Chapter 7: Riemann Mapping Theorem

Course 414, 2003–04

March 30, 2004

7.1 Theorem (Hurwitz' Theorem). *Let G be a connected open set in $\mathbb{C}$ and $(f_n)_n$ a sequence in $H(G)$ which converges to $f \in H(G)$ (uniformly on compact subsets of G). Suppose $f \not\equiv 0$, $\bar{D}(a,R) \subset G$ and $f(z)$ is never zero on $|z-a| = R$. Then there exist n_0 such that for $n \geq n_0$, f_n and f have the same number of zeros (counting multiplicities) in $D(a,R)$.*

Proof. Since $f(z)$ is never zero on the circle, we have

$$\inf_{|z-a|=R} |f(z)| = \delta > 0.$$

For n large enough (say for $n \geq n_0$)

$$\sup_{|z-a|=R} |f_n(z) - f(z)| < \delta/2$$

and thus on the circle $|z-a| = R$ we have

$$|f(z) - f_n(z)| < \delta/2 < \delta \leq |f(z)|.$$

By Rouchés theorem, f_n and f have the same total number of zeros inside the circle $|z-a| = R$ (counting multiplicity) for $n \geq n_0$. □

7.2 Corollary. *Let G be a connected open set in $\mathbb{C}$ and (f_n) a sequence in $H(G)$ such that each f_n is never $zero$ in G. Suppose $f_n \to f$ in $H(G)$. If $f(z)$ is ever zero in G, then $f \equiv 0$.*

Proof. This follows immediately from 7.1.

If f is not identically zero but $f(a) = 0$, then by the identity theorem we can choose $R > 0$ sufficiently small that $z = a$ is the only zero of f in the closed disk $\bar{D}(z,R)$ and also $\bar{D}(z,R) \subset G$. Counting multiplicity, f will have a positive number of zeros in $D(a,R)$ and by Hurwitz so must f_n for n large. □

7.3 Definition. *If $G \subset \mathbb{C}$ is open and $f : G \to \mathbb{C}$ is an injective analytic function, then f is called a* conformal mapping *from G to $f(G)$.*

Recall that $f(G)$ is necessarily open and $f^{-1}\colon f(G) \to G$ is automatically analytic by the open mapping theorem. Also, $f'(z)$ is never zero in G and this leads to the angle-preserving property of conformal mapping that gives them their name:

If γ_1 and γ_2 are two C^1 curves in G which meet at an angle θ to one another at $a \in G$ (say $\gamma_1(0) = \gamma_2(0) = a$ and $\arg(\gamma_1'(0)) - \arg(\gamma_2'(0)) = \theta$), then $f \circ \gamma_1$ and $f \circ \gamma_2$ also meet at an angle θ at $f(a)$.

To avoid worrying about ambiguity of the argument, we could restate

$$\arg(\gamma_1'(0)) - \arg(\gamma_2'(0)) = \theta$$

as

$$e^{-i\theta}\frac{\gamma_1'(0)}{\gamma_2'(0)} > 0$$

(with θ restricted to a range such as $[0, 2\pi)$ — which means θ is the angle you would need to turn $\gamma_2'(0)$ anticlockwise to align it with $\gamma_1'(0)$).

The reasoning for the angle preserving property is that

$$e^{-i\theta}\left(\frac{(f\circ\gamma_1)'(0)}{(f\circ\gamma_2)'(0)}\right) = e^{-i\theta}\left(\frac{f'(a)(\gamma_1)'(0)}{f'(a)(\gamma_2)'(0)}\right) = e^{-i\theta}\left(\frac{\gamma_1'(0)}{\gamma_2'(0)}\right) > 0.$$

The term *conformal* really means angle-preserving at each point, but it is usual in complex analysis to use it for injective analytic functions.

7.4 Corollary. *If $G \subset \mathbb{C}$ is a connected open set and $(f_n)_n$ is a sequence of injective functions in $H(G)$ such that $f_n \to f \in H(G)$, then the limit f is either constant or injective.*

Proof. Suppose f is not constant and not injective. Then there exist $a, b \in G$, $a \neq b$, with $f(a) = f(b) = w$. Taking $f_n(z) - w$ and $f(z) - w$ instead of the original f_n and f, we can assume that $w = 0$.

We can find a positive $\delta < |a - b|/2$ so that $\bar{D}(a, \delta) \subset G$, $\bar{D}(b, \delta) \subset G$ and $f(z)$ never 0 on the circles $|z - a| = \delta$, $|z - b| = \delta$. Applying Hurwitz' theorem (7.1), we find that for all large n there exists $a_n \in D(a, \delta)$ with $f_n(a_n) = 0 = w$. Similarly, for all large n there is $b_n \in D(b, \delta)$ with $f_n(b_n) = 0 = w$. By the choice of δ, $D(a, \delta) \cap D(b, \delta) = \emptyset$ and therefore $a_n \neq b_n$. Thus $f_n(a_n) = f_n(b_n) = 0$ and this contradicts injectivity of f_n. □

7.5 Examples. (i) The map

$$\begin{aligned} f : \{z \in \mathbb{C} \colon -\pi/2 < \Im(z) < \pi/2\} &\to \{z \in \mathbb{C} \colon \Re(z) > 0\} \\ f(z) &= e^z \end{aligned}$$

is a conformal mapping (of its domain onto its target set, the right half plane).

(ii) There is no conformal map ϕ of the unit disc $D(0, 1)$ onto the whole complex plane $\mathbb{C}$ (because the inverse function $\phi^{-1} \colon \mathbb{C} \to D(0, 1)$ would be a bounded entire function which was not constant, contradicting Liouvilles theorem).

(iii) If $ad - bc \neq 0$, then the function

$$\phi(z) = \frac{az + b}{cz + d}$$

is called a *Möbius transformation* or a linear fractional transformation. Since these maps are rational functions, we can (and often do) regard them as analytic functions

$$\phi\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}.$$

The condition $ad - bc \neq 0$ is there to rule out constant functions and the 'missing' values are $\phi(\infty) = a/c$, $\phi(-d/c) = \infty$ (unless $c = 0$, in which case $a \neq 0$ and $d \neq 0$, and we have $\phi(z) = (a/d)z + (b/d)$, $\phi(\infty) = \infty$.)

7.6 Proposition. *Let ϕ, ψ be Möbius transformations. Then*

(i) *$\phi \circ \psi$ is a Möbius transformation.*

(ii) *$\phi\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a bijection and its inverse $\phi^{-1}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is also a Möbius transformation.*

(iii) *The Möbius transformations form a group (under composition).*

Proof. Exercise. □

7.7 Proposition. *Every Möbius transformation can be expressed as a composition of Möbius transformations of the following kinds:*

(i) *$z \mapsto z + a$ (translation)*

(ii) *$z \mapsto \lambda z$ ($\lambda > 0$) (dilation)*

(iii) *$z \mapsto e^{i\alpha} z$ ($\alpha \in \mathbb{R}$) (rotation)*

(iv) *$z \mapsto \frac{1}{z}$.*

Proof. If $\phi(z) = \frac{az+b}{cz+d}$ and $c \neq 0$, we can write

$$\begin{aligned} \phi(z) &= \frac{\frac{a}{c}(cz + d) + \left(b - \frac{ad}{c}\right)}{cz + d} \\ &= \frac{a}{c} + \left(\frac{b}{c} - \frac{ad}{c^2}\right) \frac{1}{z + \frac{d}{c}} \\ &= \phi_5 \circ \phi_4 \circ \phi_3 \circ \phi_2 \circ \phi_1(z), \end{aligned}$$

where

$$\begin{aligned}
\phi_1(z) &= z + \frac{d}{c} \\
\phi_2(z) &= \frac{1}{z} \\
\phi_3(z) &= \left|\frac{b}{c} - \frac{ad}{c^2}\right| z \\
\phi_4(z) &= e^{i\alpha} z \quad \left(\alpha = \arg\left(\frac{b}{c} - \frac{ad}{c^2}\right)\right) \\
\phi_5(z) &= \frac{a}{c} + z
\end{aligned}$$

In the case $c = 0$,

$$\phi(z) = (a/d)z + (b/d) = \left|\frac{a}{d}\right| e^{i\beta} z + \frac{b}{d},$$

with $\beta = \arg(a/d)$, and thus ϕ is a composition of a rotation, a dilation and a translation. □

7.8 Proposition. *Let $P : \mathbb{C} \to S^2$ denote the stereographic projection map from $\mathbb{C}$ to the Riemann sphere. If C is a circle in $\mathbb{C}$, then its image $P(C)$ is a circle on S^2. If L is a straight line in $\mathbb{C}$, then $P(L) \cup \{(0, 0, 1)\}$ is also a circle on S^2.*

Conversely, if C_1 is a circle on S^2, then $P^{-1}(C_1)$ is either a circle or a line in $\mathbb{C}$.

In other words, circles on the Riemann sphere correspond to circles and lines in the plane under stereographic projection (but the point at ∞ or the north pole has to be added to lines to close the circle on the sphere).

Proof. A circle (or line) in $\mathbb{C}$ has an equation

$$A(x^2 + y^2) + 2\beta x + 2\gamma y + C = 0$$

with $A, \beta, \gamma, C \in \mathbb{R}$. In complex terms, this can be rewritten in the form

$$Az\bar{z} + \bar{B}z + B\bar{z} + C = 0.$$

($B = \beta + i\gamma$.) If $A \neq 0$ we can multiply across by A and rearrange this as

$$|Az + B|^2 = (Az + B)(A\bar{z} + \bar{B}) = A^2|z|^2 + A\bar{B}z + AB\bar{z} + |B|^2 = |B|^2 - AC$$

and the condition $|B|^2 > AC$ corresponds to a positive radius (not an empty 'circle' or a single point). If $A = 0$, then $|B|^2 > AC = 0$ is the condition for the equation to be that of a genuine line.

Now

$$P^{-1}(\xi, \eta, \zeta) = \frac{\xi + i\eta}{1 - \zeta},$$

and we can use this in the equation of the circle/line to say that the points (ξ, η, ζ) on the stereographic projection of the circle/line onto the sphere must be those that satisfy

$$A\frac{(\xi+i\eta)(\xi-i\eta)}{(1-\zeta)^2} + \bar{B}\frac{\xi+i\eta}{1-\zeta} + B\frac{\xi-i\eta}{1-\zeta} + C = 0$$

(and for lines the north pole $P(\infty)$ as well). Observe that

$$\frac{(\xi+i\eta)(\xi-i\eta)}{(1-\zeta)^2} = \frac{\xi^2+\eta^2}{(1-\zeta)^2} = \frac{1-\zeta^2}{(1-\zeta)^2} = \frac{1+\zeta}{1-\zeta}$$

because $\xi^2+\eta^2+\zeta^2 = 1$. Using this, we can simplify the equation to get

$$2\beta\xi + 2\gamma\eta + (A-C)\zeta + (A+C) = 0.$$

This is the equation of a plane in $\mathbb{R}^3$, which intersects the sphere in a circle. Hence P maps circles and lines to circles.

Conversely, starting with the equation of any plane in $\mathbb{R}^3$

$$a_1\xi + a_2\eta + a_3\zeta + c = 0.$$

We can choose $\beta = a_1/2$, $\gamma = a_2/2$, $A = (a_3+c)/2$ and $C = (c-a_3)/2$. The point on the plane closest to the origin is

$$(\xi, \eta, \zeta) = \frac{-c}{a_1^2+a_2^2+a_3^2}(a_1, a_2, a_3)$$

and this is inside the sphere exactly when

$$\sqrt{a_1^2+a_2^2+a_3^2} > |c|.$$

So the condition for the plane to intersect the sphere in more than one point is

$$c^2 = (A+C)^2 < a_1^2+a_2^2+a_3^2 = 4\beta^2+4\gamma^2+(A-C)^2$$

or

$$4AC < 4(\beta^2+\gamma^2).$$

This is exactly what is needed to get a genuine circle or line. □

7.9 Theorem. *Let $\mathcal{C}$ denote the class of all circles in $\hat{\mathbb{C}}$ (by which we mean all circles in $\mathbb{C}$ together with the sets $L \cup \{\infty\}$ with L a line in $\mathbb{C}$). Then Möbius transformations map $\mathcal{C}$ to $\mathcal{C}$.*

Proof. The equation of a circle or line in $\mathbb{C}$ can be written in the form

$$Az\bar{z} + \bar{B}z + B\bar{z} + C = 0,$$

with $A, C \in \mathbb{R}$, $B \in \mathbb{C}$, and $|B|^2 > AC$.

It is quite easy to verify that an equation of this kind transforms to another equation of the same type under the basic kinds of Möbius transformations (translations, rotations, dilations and inversion in the unit circle). Since every Möbius transformation is a composition of these, the result follows. □

7.10 Proposition. *Given two lists of three distinct point in $\hat{\mathbb{C}}$, say z_1, z_2, z_3 and w_1, w_2, w_3, there exists a unique Möbius transformation ϕ satisfying $\phi(z_j) = w_j$ ($j = 1, 2, 3$).*

Proof. It is sufficient (for the existence part) to show that we can always find ϕ with

$$\phi(z_1) = 0,\ \phi(z_2) = 1,\ \phi(z_3) = \infty,$$

because if we can do this, we can find ψ which does the same for w_1, w_2, w_3 and then we will have $(\psi^{-1} \circ \phi)(z_j) = w_j$.

If $z_1, z_2, z_3 \in \mathbb{C}$, the required ϕ is provided by

$$\phi(z) = \frac{z_2 - z_3}{z_2 - z_1}\frac{z - z_1}{z - z_3}.$$

If one of the z_j is ∞, we have the following formulae which work

$$\phi(z) = \begin{cases} (z_2 - z_3)\dfrac{1}{z - z_3} & \text{if } z_1 = \infty \\ \dfrac{z - z_1}{z - z_3} & \text{if } z_2 = \infty \\ \dfrac{1}{z_2 - z_1}(z - z_1) & \text{if } z_3 = \infty \end{cases}$$

To establish the uniqueness part of the result, suppose ϕ and ψ are two Möbius transformations that map the z_j's to the corresponding w_j's. Then $\psi^{-1} \circ \phi$ is a Möbius transformation that fixes the z_j, $j = 1, 2, 3$. But if

$$(\psi^{-1} \circ \phi)(z) = \frac{az + b}{cz + d},$$

the equation

$$\frac{az + b}{cz + d} = z$$

multiplies out to

$$az + b = cz^2 + dz.$$

This is a quadratic equation which can only have two solutions unless it simplifies to $0 = 0$ — which means $b = c = 0$, $a = d$ and

$$\frac{az + b}{cz + d} = \frac{az}{d} = z.$$

This reasoning works too if we consider the case where one of the z_j is ∞ — if $(az+b)/(cz+d) = z$ has ∞ as a solution, then $c = 0$ and then the equation we get on multiplying out is linear (only one finite solution unless it reduces to $0 = 0$). □

7.11 Theorem. *If $|a| < 1$ and $|\lambda| = 1$, then*

$$\phi(z) = \lambda\frac{z - a}{1 - \bar{a}z}$$

is a Möbius transformation that maps the unit disc $D(0, 1)$ bijectively to itself and also maps the unit circle $|z| = 1$ to itself.

Proof. If $|z| = 1$, then

$$\begin{aligned} |\phi(z)| &= \left|\lambda \frac{z-a}{1-\bar{a}z}\right| \\ &= \left|\frac{z-a}{1-\bar{a}z}\right| \\ &= \left|\frac{1}{\bar{z}}\frac{z-a}{1-\bar{a}z}\right| \text{ since } |z| = |\bar{z}| = 1 \\ &= \left|\frac{z-a}{\bar{z}-\bar{a}|z|^2}\right| \\ &= \left|\frac{z-a}{\bar{z}-\bar{a}}\right| \text{ since } |z| = 1 \\ &= 1 \end{aligned}$$

Thus ϕ maps the unit circle into itself. The image of the unit circle under ϕ must be the whole circle by 7.9, and thus ϕ maps the unit circle to itself bijectively. Since $\phi\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a bijection, we must also have a bijection $\phi\colon \hat{\mathbb{C}} \setminus \{|z| = 1\} \to \hat{\mathbb{C}} \setminus \{|z| = 1\}$.

Note that ϕ is continuous and $\hat{\mathbb{C}} \setminus \{|z| = 1\}$ has two connected components — the disc $D(0,1)$ and the complement of the closed unit disc in $\hat{\mathbb{C}}$. Now $\phi(0) = -\lambda a \in D(0,1)$, and hence $\phi(D(0,1)) \subset D(0,1)$. Since $\phi(1/\bar{a}) = \infty$, the same reasoning tells us that ϕ maps the exterior of the closed disc into itself. Since ϕ is a bijection of $\hat{\mathbb{C}}$ to $\hat{\mathbb{C}}$, we must have that $\phi(D(0,1)) = D(0,1)$ (and also that ϕ maps the exterior of the unit disc onto itself). □

7.12 Corollary. *If $a, b \in D(0,1)$, then there exist a conformal map $\phi\colon D(0,1) \to D(0,1)$ with $\phi(a) = b$.*

In fact the set of conformal maps of the unit disc onto itself forms a group under composition (easy to check) and this corollary can be stated as saying that this group of self-transformations of the unit disc acts *transitively* on the disc.

Proof. Take

$$\phi_1(z) = \frac{z-a}{1-\bar{a}z}, \quad \phi_2(z) = \frac{z-b}{1-\bar{b}z}.$$

Then $\phi_1(a) = 0$, $\phi_2(b) = 0$ and ϕ_1, ϕ_2 are conformal maps of the disc to itself. The one we want is $\phi_2^{-1} \circ \phi_1$, which satisfies

$$(\phi_2^{-1} \circ \phi_1)(a) = b.$$

□

7.13 Theorem (Schwarz Lemma). *Let $f : D(0,1) \to D(0,1)$ be an analytic function which maps the unit disc $D(0,1)$ to itself. If $f(0) = 0$, then*

(i) $|f(z)| \leq |z|$ *for* $0 < |z| < 1$*;*

(ii) $|f'(0)| \leq 1$*;*

(iii) If equality holds in (ii) or if equality holds in (i) for any single $z \neq 0$, then

$$f(z) \equiv \lambda z$$

where λ is a constant of modulus $|\lambda| = 1$.

Proof. Let $g(z) = \frac{f(z)}{z}$.

Then $g(z)$ is analytic for $0 < |z| < 1$ and it is has a removable singularity at $z = 0$ since $f(0) = 0$. It becomes analytic at 0 if we define

$$g(0) = \lim_{z \to 0} \frac{f(z)}{z} = f'(0).$$

Fix a number $0 < r < 1$. For $|z| = r$

$$|g(z)| = \left|\frac{f(z)}{z}\right| < \frac{1}{r}.$$

By the maximum modulus theorem for the analytic function g, it follows that $|g(z)| < 1/r$ for $|z| \leq r$. Fix $z \in D(0,1)$ and let $r \to 1^-$ to get

$$|g(z)| \leq 1.$$

This is true for all $z \in D(0,1)$ and we deduce

$$\left|\frac{f(z)}{z}\right| \leq 1, \quad (0 < |z| < 1)$$

and from this

$$|f(z)| \leq |z| \quad (0 \neq z \in D(0,1)).$$

This proves (i) and the fact that $|g(0)| \leq 1$ proves (ii).

To prove (iii), observe that if we get equality, then we must have a point$z \in D(0,1)$ where $|g(z)| = 1$ — in other words, a point where the maximum modulus of g is attained. By the maximum modulus theorem again, this means that g must be a constant function — $g(z) \equiv \lambda$ and then we must have $|\lambda| = 1$. Hence $f(z) = \lambda z$. □

7.14 Corollary. *Let $\phi\colon D(0,1) \to D(0,1)$ be any conformal map of the unit disc onto itself. Then there exist $|\lambda| = 1$ and $a \in D(0,1)$ such that*

$$\phi(z) = \lambda \frac{z - a}{1 - \bar{a}z}$$

(that is, ϕ is a Möbius transformation of the kind studied above).

Proof. Suppose $\phi(0) = a$. Then let

$$\psi(z) = \frac{\phi(z) - a}{1 - \bar{a}\phi(z)}.$$

Being the composition of two conformal mappings, $\psi\colon D(0,1) \to D(0,1)$ is a conformal map and $\psi(0) = 0$.

By the Schwarz lemma (7.13), $|\psi'(0)| \leq 1$.

But also, $\psi^{-1}\colon D(0,1) \to D(0,1)$ is conformal and has $\psi^{-1}(0) = 0$. Again by the Schwarz lemma,

$$\begin{aligned} |(\psi^{-1})'(0)| &\leq 1 \\ \left|\frac{1}{\psi'(\psi^{-1}(0))}\right| &\leq 1 \\ \left|\frac{1}{\psi'(0)}\right| &\leq 1 \\ |\psi'(0)| &\geq 1 \end{aligned}$$

Combining with $|\psi'(0)| \leq 1$ gives $|\psi'(0)| = 1$. Therefore, by the Schwarz lemma,

$$\begin{aligned} \psi(z) &\equiv \lambda z \ \ (|\lambda| = 1) \\ \frac{\phi(z) - a}{1 - \bar{a}\phi(z)} &\equiv \lambda z \\ \phi(z) &= \frac{\lambda z + a}{1 + \lambda\bar{a}z} \\ &= \lambda\frac{z + \bar{\lambda}a}{1 + \lambda\bar{a}z} \end{aligned}$$

□

7.15 Theorem (Riemann Mapping Theorem). *Let $G \subset \mathbb{C}$ be a simply connected connected open set with $G \neq \mathbb{C}$. Let $a \in G$ be arbitrary. Then there exists a (unique) conformal map $f : G \to D(0,1)$ of G onto the unit disc which satisfies*

$$f(a) = 0 \text{ and } f'(a) > 0.$$

Proof of uniqueness. If there was a second conformal map $g : G \to D(0,1)$ with the given properties, then

$$f \circ g^{-1}\colon D(0,1) \to D(0,1)$$

would be a conformal map and it would satisfy

$$(f \circ g^{-1})(0) = f(a) = 0.$$

Hence by 7.14,

$$(f \circ g^{-1})(z) \equiv \lambda z \text{ for some } \lambda \text{ with } |\lambda| = 1.$$

Looking at the derivative at the origin, we find

$$(f \circ g^{-1})'(0) = f'(g^{-1}(0))(g^{-1})'(0) = f'(a)\frac{1}{g'(g^{-1}(0))} = \frac{f'(a)}{g'(a)} > 0.$$

But, calculating it another way, we get λ. Hence λ is a positive real number. But also $|\lambda| = 1$ and that means $\lambda = 1$. Thus $f \circ g^{-1}$ is the identity function z and $f = g$.

The proof of existence will be divided into several steps.

Lemma 7.16. *If $G \subset \mathbb{C}$ is a simply connected connected open set and $f : G \to \mathbb{C}$ is analytic and never vanishes, then f has an analytic square root — that is, there exists an analytic $g : G \to \mathbb{C}$ with $g(z)^2 \equiv f(z)$.*

Proof. There exists a branch of $\log f$ in G (see chapter 2).

$$g(z) = e^{(1/2)\log f(z)}$$

will do. □

7.17 Lemma. *If $G \subset \mathbb{C}$ ($G \neq \mathbb{C}$) is a connected open set with the property that every nowhere-vanishing analytic function on G has an analytic square root, then there exists an injective analytic function f on G with $f(G) \subset D(0,1)$.*

Proof. Pick $b \in \mathbb{C} \setminus G$. Then there exists $g : G \to \mathbb{C}$ analytic with $g(z)^2 = z - b$. This g is injective since

$$g(z_1) = g(z_2) \Rightarrow z_1 - b = g(z_1)^2 = g(z_2)^2 = z_2 - b \Rightarrow z_1 = z_2.$$

Now, by the open mapping theorem, $g(G)$ is open. Pick $w_0 \in g(G)$ and choose $r > 0$ so that $D(w_0, r) \subset g(G)$. Then $D(-w_0, r) \subset \mathbb{C} \setminus g(G)$ because if there exists a point $w \in D(-w_0, r) \cap g(G)$, then $w = g(z_1)$ for some $z_1 \in G$ and also $-w \in D(w_0, r) \subset g(G)$ which tells us that $-w = g(z_2)$ for some $z_2 \in G$. Now

$$\begin{aligned} g(z_1) = -g(z_2) &\Rightarrow g(z_1)^2 = g(z_2)^2 \\ &\Rightarrow z_1 - b = z_2 - b \\ &\Rightarrow z_1 = z_2 \\ &\Rightarrow g(z_2) = g(z_1) = -g(z_2) \\ &\Rightarrow g(z_2) = 0 \\ &\Rightarrow 0 = g(z_2)^2 = z_2 - b \\ &\Rightarrow z_2 = b \in \mathbb{C} \setminus G \end{aligned}$$

and this contradicts $z_2 \in G$.

Put

$$f(z) = \frac{r}{2(g(z) + w_0)}.$$

Then f is analytic on G ($|g(z) + w_0| \geq r$ for all $z \in G$ is proved above, and shows that the denominator is never 0), injective on G since g is and also f satisfies $|f(z)| \leq 1/2 < 1$ for $z \in G$. □

7.18 Lemma. *With the same hypothesis on G as in Lemma 7.17, and for $a \in G$ a given point, there exists an injective analytic function on G satisfying*

$$\begin{aligned} f(G) &\subset D(0,1) \\ f(a) &= 0 \\ f'(a) &> 0 \end{aligned}$$

Proof. Replace the $f(z)$ obtained in Lemma 7.17 by

$$\lambda \frac{f(z) - f(a)}{1 - \overline{f(a)} f(z)}$$

with $|\lambda| = 1$, λ suitably chosen to make the derivative at a positive. □

7.19 Lemma. *Let $G \subset \mathbb{C}$ be a connected open set , $G \neq \mathbb{C}$, with the property that every nowhere-vanishing analytic function has an analytic square root. Let $a \in G$ be fixed. Then there exists a conformal map $f : G \to D(0,1)$ of G onto the unit disc with the properties $f(a) = 0$ and $f'(a) > 0$.*

Observe that this lemma will complete the proof of the Riemann mapping theorem (7.15).

Proof. Let $\mathcal{F}$ denote the family of all analytic functions $f : G \to \mathbb{C}$ such that either $f \equiv 0$ or else

$$f \text{ is injective, } f(G) \subset D(0,1), f(a) = 0 \text{ and } f'(a) > 0.$$

By Lemma 7.18, $\mathcal{F}$ contains nonzero functions.

By Montels theorem $\mathcal{F}$ is relatively compact in $H(G)$. By Corollary 7.4 to Hurwitz' theorem, we can see that all functions in the closure of $\mathcal{F}$ in $H(G)$ are either constant or injective. Since all the function in $\mathcal{F}$ have the value 0 at a, this must be true of all functions in the closure of $\mathcal{F}$. Hence the only constant function in the closure is 0 (which is in $\mathcal{F}$).The other functions f in the closure must have $f(G) \subset \bar{D}(0,1)$. Since $f(G)$ is open (open mapping theorem), $f(G) \subset D(0,1)$. Also, because the map $f \mapsto f'(a) : H(G) \to \mathbb{C}$ is continuous, all functions in the closure of $\mathcal{F}$ must satisfy $f'(a) \geq 0$. Since they are injective if not identically zero, $f'(a) \neq 0$ and so $f'(a) > 0$ (unless $f \equiv 0$). The conclusion from these observations is that $\mathcal{F}$ is actually closed in $H(G)$. Hence $\mathcal{F}$ is compact in $H(G)$.

Since the map $f \mapsto f'(a) : \mathcal{F} \to \mathbb{R}$ is a continuous function on a compact set, it must attain its maximum value. Let $f \in \mathcal{F}$ be a function with $f'(a)$ as large as possible. As remarked earlier, f is not the zero function and $f'(a) > 0$.

We now show that this f maps G onto the unit disc (and so it is the required Riemann mapping).

Suppose for the sake of obtaining a contradiction that $f(G) \neq D(0,1)$ and choose $w \in D(0,1) \setminus f(G)$. By the assumption on G about analytic square roots, we can find $h \in H(G)$ with

$$h(z)^2 = \frac{f(z) - w}{1 - \bar{w} f(z)}. \tag{1}$$

Let

$$g(z) = \lambda \frac{h(z) - h(a)}{1 - \overline{h(a)} h(z)}$$

where $|\lambda| = 1$ and λ is chosen to ensure that $g'(a) > 0$.

Now $g \in \mathcal{F}$ and

$$g'(a) = |g'(a)| = |(\phi \circ h)'(a)|,$$

where $\phi(\zeta) = (\zeta - h(a))/(1 - \overline{h(a)}\zeta)$. This leads to

$$g'(a) = |\phi'(h(a))h'(a)| = \frac{1}{1 - |h(a)|^2} |h'(a)|.$$

Differentiating (1) gives

$$2h(z)h'(z) = \frac{1 - \bar{w} f(z) + \bar{w}(f(z) - w)}{(1 - \bar{w} f(z))^2} f'(z).$$

Evaluating at $z = a$ we get

$$2h(a)h'(a) = (1 - |w|^2) f'(a).$$

$$|h'(a)| = \frac{f'(a)(1 - |w|^2)}{2|h(a)|} = \frac{f'(a)(1 - |w|^2)}{2\sqrt{|w|}}$$

(recall $h(a)^2 = (f(a) - w)/(1 - \bar{w} f(a)) = -w$).

Thus

$$\begin{aligned} g'(a) &= \frac{1}{1 - |h(a)|^2} |h'(a)| \\ &= \frac{1}{1 - |w|} \frac{1 - |w|^2}{2\sqrt{|w|}} f'(a) \\ &= \frac{1 + |w|}{2\sqrt{|w|}} f'(a) \\ &> f'(a). \end{aligned}$$

The last step is justified because

$$\begin{aligned} 1+|w| &> 2\sqrt{|w|} \\ 1-2\sqrt{|w|}+|w| &> 0 \\ (1-\sqrt{|w|})^2 &> 0. \end{aligned}$$

The fact that $g'(a) > f'(a)$ contradicts the choice of $f \in \mathcal{F}$ as maximising $f'(a)$. This contradiction shows that $f(G) = D(0,1)$. □

7.20 Theorem. *Let $G \subset \mathbb{C}$ be a connected open set. Then the following are equivalent properties for G*

(i) *G is simply connected*

(ii) *$\mathrm{Ind}_\gamma(a) = 0$ for each piecewise C^1 closed curve in G and each $a \in \mathbb{C} \setminus G$.*

(iii) *For each analytic function $f : G \to \mathbb{C}$ and each piecewise C^1 closed curve γ in G, $\int_\gamma f(z)\,dz = 0$.*

(iv) *Every analytic function $f : G \to \mathbb{C}$ has an antiderivative.*

(v) *For every analytic function $f : G \to \mathbb{C}$ which never vanishes in G, there is a branch of $\log f$ in G.*

(vi) *Every never-vanishing analytic $f : G \to \mathbb{C}$ has an analytic square root.*

(vii) *G is homeomorphic to the unit disc $D(0,1)$ (that is, there exists a continuous bijection $f : G \to D(0,1)$ with a continuous inverse f^{-1}).*

(viii) *$\hat{\mathbb{C}} \setminus G$ is connected (this says exactly that G has no "holes" — no bounded connected components of $\mathbb{C} \setminus G$).*

Proof. We know

$$\text{(i)} \Rightarrow \text{(ii)} \iff \text{(iii)} \iff \text{(iv)} \iff \text{(v)}$$

(see Chapter 2).

(v) ⇒ (vi) is easy — $e^{(1/2)\log f}$ is an analytic square root of f.

(vi) ⇒ (vii): In the case $G \neq \mathbb{C}$, Lemma 7.19 gives this. For $G = \mathbb{C}$, $f(z) = z/(1+|z|)$ is a homeomorphism of $\mathbb{C}$ to the unit disc.

(vii) ⇒ (i) is straightforward — since the disc is simply connected, so is anything homeomorphic to the disc.

It remains therefore to show that (viii) is equivalent to the rest.

(viii) ⇒ (ii): First, the assumption implies that every connected component of $\mathbb{C} \setminus G$ is unbounded. A bounded connected component E of $\mathbb{C} \setminus G$ would be open and closed in $\mathbb{C} \setminus G$. Hence E would be closed in $\mathbb{C}$ since $\mathbb{C} \setminus G$ is closed. As E is also bounded, it would be closed in $\hat{\mathbb{C}}$. Hence E is closed in $\hat{\mathbb{C}} \setminus G$. E open in $\mathbb{C} \setminus G$ implies $E = (\mathbb{C} \setminus G) \cap U$ with $U \subset \mathbb{C}$ open.

Hence $E = (\hat{\mathbb{C}} \setminus G) \cap U$ is open in $\hat{\mathbb{C}} \setminus G$. Being an open and closed subset it must be a union of connected components of $\hat{\mathbb{C}} \setminus G$, but that is impossible if $\hat{\mathbb{C}} \setminus G$ is connected and contains $\infty \notin E$.

To show (ii) take any piecewise C^1 closed curve γ in G. We know $\text{Ind}_\gamma(w)$ is constant on connected components of $\mathbb{C} \setminus \gamma$. Hence constant on connected components of $\mathbb{C} \setminus G \subset \mathbb{C} \setminus \gamma$. (Each connected component of the smaller set is contained in a connected component of the larger one.) But we also know that there is just one unbounded connected component of $\mathbb{C} \setminus \gamma$ and $\text{Ind}_\gamma = 0$ there. As all components of $\mathbb{C}\setminus G$ are unbounded, they must all be in the unbounded component of $\mathbb{C} \setminus \gamma$. Hence $\mathbb{C} \setminus G$ is contained in the unbounded component of $\mathbb{C} \setminus \gamma$ and so we must have $\text{Ind}_\gamma(w) = 0$ for all $w \in \mathbb{C} \setminus G$.

(iv) $\Rightarrow$ (viii): Suppose that (iv) holds but $\hat{\mathbb{C}} \setminus G$ is not connected. Then we can write

$$\hat{\mathbb{C}} \setminus G = K \cup L$$

where K and L are disjoint nonempty relatively closed subsets of $\hat{\mathbb{C}} \setminus G$. Since G is open in $\mathbb{C}$, it is also open as a subset of $\hat{\mathbb{C}}$ and hence $\hat{\mathbb{C}} \setminus G$ is closed in $\hat{\mathbb{C}}$. Hence K and L are closed in $\hat{\mathbb{C}}$ (even compact). Now $\infty \in K$ or $\infty \in L$ — say $\infty \in L$. Then $K \subset \mathbb{C}$ and K is compact. $\tilde{G} = G \cup K$ is open in $\mathbb{C}$.

Essentially, the idea is that we can find a finite number of closed curves γ_j in $\tilde{G} \setminus K = G$ with

$$\sum_j \text{Ind}_{\gamma_j}(a) = 1 \text{ for all } a \in K$$

This is contradicts (iv) because $f(z) = 1/(z-a)$ is analytic on G, hence has an antiderivative by (iv) and so

$$\text{Ind}_{\gamma_j}(a) = \frac{1}{2\pi i} \int_{\gamma_j} \frac{1}{z-a}\, dz = 0$$

for each j.

The proof that we can find such γ_j is a bit tricky and so we will check out something a little simpler which is good enough for our purpose. We use here the notation that for $a, b \in \mathbb{C}$, $[a, b]$ denotes the line segment from a to b, or the parameterised path $\Gamma(t) = (1-t)a + tb$ $(0 \le t \le 1)$. □

Lemma 7.21. *If $\tilde{G} \subset \mathbb{C}$ is open and $K \subset \tilde{G}$ is compact then we can find a finite number of closed line segments $\Gamma_j = [a_j, b_j]$ $(j = 1, 2, \ldots, n)$ in $\tilde{G} \setminus K$ such that*

1. *For each $f : \tilde{G} \to \mathbb{C}$ analytic and $a \in K$,*

$$\sum_{j=1}^{n} \int_{\Gamma_j} \frac{f(z)}{z-a}\, dz = 2\pi i f(a) \tag{2}$$

2. *The number of times any given point w occurs in the list $a_1, a_2, \ldots, a_n$ of initial points of the line segments is the same as the number of times w occurs in the list $b_1, b_2, \ldots, b_n$ of the end points.*

Armed with this Lemma, we can use it for $f(z) = 1$ to get

$$\sum_{j=1}^{n} \int_{\Gamma_j} \frac{1}{z-a}\, dz = 2\pi i \qquad (a \in K).$$

Using our assumption (iv) tells us that $1/(z-a)$ has an antiderivative (call it g) on $G = \tilde{G} \setminus K$ and so we have

$$\sum_{j=1}^{n} \int_{\Gamma_j} \frac{1}{z-a}\, dz = \sum_{j=1}^{n} g(b_j) - g(a_j) = \sum_{j=1}^{n} g(b_j) - \sum_{k=1}^{n} g(a_k) = 0$$

by the second property in the Lemma. This contradiction shows that (iv) $\Rightarrow$ (viii), once we verify the Lemma.

Proof. (of Lemma 7.21)

If $\tilde{G} = \mathbb{C}$ we can take $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ to be the four sides of a large square, where the square contains K in its interior and the sides are traversed anticlockwise. Statement 1 of the Lemma is true by the Cauchy integral formula.

The case where K is empty is also trivial.

Assuming that $\tilde{G} \neq \mathbb{C}$ and $K \neq \emptyset$. there is a positive shortest distance

$$\delta = \inf\{|a-w| : a \in K, w \in \mathbb{C} \setminus \tilde{G}\}$$

from K to the complement of $\tilde{G}$.

We now consider an infinite grid on the plane of squares with sides parallel to the real and imaginary axes and a fixed side length $r < \delta/\sqrt{2}$. To be exact, take all the squares of the grid with vertices at the points $nr + imr$ $(n, m \in \mathbb{Z})$. By boundedness of K only a finite number of these square can have a point of K in them. We list all the squares $S_1, S_2, \ldots, S_N$ of the grid which intersect K at any point (inside or on the edge of S_j). To make things more precise let us say that S_j means the set of points inside or on the edge of the square. Because the squares have diagonal $r\sqrt{2} < \delta$, no two points of S_j can be separated by as much as distance δ. Since there is one of the points of S_j that belongs to K, we can conclude from the choice of δ that $S_j \subset \tilde{G}$.

Now write down for each S_j the 4 sides of it so that the boundary square is traversed anti-clockwise. We end up with $4N$ line segments $\Gamma_1^*, \Gamma_2^*, \ldots, \Gamma_{4N}^*$ and let us write

$$\Gamma_j^* = [a_j^*, b_j^*]$$

From Cauchys integral formula we have that

$$\int_{\partial S_j} \frac{f(z)}{z-a}\, dz = 2\pi i f(a)$$

for a in the interior of S_j (and any $f \in H(\tilde{G})$). The integral is zero for a outside S_j and so

$$\sum_{j=1}^{N} \int_{\partial S_j} \frac{f(z)}{z-a}\, dz = \sum_{k=1}^{4N} \int_{\Gamma_k^*} \frac{f(z)}{z-a}\, dz = 2\pi i f(a)$$

holds for a in the union of the interiors of the S_j.

Also observe that each corner of S_j occurs once as a starting point and once as an ending point of a side of S_j. Hence the number of times any given point occurs among a_k^* ($1 \le k \le 4N$) is the same as the number of times it occurs among b_k^* ($1 \le k \le 4N$).

Some segments may occur twice in the list $\Gamma_1^*, \Gamma_2^*, \dots, \Gamma_{4N}^*$. This happens if two adjacent squares of the original grid meet K, but then the two segments will be traversed in opposite directions. We can omit both from the list $\Gamma_1^*, \Gamma_2^*, \dots, \Gamma_{4N}^*$ without disturbing either the property

$$\sum_k \int_{\Gamma_k^*} \frac{f(z)}{z-a}\, dz = 2\pi i f(a)$$

or the property that every endpoint occurs as often as every initial point. Removing all such pairs of segments, we end up with a list $\Gamma_1, \Gamma_2, \dots \Gamma_n$ (where $\Gamma_j = [a_j, b_j]$).

We then have (2) for a in the union of the interiors of the S_j. Both sides of (2) give continuous functions of a. The left side is continuous for a not on $\bigcup_j \Gamma_j$ while the right hand side is continuous on $\tilde{G}$. It follows that the equality persists at all points $a \in \tilde{G}$ which are limits of points in the union of the interiors of the S_j, but not in $\bigcup_j \Gamma_j$. This will include all $a \in K$ because if $a \in K$ then a belongs in at least one S_j. If a is an interior point of S_j, then we already know (2) holds. If a is on the edge of S_j but not a corner of S_j, the adjacent square on the original grid must also be one of the S_k and the side containing a would be canceled out. If a is a corner of S_j then all 4 squares of the grid that meet at that corner are among the S_k and all 4 segments that contain a (East, North, West and South of a) must have been canceled out.

Thus all remaining segments Γ_j are in $\tilde{G} \setminus K$ and (2) holds for $a \in K$. □

7.22 Remark. As a consequence of the Riemann mapping theorem any questions about analytic functions on a simply connected connected open set $G \subset \mathbb{C}$ can be reduced to questions on two simple domains — $G = \mathbb{C}$ and $G = D(0,1)$. However, translating some problems from G to the unit disc may require some specific knowledge about the "Riemann mapping function" $f : G \to D(0,1)$, or equivalently information about its inverse $f^{-1} \colon D(0,1) \to G$.

This leads to the study of problems concerning injective analytic functions on the unit disc. One such problem attracted a lot of attention since the early part of the 20th century, although it actually does not seem to have much practical value. It was known as the Bieberbach conjecture and relates to the power series coefficients of injective functions on the unit disc. It said that if $\sum_{n=0}^{\infty} a_n z^n$ is the power series of an injective analytic function on the disc, then

$$\frac{|a_n|}{|a_1|} \le n$$

It is possible to have equality. The function

$$k(z) = \sum_{n=1}^{\infty} n z^n = \frac{z}{(1-z)^2}$$

is injective on the unit disc. It is known as the Koebe function.

Bieberbach proved that $|a_2| \leq 2|a_1|$ for any injective analytic function on the disc and that if $|a_2| = 2|a_1|$, then

$$\sum_{n=0}^{\infty} a_n z^n = a_0 + a_1 \bar{\alpha} k(\alpha z) \tag{3}$$

for some α, $|\alpha| = 1$. In other words he proved his conjecture for the case $n = 2$ and showed that equality holds only if the function is closely related to the Koebe function. He conjectured that if $|a_n| = n|a_1|$ for any $n \geq 2$, then the function has to be as in (3).

A lot of work was done, much of it related to the Bieberbach conjecture in some way, where the Koebe function (or really the functions (3)) was the extremal case. The Bieberbach conjecture itself was making slow progress (it was proved for most values of n up to $n = 7$) until Louis de Branges proved the whole conjecture in 1985 (Acta Math. volume 154, pages 137-152).

It does not seem that knowing the Bieberbach conjecture has any important consequence, but the result attracted a lot of attention because the conjecture remained unsolved for so long. The original result of Bieberbach for $n = 2$ can be used to prove the following elegant result (which is quite useful).

7.23 Theorem (Koebe $1/4$-theorem). *If f is an injective analytic function on the unit disc, then the range of f contains the disc*

$$D\left(f(0), \frac{1}{4}|f'(0)|\right).$$

Unless f has the form (3), the range of f will contain a disc about $f(0)$ of radius $r > (1/4)|f'(0)|$.

I omit the proof of this as I have also omitted the (reasonably elementary) material necessary to show $|a_2| \leq 2|a_1|$.

One other theorem that was mentioned earlier, but we have not covered, is Runges theorem.

Theorem 7.24 (Runges theorem). *Let $G \subset \mathbb{C}$ be open and $E \subset \hat{\mathbb{C}} \setminus G$ a subset with the property that the closure of E intersects every connected component of $\hat{\mathbb{C}} \setminus G$. Then the set of rational functions with poles only at points of E (here we include $z = \infty$ as a possible pole) is dense in $H(G)$.*

Again we omit the full proof, although Lemma 7.21 helps quite a lot in the proof. Here is a sketch. If $K \subset G$ is compact, and $f \in H(G)$, then Lemma 7.21 expresses $f(a)$ for $a \in K$ as an integral (hence approximable by a Riemann sum) of functions

$$g_z(a) = \frac{1}{2\pi i} \frac{f(z)}{z - a}$$

with poles at points $z \in G \setminus K$. The essential part of the proof (not a very short part) consists in showing the functions $g_z \colon K \to \mathbb{C}$ can be approximated uniformly on K by rational functions with poles in E.

The conclusion of this argument is that for $f \in H(G)$ and $K \subset G$ compact, it is possible to find a rational function $r(z)$ with poles only in E so that

$$\sup_{z\in K} |f(z) - r(z)|$$

is small. Finally, we do this for an exhaustive sequence $K_1 \subset K_2 \subset K_2 \subset \cdots$ of compact subsets of G to produce a sequence of rational functions $r_n(z)$, each with poles only on E so that

$$\sup_{z\in K_n} |f(z) - r_n(z)| < \frac{1}{n}.$$

It follows then that $r_n \to f$ in $H(G)$.

Corollary 7.25. *Let $G \subset \mathbb{C}$ be open and assume $\hat{\mathbb{C}} \setminus G$ is connected. The the polynomial functions are dense in $H(G)$.*

Proof. Apply Runges theorem with $E = \{\infty\}$ and note that rational functions with poles only at ∞ are polynomials. □

Remark 7.26. One reason to advertise the existence of Runges theorem in advance was to state that it can be used to exhibit pointwise convergent sequences of analytic functions with discontinuous limits.

Let $G_n = \{z \in \mathbb{C} : \Re z \neq 1/(2n)\}$ (the complex plane minus a vertical line) and let $K_n \subset G_n$ be the compact set

$$K_n = L_n \cup R_n, \quad L_n = \{z \in \mathbb{C} : |z| \leq n, \Re z \leq 0\}, R_n = \{z \in \mathbb{C} : |z| \leq n, \Re z \geq 1/n\}$$

Let $f_n \colon G_n \to \mathbb{C}$ be the analytic function

$$f_n(z) = \begin{cases} 0 & \Re z < 1/(2n) \\ z & \Re z > 1/(2n) \end{cases}$$

Then $\hat{\mathbb{C}} \setminus G_n$ is connected and so the Corollary to Runges theorem says we can find a polynomial $p_n(z)$ so that

$$\sup_{z\in K_n} |f(z) - p_n(z)| < \frac{1}{n}$$

It follows that

$$\lim_{n\to\infty} p_n(z) = \begin{cases} 0 & z \in \bigcup L_n = \{z \in \mathbb{C} : \Re z \leq 0\} \\ z & z \in \bigcup R_n = \{z \in \mathbb{C} : \Re z > 0\} \end{cases}$$

So the pointwise limit fails to be continuous along the imaginary axis (except at the origin and it fails to be analytic along the whole imaginary axis).