1 so that $|B_1 \geq \alpha_1/2$. Let α_2 be the supremum of the diameters of the balls in U which are disjoint from B_1 and choose B_2 from among these so that $|B_2| \geq \alpha/2$. Continue in this way.

The resulting collection is disjoint by choice and has the property that $\{5B_j\}$ covers E. To see this suppose $x \in E \setminus \cup_j 5B_j$. There is a ball $B_x \subset U$ which was never chosen. If it was not chosen because it intersected some previously chosen B_j, then $|B_x| \leq 2|B_j|$ so $B_x \subset 5B_j$, a contradiction. Therefore it was not chosen because $|B_j| \leq 2\alpha_n$ for every n. But this implies we have chosen infinitely many disjoint balls in U with diameters bounded away from zero, another contradiction. Thus $\{5B_j\}$ cover E.

Thus

$$\sum_j \mathcal{H}^n(B_j) = 5^{-n} \sum_j \mathcal{H}^n(5B_j) \geq \mathcal{H}^n(E).$$

Choose a finite subcollection $\{B_1, \ldots, B_{N_1}\}$ which covers $\frac{1}{2}5^{-n}$ of the measure of E. Let $E_1 = E \setminus \cup_j B_j$. Then if $\epsilon = \frac{1}{2}5^{-n}$, we get

$$\mathcal{H}^n(E_1) \leq \mathcal{H}^n(U \setminus \cup_j B_j) \leq \mathcal{H}^n(E) + \epsilon - 5^{-n}\mathcal{H}^n(E) \leq (1 - \frac{1}{2}5^{-n})\mathcal{H}^n(E).$$

Now repeat the argument to E_1 (first verify that the balls which are disjoint from the chosen ones cover E_1 in the sense of Vitali), obtaining a second set of disjoint balls $\{B_j^1\}$ so that if $E_2 = E_1 \setminus \cup_j B_j^1$, then

$$\mathcal{H}^n(E_2) \leq (1 - 5^{-n}/2)\mathcal{H}^n(E_1) \leq (1 - 5^{-n}/2)^2\mathcal{H}^n(E).$$

Continuing in the obvious way gives a disjoint collection of balls which covers almost every point of E. □

3. The distortion theorems

In this section we give the "usual" proof of Koebe's $\frac{1}{4}$ theorem, via the area theorem and deduce the sharp version of the distortion estimates. A weaker version of this was given in Lemma 5, and was all that was needed in the text.

Recall Green's theorem,

$$\iint_\Omega u\Delta v + v\Delta u dxdy = \int_{\partial\Omega} u\frac{\partial v}{\partial n} + v\frac{\partial u}{\partial n}ds, \tag{56}$$

where n denotes the inward pointing normal vector of $\partial\Omega$.

$$COMPLEXVERSION$$

We will also use Green's theorem in the following form:

$$\int_{\partial\Omega} f(x,y)dx + g(x,y)dy = \iint_\Omega \frac{\partial g}{\partial x} - \frac{\partial f}{\partial g}dxdy \tag{57}$$

and its simple consequence that the area of a region Ω is given by

$$\text{area}(\Omega) = \frac{1}{2}\int_{\partial\Omega} xdy - ydx = \frac{1}{2i}\int \partial\Omega\bar{z}dz. \tag{58}$$

We now come to some well known (but perhaps not as well known as the results above) estimates for univalent mappings. The basic idea is to show that a univalent map f on $\mathbb{D}$ is well approximated by its linear Taylor approximation $f(z_0)+f'(z_0)(z-z_0)$ in a hyperbolic neighborhood of z_0, with estimates that do not depend on f or z. These so called "distortion estimates" are fundamental to most arguments in geometric function theory. The first step is to prove:

THEOREM 127 (Area theorem). *Suppose $g(z) = \frac{1}{z} + b_0 + b_1 z + \dots$ is univalent in $\mathbb{D}$. Then $\sum_{n=0}^{\infty} n|b_n|^2 \le 1$. In particular, $|b_1| \le 1$.*

PROOF. For $0 < r < 1$ let $D_r = \mathbb{C} \setminus g(D(0,r))$. If $z = g(w)$ and $w = e^{i\theta}$ then $dw = iwd\theta$, so by (58),

$$\text{area}(D_r) = \iint_{D_r} dxdy = \frac{1}{2i}\int_{\partial D_r} \bar{z}dz = \frac{-1}{2i}\int_{\partial D(0,r)} \bar{g}(w)g'(w)dw.$$

To evaluate the right hand side note that

$$\begin{aligned} g(z) &= \frac{1}{z} + b_0 + b_1 z + \dots, \\ g'(z) &= 1\frac{1}{z^2} + 0 + b_1 + 2b_2 z + \dots, \end{aligned}$$

so that

$$\begin{aligned}\int_{|w|=r} \bar{g}(w)g'(w)dw &= i\int \bar{g}(w)g'(w)wd\theta \\ &= i\int (\frac{1}{\bar{w}} + \bar{b}_0 + \bar{b}_1\bar{w} + \dots)(-\frac{1}{w} + b_1 w + 2b_2 w + \dots)d\theta \\ &= 2\pi i(-\frac{1}{r^2} + |b_1|^2 r^2 + 2|b_2|r^4 + \dots 0\end{aligned}$$

Thus,

$$0 \le \text{area}(D_r) = \pi(\frac{1}{r^2} - \sum_{n=1}^{\infty} n|b_n|^2 r^{2n}).$$

Taking $r \to 1$ gives the result. □

COROLLARY 128. *If $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ is univalent on the unit disk, then $|a_2| \le 2$.*

PROOF. Let $g(z) = (f(z^2))^{-1/2} = 1/z - a_2 z/2 + \dots$. We claim g is one-to-one. To see this suppose $g(z) = g(w)$. Then $f(z^2) = f(w^2)$, so $z = \pm w$. Note that g is odd, so $z = w$. Since $b_1 = a_2/2$, the previous result implies $|a_2| \le 2$. □

THEOREM 129 (Koebe 1/4 theorem). *If f is univalent on $\mathbb{D}$, then*

$$\frac{1}{4}|f'(z)|(1-|z|^2) \le \text{dist}(f(z), \partial\Omega) \le |f'(z)|(1-|z|^2).$$

PROOF. By precomposing with a Möbius transformation and postcomposing by a linear map, we may assume $z = 0$, $f(0) = 0$ and $f'(0) = 1$. Then the right hand inequality is just Schwarz's lemma applied to f^{-1}. To prove the left hand inequality, suppose f never equals w in $\mathbb{D}$. Then

$$g(z) = \frac{wf(z)}{w - f(z)} = z + (a_2 + \frac{1}{w})z^2 + \dots,$$

is univalent with $f(0) = 0$ and $f'(0) = 1$. Applying Corollary 128 to both f and g gives

$$\frac{1}{|w|} \le |a_2| + |a_2 + \frac{1}{w}| \le 2 + 2 = 4.$$

Thus the omitted point w lies outside $D(0, 1/4)$, as desired. □

Because of Koebe's theorem we have

$$d\rho_\Omega \le d\tilde{\rho}_\Omega \le 4d\rho_\Omega. \tag{59}$$

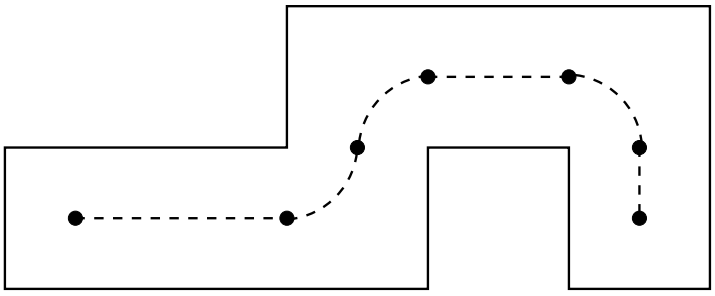


FIGURE 2. Its easy to compute the quasihyperbolic length of this curve (which is $3 + \frac{3}{2}\pi$), and a little more involved to show it is a quasi-hyperbolic geodesic, but together these facts give an estimate of its hyperbolic length.

LEMMA 130. *Suppose f is univalent on $\mathbb{D}$, $f(0) = 0$ and $f'(0) = 1$. Then*

$$\frac{1-|z|}{(1+|z|)^3} \leq |f'(z)| \leq \frac{1+|z|}{(1-|z|)^3},$$

PROOF. Fix a point $w \in \mathbb{D}$ and write the Koebe transform of f,

$$F(z) = \frac{f(\tau(z)) - f(w)}{(1-|w|^2)f'(w)},$$

where

$$\tau(z) = \frac{z+w}{1-\bar{w}z}.$$

This is univalent, so by Corollary 128, $|a_2(w)| \leq 2$. Differentiation and setting $z = 0$ shows

$$F'(z) = \frac{f'(\tau(z))\tau'(z)}{(1-|w|^2)f'(w)},$$

$$F''(z) = \frac{f''(\tau(z))\tau'(z)^2 + f'(\tau(z))\tau''(z)}{(1-|w|^2)f'(w)},$$

$$\tau'(0) = 1-|w|^2, \tau''(0) = -2(1-|w|^2),$$

$$F''(0) = \frac{f''(w)}{f(w)}(1-|w|^2) - 2\bar{w}.$$

This implies that the coefficient of z^2 (as a function of w) in the power series of F is

$$a_2(w) = \frac{1}{2}((1-|w|^2)\frac{f''(w)}{f'(w)} - 2\bar{w}).$$

Using $|a_2| \leq 2$ and multiplying by $w/(1-|w|^2)$, we get

$$\left|\frac{wf''(w)}{f'(w)} - \frac{2|w|^2}{1-|w|^2}\right| \leq \frac{4|w|}{1-|w|^2}.$$

Thus

$$\frac{2|w|^2-4|w|}{1-|w|^2} \leq \frac{wf''(w)}{f'(w)} \leq \frac{4|w|+2|w|^2}{1-|w|^2}.$$

Now divide by $|w|$ and use partial fractions,

$$\frac{-1}{1-|w|}+\frac{-3}{1+|w|} \leq \frac{1}{|w|}\frac{wf''(w)}{f'(w)} \leq \frac{3}{1-|w|}+\frac{1}{1+|w|}$$

Note that

$$\begin{aligned}\frac{\partial}{\partial r}\log|f'(re^{i\theta})| &= \frac{\partial}{\partial r}\mathrm{Re}\log f'(z)\\ &= \mathrm{Re}\frac{z}{|z|}\frac{\partial}{\partial z}\log f'(z)\\ &= \frac{1}{|z|}\mathrm{Re}(\frac{zf''(z)}{f'(z)})\end{aligned}$$

Since $w=re^{i\theta}$ and $f'(0)=1$, we can integrate to get

$$\log(1-r)-3\log(1+r) \leq \log|f'(re^{i\theta})| \leq -3\log(1-r)+\log(1+r).$$

Exponentiating gives the result. □

4. Extremal problems in geometric function theory

The area theorem can be seen as computing

$$\sup_f |f''(0)|,$$

where the supremum is over all univalent maps of the disk with $f(0)=0$ and $|f'(0)|=1$. Geomertric function theory has a long history of dealing with problems of this type, where we seek to find the infimum or supremum of some quantity over the class of univalent functions. Here we describe a few such problems (although we shall not describe the answers in much detail).

DeBrange's Theorem: One of the most famous problems in geometric function theory was the Bieberbach conjecture, which was proven by de Brange in 198?. If $f(z)=z+a_2z^2+a_3z^3+\ldots$, then $|a_n| \leq n$. This is best possible becuase of the Koebe map from $\mathbb{D}$ onto $\mathbb{C}\setminus[\frac{1}{4},\infty)$.

The Bloch-Landau constant: Suppose Ω is simply connected and $f:\mathbb{D}\to\Omega$ is conformal. The Koebe $\frac{1}{4}$ theorem implies that any simply connected domain contains

1929	Landau	¿ .566
1956	Riech	¿ .569
1961	Jenkins	¿ .5705
1968	Toppila	¿ .5708
1989	Zhang	¿ .57088
1999	Xiang	¿ .570584
1935	Robinson	¡ .658
1945	Goodman	.65647
1985	Bellar and Hummel	¡ .6564155
2008	Carroll and Ortega-Cerdá	¡ .6563937

1973	Metzger	3
197	Brennan	some $p > 3$
1985	Pommerenke	3.399
1999	Bertilsson	3.422

an disk of radius bigger than $\frac{1}{4}|f'(0)|$. What is the largest constant C so that Ω must contain a disk of radius $C|f'(0)|$? This number is called the schlicht Bloch-Landau constant. The exact value is still unknown, but here is table that summarizes some of the known results.

Halls' lemma: Suppose $E \subset \mathbb{D}$ is a closed set whose radial projection onto the unit circle covers the whole circle. What is the largest that $\omega(0, \mathbb{T}, \mathbb{D} \setminus E)$ can be? This was solved by Marshall and Sundberg,..

The omitted area problem:

Brennan's conjecture: Suppose Ω is simply connected and $g : \Omega \to \mathbb{D}$ is conformal. Then

$$\int_\Omega |g'|^2 dxdt = \pi,$$

so clearly $|g'| \in L^2(\Omega, dxdy)$. Brennan's conjecture asks if $|g'| \in L^p(\Omega, dxdy)$ for all $p < 4$. Gehring and Hayman showed this true for $p \in (\frac{4}{3}, 2]$. The lower bound is sharp but the upper bound as steadily increased as shown in the following table:

The Carleson-Jones conjecture:

The integral means spectrum:

5. The strong law of large numbers

In Chapter 5 we used the strong law of large numbers to compute the dimension of certain subsets of $[0,1]$ defined in terms of digit frequences. The following is a self-contained proof of the needed result.

THEOREM 131. *Let $\{f_n\}$, $n = 1, 2 \dots$ be a sequence of orthogonal functions on a probability space $(X, d\nu)$ and suppose $E(f^2) = \int |f|^2 d\nu \leq 1$. Then*

$$\frac{1}{n} S_n = \frac{1}{n} \sum_{k=1}^{n} f_k \to 0,$$

a.e. (with respect to ν) as $n \to \infty$.

PROOF. We begin with the simple observation that if $\{g_n\}$ is a sequence of functions on a probability space $(X, d\nu)$ such that

$$\sum_n \int |g_n|^2 d\nu < \infty,$$

then $\sum_n |g_n|^2 < \infty$ a.e. $(d\nu)$ and hence $g_n \to 0$ a.e. $(d\nu)$.

Using this, it is easy to verify the law of large numbers (LLN) for $n \to \infty$ along the sequence of squares. Namely,

$$\int (\frac{1}{n} S_n)^2 d\nu = \frac{1}{n^2} \int |S_n|^2 d\nu = \frac{1}{n^2} \sum_{k=1}^{n} \int |f_k|^2 d\nu \leq \frac{1}{n}.$$

Therefore if we set $g_n = \frac{1}{n^2} S_{n^2}$, we have

$$\int \frac{1}{n^2} |S_{n^2}|^2 d\nu \leq \frac{1}{n^2}.$$

Since the right hand side is summable, the observation above implies $g_n \to 0$ a.e. $(d\nu)$. This is the same as $\frac{1}{n^2} S_{n^2} \to 0$, a.e..

To deal with limit over all the integers take $m^2 \leq n < (m+1)^2$ and set $m(n) = \lfloor \sqrt{n} \rfloor$. Then

$$\begin{aligned} \int |\frac{1}{m^2} S_n - \frac{1}{m^2} S_{m^2}|^2 d\nu &= \frac{1}{m^4} \int |\sum_{k=m^2+1}^{n} f_k|^2 d\nu \\ &= \frac{1}{m^4} \int \sum_{k=m^2+1}^{n} |f_k|^2 d\nu \\ &\leq \frac{2}{m^3} \end{aligned}$$

since the sum has at most $2m$ terms, each of size at most 1. Put

$$g_n = \frac{S_n}{m(n)^2} - \frac{S_{m(n)^2}}{m(n)^2}.$$

Then since each $m = m(n)$ is associated to at most $2m+1$ different n's we get

$$\sum_{n=1}^{\infty} \int |g_n|^2 d\mu \leq \sum_{n=1}^{\infty} \frac{2}{m(n)^3} \leq \sum_{m} (2m+1)\frac{2}{m^3} < \infty,$$

so by the initial observation, $g_n \to 0$ a.e. with respect to ν. This implies $\frac{1}{m(n)^2} S_n \to 0$ a.e., which in turn implies $\frac{1}{n} S_n \to 0$ a.e., which is what we wanted. □

This version is sometimes called the strong law of large numbers because it gives a.e. convergence, as opposed to the weak version which only says that $\frac{1}{n} S_n$ converges to 0 in L^2.

As a remark we should note that better estimates for the decay of S_n are possible if we assume that the functions $\{f_n\}$ are independent with respect to the measure ν. This means that for any n and any collection of measurable sets $\{A_1, \ldots, A_n\}$ we have

$$\nu(x \in X : f_j(x) \in A_j, j = 1, \ldots, n\}) = \prod_{j=1}^{n} \nu(\{x \in X : f_j(x) \in A_j\}).$$

Roughly, this says that knowing the values of any of the f_j's at x does not give us any information about the values of the remaining functions there.

By 1915 Hausdorff had proved that if $\{f_n\} \in L^2(\nu) \cap L^1(\nu)$ are independent, orthonormal (orthogonal and have L^2 norm 1) and satisfy $\int f_n d\nu = 0$ then

$$\lim_{N\to\infty} \frac{1}{N^{\frac{1}{2}+\epsilon}} \sum_{n=0}^{N} f_n(x) = 0, \text{ for a.e. } x$$

and for every $\epsilon > 0$. After that Hardy-Littlewood, and independently Khinchin, proved

$$\lim_{N\to\infty} \frac{1}{\sqrt{N \log N}} \sum_{n=0}^{N} f_n(x) = 0 \text{ for a.e. } x.$$

The "final" result, found by Khinchin for a special case in 1928 and proved in general by Hartman-Wintree in 1941 says

$$\limsup_{N\to\infty} \frac{1}{\sqrt{2N \log\log N}} \sum_{n=0}^{N} f_n(x) = 1 \text{ for a.e. } x.$$

APPENDIX D

Background material

1. Real Analysis

LEMMA 132. *If f is continuous on a compact set, then it is uniformly confinious.*

THEOREM 133 (Theorem 3.3.8, [?]). *Suppose f is C^2 on an open set Ω and suppose $K \subset \Omega$ is compact. for $p, q \in K$, define*

$$R(p) = f(p) - f(q) - Dff(q)(p - q).$$

Then

$$\lim_{|p-q|\to 0} \frac{R(p)}{p - q} = 0.$$

The convergence is uniform, i.e., for any $\epsilon > 0$, there is an $\delta > 0$,so that

$$p, q \in K, |p - q| < \delta \Rightarrow |R(p)| \leq \epsilon |p - q|.$$

LEMMA 134. *If f_n converges uniformly to f on X then $\lim_{n\to\infty} \int_X f_n dx = \int_X f dx$.*

LEMMA 135. *If $\{f_n\}$ satisfy $\max_X |f_n| \leq a_n$ and $\sum_n a_n < \infty$, then $\int_X \sum_{n=0}^{\infty} f_n dx = \sum_{n=0}^{\infty} \int_X f_n dx$.*

2. Topology

let $p : E \to B$ be continuous and surjective. An open set $U \subset B$ is evenly covered if the inverse image $p^{-1}(U)$ can be written as a disjoint union of sets V_α so that p restricted to each V_α is a homeomorphism onto U. If every point b of B has a neighborhood U that is evenly covered by p, then p is called a covering map.

A space X is simply connected if it is path connected and if its fundamental group is tirvial, i.e., every closed loop in X can be homotoped to a point.

THEOREM 136 (Theorem 11.4.1, [?]). *If U is a bounded planar domain, then the following are equivalent:*

(1) *U is simply connected.*

(2) $\mathbb{C} \setminus U$ *is connected.*

LEMMA 137 ([**?**], Lemma 8.4.1). *Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. Any path $f[0,1] \to B$ beginning at b_0 has a unique lift to a path $\tilde{f}$ in E beginning at e_0.*

LEMMA 138 (Exercise 8.4.12(a), [**?**]). *Let $p : E \to B$ be a convering map; let $p(e_0) = b_0$. Let $f : (Y, y_0 \to (B, b_0)$ be continuous. If Y is locally path connected and simply connected then f can be lifted uniquely to a continuous map $\tilde{f} : (Y, y_0), \to (E, e_0)$.*

Bibliography