GEOMETRIC FUNCTION THEORY

WS 1995/96

J.B. Cooper

I. Basic facts on holomorphic functions—the Riemann mapping theorem

II. Cross ratios, Möbius transformations, Circle geometry

III. Domains as Riemann surfaces, the Poincaré metric, the Picard theorems

I. Facts on analytic functions

*Notation.*We use both letters z and P to represent points in the plane (z denotes the corresponding complex number). As a general rule we write z when we are emphasising the algebraic properties of the complex plane, P when we are emphasising geometrical aspects. Sometimes we will write z_P to denote the complex number corresponding to the point P in the plane.
$U(P, r)$ denotes the open disc $\{z \in \mathbf{C} : |z - z_P| < r\}$. ∂U denotes its boundary. D is the unit disc i.e. the case $P = 0$, $r = 1$.
Ω or U denotes a generic domain in $\mathbf{C}$ (i.e. an open, connected subset).

Let f be a function from U into $\mathbf{C}$. We shall identify without explicit comment f with the vector field (u, v) or pair of functions of two real variables where

$$f(x + iy) = u(x, y) + iv(x, y).$$

Example. For $f(z) = z^3$ resp. $f(z) = e^z$, we have $u(x, y) = x^3 - 3xy^2 \quad v(x, y) = 3x^2y - y^3$ resp. $u(x, y) = e^x \cos y \quad v(x, y) = e^x \sin y$.

We define f to be **holomorphic** if it is C^1 and satisfies the Cauchy-Riemann equation

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

(In fact it suffices that f be differentiable and satisfy the above equations. It is usually proved in an elementary course on function theory that the continuity of the partial derivatives then follows).

We remark that the above condition signifies that the Jacobi matrix of f has the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ of a similarity, more precisely a rotation followed by a dilation.

*Notation.*We employ the differential operators

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

(Thus $f = u + iv$ satisfies the Cauchy-Riemann equations if and only if $\frac{\partial f}{\partial \bar{z}} = 0$).

Examples.

$$\frac{\partial}{\partial z} z = 1, \quad \frac{\partial}{\partial z} \bar{z} = 0, \quad \frac{\partial}{\partial \bar{z}} z = 0, \quad \frac{\partial}{\partial \bar{z}} \bar{z} = 1.$$

If f is complex differentiable

$$f'(z) = \frac{\partial f}{\partial z} = \frac{\partial f}{\partial x} = -i\frac{\partial f}{\partial y}.$$

Also $\Delta f = 4\dfrac{\partial}{\partial z}\dfrac{\partial}{\partial \bar{z}} = 4\dfrac{\partial}{\partial \bar{z}}\dfrac{\partial}{\partial z}$.

The (real) chain rule takes the following form: if f, g are C^1, then

$$\frac{\partial}{\partial z}(f \circ g) = \frac{\partial f}{\partial z}(g(z))\frac{\partial g}{\partial z}(z) + \frac{\partial f}{\partial \bar{z}}(g(z))\frac{\partial \bar{g}}{\partial z}(z)$$
$$\frac{\partial}{\partial \bar{z}}(f \circ g) = \frac{\partial f}{\partial z}(g(z))\frac{\partial g}{\partial \bar{z}}(z) + \frac{\partial f}{\partial \bar{z}}(g(z))\frac{\partial \bar{g}}{\partial \bar{z}}(z).$$

Exercise. If f is holomorphic, then $\Delta(\ln |f|^2) = 0$. For

$$(\Delta(\ln |f|^2) = \Delta \ln f + \Delta \ln \bar{f} = 4\frac{\partial}{\partial z}\frac{\partial}{\partial \bar{z}} \ln f + 4\frac{\partial}{\partial \bar{z}}\frac{\partial}{\partial z} \ln \bar{f} = 0.$$

Exercise. If ϕ is harmonic and f is analytic, then $\phi \circ f$ is harmonic. (Three suggestions for a proof:

1) use the real version of the chain rule and the Cauchy Riemann equations.

2) use the fact that being harmonic is a local property and that a function is harmonic if and only if it is locally the real part of an analytic function.

3) use the complex chain rule).

Definition. If γ is a (piecewise) C^1 curve in U then we define

$$\int_\gamma f(z)\, dz$$

in the usual way. Using the above definition and the theorem of Stokes, we obtain immediately:

Proposition. *If f is holomorphic on U and γ is the boundary of a two-dimensional cube in U (cf. Analysis I), then $\int_\gamma f(z)\, dz = 0$.*

From this we deduce in the usual way

Corollary. *If f and γ are as above, then*

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

whenever the winding number of γ with respect to z is 1 (i.e. $1 = \frac{1}{2\pi i}\int_\gamma \frac{1}{\zeta - z}\, d\zeta$).

From this result one deduces the existence of the Taylor series and the Laurent series as in an elementary course on complex variables.

These allow us to classify the isolated singularities of a function as follows:

Definition. Let $F : U \to \mathbf{C}$ be a holomorphic function where U is a punctured neighbourhood of z_0. We suppose that F has Laurent series

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

Then there are three possibilities:

(1) a **removable singularity**. This is the case if $a_n = 0$ whenever n is negative;
(2) a **pole**. This is the case when at least one a_n (n negative) is non-zero and only finitely many are non-zero;
(3) an **essential singularity**. The remaining case (i.e. infinitely many a_n (n negative) are non-vanishing).

The growth behaviour of F near z_0 can be used to distinguish between the three cases. If F is bounded near z_0 then the latter is a removable singularity. If $|F|$ tends to ∞ as $z \to z_0$, then it is a pole. If it is unbounded but does not tend to ∞, then we have an essential singularity.

Further consequences of the Cauchy integral formula are

Proposition. *If f is a holomorphic, non-zero function on U, then its zero set $Z(f)$ is discrete.*

Liouville's theorem. *A bounded entire function is constant.*

The principle of the argument. *Let F be analytic on U, γ a simple closed curve which is the boundary of a two dimensional cube in U and suppose that F has k zeroes in the interior of γ and none on the curve. Then*

$$k = \frac{1}{2\pi i} \int_\gamma \frac{F'(\zeta)}{F(\zeta} \, d\zeta.$$

This implies the following result of Hurwitz:

Proposition. *Suppose that the sequence (F_n) of holomorphic functions converges almost uniformly (i.e. uniformly on compact subsets) in U to F and that none of the F_j has zeroes. Then either F has no zeroes or it is the constant function 0.*

Another consequence of the Cauchy integral formula is:

Maximum principle. *If $F : U \to \mathbf{C}$ is such that there is a $P \in U$ so that $|F(P)| \geq |F(z)|$ for each $z \in U \setminus \{P\}$, then F is constant.*

From this we can deduce

Schwarz' Lemma. *Let $F : D \to D$ be holomorphic with $F(0) = 0$. Then $|F(z)| \leq |z|$ and $|F'(0)| \leq 1$. If we have equality in the first equation for some non-zero z or in the second, then F has the form $z \mapsto \alpha z$ für $\alpha \in \partial U$.*

Proof. We apply the maximum principle to the function

$$G \mapsto \begin{cases} F(z)/z & (z \neq 0) \\ F'(0) & (z = 0). \end{cases}$$

Below we shall discuss general Möbius transformations in some detail. For the present moment we will consider the special function

$$\phi_a : z \mapsto \frac{z-a}{1-\bar{a}z},$$

where $|a| < 1$. This maps D conformally onto D and a to 0. This follows easily from general facts (see below) and the following calculation which shows that ∂D is mapped onto itself. Suppose that $|z| = 1$. Then

$$|\phi_a(z)| = \left|\frac{z-a}{1-\bar{a}z}\right| = \left|\frac{1}{\bar{z}}\frac{z-a}{1-\bar{a}z}\right| = \left|\frac{z-a}{\bar{z}-\bar{a}}\right| = 1.$$

Also it is easily calculated that the inverse of ϕ_a is of the same form, in fact it is ϕ_{-a}.

We also use the notation ρ_τ ($\tau \in \mathbf{R}$) to denote the rotation $z \mapsto e^{i\tau}z$ which is also clearly a conformal mapping on D. On the other hand we have:

Proposition. *Suppose that $F : D \to D$ is conformal. Then it has the form*

$$F : z \mapsto \phi_a \circ \rho_\tau(z)$$

for some a and τ.

Proof. We put $b = F(0)$. Then $G = \phi_b \circ F$ is as in the Lemma of Schwarz. Hence $|G'(0)| \leq 1$. But its inverse G^{-1} satisfies the same conditions and so, by the chain rule, $|G'(0)| \geq 1$. Hence it is equal to one and so G is a rotation from which the result follows.

Using the above mappings we can obtain the following version of the Schwarz-Lemma which is invariant under such conformal mappings.

Proposition—Schwarz, Pick. *Let $F : D \to D$ be holomorphic with $F(z_1) = w_1$, $F(z_2) = w_2$, $F(z) = w$. Then*

$$\left|\frac{w_1 - w_2}{1 - \bar{w}_1 w_2}\right| \leq \left|\frac{z_1 - z_2}{1 - \bar{z}_1 z_2}\right|,$$

and

$$|F'(z)| \leq \frac{1 - |w|^2}{1 - |z|^2}.$$

If there is equality in the first case for a pair of distinct z's or in the second for some z then F is conformal (and so of the above form).

Proof. We use the Möbius trausformations

$$\phi(z) = \frac{z + z_1}{1 + \bar{z}_1 z}, \quad \psi(z) = \frac{z - w_1}{1 - \bar{w}_1 z}$$

which take 0 to z_1 resp. w_1 to 0. Thus $\psi \circ F \circ \phi$ satisfes the conditions of the Lemma of Schwarz. Hence we have the inequality: $|\psi \circ F \circ \phi(z)| \leq |z|$. If we choose $z = \phi^{-1}(z_2)$, then we obtain the first inequality after a routine computation. Similarly, the inequality $|\phi'(w_1) F'(z_1) \phi'(0)| \leq 1$ which we obtain by using the chain rule produces the second result.

Normal families

We consider now the space $H(U)$ of holomorphic functions on a domain U. We regard this as a Fréchet space (complete metrisable locally convex space) with the countable family $\{p_n\}$ of seminorms where $p_n(f)$ is the supremum of the absolute value of f on K_n and (K_n) is a (countable) basis for the compact subsets of U. We shall only require the following facts about the corresponding topology—it is metrisable and convergence means uniform convergence on compact subsets of U.

Then we have the following characterisation of the relatively compact subsets of $H(U)$ which follows from the theorem of Ascoli and the Cauchy integral theorem:

Proposition. *A subset A of $C(U)$ is relatively compact if and only if it is uniformly bounded and equicontinuous on compacta. $A \subset H(U)$ is relatively compact if and only if it is uniformly bounded on compacta.*

In particular we have Montel's theorem—if A is uniformly bounded, then it is relatively compact.

Relatively compact subsets of $H(U)$ are traditionally called **normal families**. Using a standard reformulation of the condition of relative compactness in metric spaces we can recover the classical definition: A is normal if and only if each sequence (f_n) in A has a subsequence which converges almost uniformly. (The limit function is then automatically in $H(U)$).

One of the main applications of normal families is to the proof of the Riemann mapping theorem:

Proposition. *Let the domain U be a proper subset of* $\mathbf{C}$ *and simply connected (equivalently, homeomorphic to D). Then U is conformally equivalent to D.*

Proof. We begin the proof with the following two reductions.

Firstly we note that if the closure of U is a proper subset of $\mathbf{C}$ (i.e. U is not dense in $\mathbf{C}$), then we can find a disc in the exterior of U. Inversion in this disc is a conformal mapping of U onto a bounded set.

On the other hand since U is simply connected, we can define a branch of the logarithm on U. (For convenience, we assume that $0 \notin U$ and $1 \in U$—we can clearly arrange this by simple transformations. Then we define

$$\ln z = \int_\gamma \frac{1}{z}\, dz$$

where γ is a path from 1 to z in U. By the condition of simply connectedness, this is independent of the path and so is an analytic function). By the usual argument the image of U under this function (which is then conformally equivalent to U) satisfies the conditions of the previous paragraph.

Hence combining these two arguments we can reduce to the case where U is a bounded region.

We now introduce for a fixed P in U the family

$$\mathcal{F} = \{f : U \to D : f \text{ is } 1-1, \text{ holomorphic and } f(P) = 0\}.$$

This is non-empty (since we can find such a function on any bounded disc). By Montel's theorem, $\mathcal{F}$ is normal.

Now we set $M = \sup\{|f'(P)| : f \in \mathcal{F}\}$ and find a sequence (f_n) with $|{f_n}'(P)| \geq M - \dfrac{1}{n}$.

Since $\mathcal{F}$ is normal we can, by going over to a subsequence, assume that f_n converges in $H(U)$, say to f. We claim that $f \in \mathcal{F}$ and (as is obvious) $|f'(P)| = M$ i.e. the above supremum is attained. Firstly we note that f is non-constant (why?). Secondly, f takes its values in $\overline{D}$ as the limit of the f_n. But it then follows from the maximum modulus theorem that in fact its image is a subset of D.

It can easily be deduced using Hurwitz' theorem that f is also $1-1$ (Exercise).

Hence it remains to show that f is onto. This is where the concrete analysis comes in. We shall show that if f is not onto, then we can find a g in $\mathcal{F}$ with too large a value for $|g'(P)|$. Assume therefore that f is not onto. Suppose that the value β is omitted. We now compose with the Möbius transformation ϕ_β to get a function F which omits the value 0. Since this function is defined on a simply connected domain, we can again as above define a function $\ln F$ (by putting

$$\ln F(z) = \int \frac{F'(\zeta)}{F(\zeta}\, d\zeta,$$

the integral being taken along any path in U from P to z). We can then define F^α for any α by putting

$$F^\alpha(z) = \exp(\alpha \ln F(z)).$$

Putting this together we define an analytic function μ where

$$\mu(z) = (\phi_\beta \circ f(z))^{1/2} .$$

Then we put

$$\nu(z) = \frac{|\mu'(P)|}{\mu'(P)} \phi_\tau \circ \mu(z)$$

where $\tau = \mu(P)$. Then $\nu \in \mathcal{F}$ and one can compute that

$$\nu'(P)| = \frac{1 + |\beta|}{2|\beta|^{1/2}} M$$

and this is strictly greater then M. This contradiction proves the result.

We close this section by stating the two further main results of this course. They will be proved later.

Picard's little theorem. *If F is a non-constant entire function, then F misses at most one value. In other words if there are two values w_1 and w_2 which are not in the range of an entire function, then the latter is constant.*

Picard's great theorem. *If F is holomorphic in a punctured neighbourhood of P and has an essential singularity there then for each smaller punctured neighbourhood, F misses at most one complex value there.*

Appendix—The theorem of Ascoli

The classical theorem of Ascoli characterises compact subsets of $C(K)$ where K is compact:

Proposition. *A subset A of $C(K)$ is relatively (norm)-compact if and only if it is uniformly and equicontinuous.*

The proper setting for generalisations is the following situation: K is a compact space and (M, d) is a complete metric space. Then the space $C(K; M)$ of continuous (and so uniformly continuous) functions from K into M has a natural metric d_∞ where $d_\infty(f, g) = \sup_{x \in K} d(f(x), g(x))$.

A family $\mathcal{F}$ in $C(K; M)$ is relatively compact (for this metric) if and only if it is bounded and equicontinuous and for each $x \in K$, $\mathcal{F}_x = \{f(x) : f \in \mathcal{F}\}$ is relatively compact. In particular, if M itself is compact, then this reduces to the fact that $\mathcal{F}$ is equicontinuous.

From this we can easily deduce the following (using the diagonal method): Let U be locally compact and σ-compact and consider $C(U; M)$. This is a metric space (with almost uniform convergence as the corresponding concept of convergence) and a subset is equicontinuous if and only if it uniformly bounded and equicontinuous on compacta and satisfies the third condition above. If M is compact, only the equicontinuity is necessary since the other two hold automatically. In the final section we shall be particularly interested in the case where M is the sphere (in three space).

We now give some exercises which are useful in deriving Montel's theorem from the above versions of Ascoli's theorem.

Exercise. Show that $\mathcal{F} \subset H(U)$ is normal if and only if it is normal at each point. ($\mathcal{F}$ is normal at z if there is a neighbourhood of z on which the restrictions of the functions of $\mathcal{F}$ are normal). (Use the diagonal method).

Exercise. Show that if $\mathcal{F}$ is a subset of $H(U)$, then the following conditions are equivalent.

1) $\mathcal{F}$ is locally bounded i.e. for each K compact in U $\mathcal{F}$ is uniformly bounded on K;

2) each $z \in U$ has a neighbourhood on which $\mathcal{F}$ is uniformly bounded.

Exercise. Show that if $\mathcal{F}$ is a uniformly bounded subset of $H(U)$, then so is $\{f' : f \in \mathcal{F}\}$. (Use the previous exercise and the Cauchy integral formula for the derivative).

We remark that from the above it follows that if A is a locally bounded subset of $H(U)$ then it is equicontinuous (since the derivates are locally bounded and this implies that the restrictions of A to compact subsets are uniformly Lipshitz continuous).

We close with a famous result of classical function theory which can easily proved using these ideas.

Exercise. Prove the theorem of Vitali-Porter: Let (f_n) be a locally bounded sequence in $H(U)$, $f \in H(U)$ so that $\lim f_n(z)$ exists for each z in a subset A of U which has an accumulation point in U. Show that there is then an $f \in H(U)$ so that $f_n \to f$ almost uniformly. (Compare the (compact) topology of almost uniform convergence with the Hausdorff topology of pointwise convergence on A).

II. Cross ratios, Möbius transformations, Circle geometry

The complex form of the equation of a circle

We can write the equation of the circle with centre z_0 and radius r as

$$(z - z_0)(\bar{z} - \bar{z}_0) = r^2.$$

If we set

$$S(z) = z\bar{z} - z_0\bar{z} - z\bar{z}_0 + z_0\bar{z}_0 - r^2$$

this has the form $S(z) = 0$. The function S is called the **power** of z with respect to the circle.

It is convenient to consider a more general equation of the form

$$Az\bar{z} + B\bar{z} + \bar{B}z + D = 0 \quad (*)$$

with A and D real. If $A = 0$ this is a line and if $A \neq 0$ we can divide out to get the equation

$$z\bar{z} + \frac{B}{A}\bar{z} + \frac{\bar{B}}{A}z + \frac{D}{A} = 0$$

which is of the above form where $z_0 = -\frac{B}{A}$, $z_0\bar{z}_0 - r^2 = \frac{D}{A}$. Hence the equation represents a real, non-degenerate circle if and only if $A \neq 0$ and $AD - |B|^2 < 0$.

Exercise. Show that if C and C_1 are two (non concentric) circles, then the locus of those points with the same power with respect to the two circles is a straight line (called the **radical axis** of the circles). Show that if we have three circles then the corresponding three radical axes are concurrent or parallel.

It will be useful to write $(*)$ in matrix form as

$$\begin{bmatrix} 1 & \bar{z} \end{bmatrix} \begin{bmatrix} A & B \\ \bar{B} & D \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} = 0$$

where the 2×2 complex matrix $\mathcal{C} = \begin{bmatrix} A & B \\ \bar{B} & D \end{bmatrix}$ is hermitian. Two such matrices $\mathcal{C}$ and $\mathcal{C}'$ define the same circle if and only if one is a real multiple of the other.

Geometric interpretation of the power of a point. We consider the power of a point P with respect to the above circle. We assume without loss of generality that P is the origin and consider the intersection of the line $t \mapsto t\omega$ where ω is a complex number with $|\omega| = 1$ with the circle. This leads to the quadratic equation

$$At^2|\omega|^2 + Bt\omega + \bar{B}t\bar{\omega} + D = 0$$

This has in general two roots t_1 and t_2 with the property that $t_1.t_2 = \frac{D}{A}$ (which is independent of the direction). This latter quantity is just the power of P with respect to the circle. This is positive if P lies outside of the circle, zero if it lies on the circle and negative if it lies inside of the circle. In the latter case the power is the square of the length of the tangent from P to the circle. The fact that t_1t_2 is independent of direction translates into a well-known result of circle geometry (even for the case where P lies inside of the circle).

Exercise. What is the locus of the set of points whose power with respect to a given circle is constant, resp. the points so that the difference of its powers to two given circles is constant resp. so that the ratio of these powers is constant?

Möbius transformations

The typical Möbius transformation has the form

$$T = T_A : z \mapsto \frac{az+b}{cz+d}$$

where A is the invertible complex matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. (Note that two matrices induce the same Möbius transformation if and only if they are proportional). Its inverse is the Möbius transformation

$$w \mapsto \frac{dw-b}{-cw+a}$$

with matrix $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ which is (up to a constant) the inverse of A. A simple calculation shows that the composition $T_A \circ T_B$ of two Möbius transformations is the transformation T_{AB}.

Important special cases are those with matrices:
$\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$. This is the translation $z \mapsto z+t$;
$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ where $k>0$. This is the dilation $z \mapsto kz$;
$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ where $|k|=1$. This is the rotation $z \mapsto kz$;
$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Inversion $z \mapsto \frac{1}{z}$.

We remark that the last transformation is inversion in the unit circle (i.e. the mapping $z \mapsto \frac{1}{\bar{z}}$) followed by a reflection in the x-axis.

We can write the general Möbius transformation in the form

$$\frac{az+b}{cz+d} = \frac{bc+d}{c^2\left(z+\frac{d}{c}\right)} + \frac{a}{c} \quad (c \neq 0)$$

resp.

$$\frac{a}{d}(z+b) \quad (c=0)$$

from which one sees that the mappings of the above special type (i.e. translation, dilation, rotation, inversion) generate the group of Möbius transformations.

The cross ratio

If z_1, z_2, z_3, z_4 are distinct points of the plane we define:

$$(z_1, z_2, z_3, z_4) = \frac{z_1 - z_2}{z_1 - z_4} \div \frac{z_3 - z_2}{z_3 - z_4}$$

This is the image of z_1 under the (uniquely determined) Möbius transformation

$$z \mapsto \frac{z - z_2}{z - z_4} \div \frac{z_3 - z_2}{z_3 - z_4}$$

which maps $z_2 \mapsto 0$, $z_3 \mapsto 1$, $z_4 \mapsto \infty$.

Proposition. *If T is a Möbius transformation, then*

$$(Tz_1, Tz_2; Tz_3, Tz_4) = (z_1, z_2; z_3, z_4).$$

Proof. This can be proved simply by substituting the value of $T(z_1)$ etc. and carrying out the corresponding calculation.

One can simplify the proof by using the above result which shows that it suffices to consider transformations of the four special forms described above. Of these only the case of inversion needs to be calculated and this is very simple.

Proposition. *Let T be a Möbius transformation. Then T maps circles or straight lines onto circles (or straight lines).*

Proof. We use the remark of the previous proof to reduce to the only non-trivial case, the transformation $z \mapsto \frac{1}{z}$. Then if we substitute $w = \frac{1}{z}$ in the equation

$$Az\bar{z} + B\bar{z} + \bar{B}z + D = 0$$

then we get

$$Dw\bar{w} + Bw + \bar{B}\bar{w} + A = 0$$

which is the equation of a circle (if $D \neq 0$) or a line ($D = 0$). (We remark that the two cases correspond to whether the origin lies on the circle or not).

In fact a routine calculation show that if C is the circle with matrix

$$\mathcal{C} = \begin{bmatrix} A & B \\ \bar{B} & D \end{bmatrix}$$

and T is the Möbius transformation with matrix $\mathcal{A}$, then the pre-image of C has matrix $\mathcal{A}^*\mathcal{C}\mathcal{A}$.

Proposition. *The points z_1, z_2, z_3, z_4 lies on a circle (or line) if and only if their cross-ratio is real.*

Proof. We fix z_2, z_3, z_4 and note that for $z \in \mathbf{C}$, then the cross-ratio of z, z_2, z_3, z_4 is real if and only if

$$(z, z_2; z_3, z_4) = \overline{(z, z_2; z_3, z_4)}$$

and this simplifies to the equation of a circle or line.

Exercise. Give an alternative proof of this by showing that there is a Möbius transformation T which takes z_2, z_3 and z_4 to points w_2, w_3, w_4 on the real line. Then $w = T(z)$ lies on the real line if and only if its cross-ratio with the other w's is real. Then apply the above proposition.

Exercise. Suppose that we have four circles S_1, S_2, S_3, S_4 with $S_1 \cap S_2 = \{z_1, w_1\}$, $S_2 \cap S_3 = \{z_2, w_2\}$, $S_3 \cap S_4 = \{z_3, w_3\}$, $S_4 \cap S_1 = \{z_4, w_4\}$. Then if z_1, z_2, z_3, z_4 lie on a circle (or line), so do w_1, w_2, w_3, w_4 (the cross-ratios

$$(z_1, w_2; z_2, w_1), \quad (z_2, w_3; z_3, w_2), \quad (z_3, w_4; z_4, w_3), \quad (z_4, w_1; z_1, w_4)$$

are real. Multiplying gives that the product $(z_1, z_3; z_2, z_4)(w_1, w_3; w_2, w_4)$ is real).

We remark that using this one can prove the following fascinating chain of results due to W.K. Clifford (for the details see Yaglom).

Suppose that we have three straight lines in general position. Then they determine a circle (the circumcircle of the corresponding triangle).

Suppose now that we have four such lines. Then by omitting one line at a time we obtain four configurations as above and so four circles. The first result of Clifford states that these four circle have a common point. We call this the **central point** of the configuration.

We now consider a configuration of five lines. This determines five configurations of four lines and hence five central points. The next result of Clifford is that these five points lie on a circle.

This series of results can be continued indefinitely, the statements alternating between the fact that a given family of circles have a common point and a given family of points all lie on a circle. The reader is invited to give an exact formulation.

Similarity

We say that two Möbius transformations H and H_1 are **similar** if there is a transformation S so that $H_1 = S^{-1}HS$. For the corresponding matrices T and T_1, this means that there is a non-zero complex number q and an invertible matrix A so that $T_1 = qA^{-1}TA$. Under normal matrix similarity, the trace and the determinant are invariants. However the presence of the factor q above means that this will no longer be the case for this notion of similarity. However, it is easy to see that the quotient of the square of the trace by the determinant *is* invariant. In fact it is more convenient to use the expression

$$\sigma(T) = \frac{(\operatorname{Tr} T)^2}{\det T} - 4 = \frac{(a-d)^2 + 4bc}{ad - bc}.$$

(Thus if $T = \mathrm{Id}$, then $\sigma = 0$—this explains the presence of the term -4).

We shall now show that each Möbius transformation is similar to one of a particularly transparent form, a fact which is useful, for example, in examining the effect of iterating such transformations. In order to do this we consider the fixed point set of T i.e. the solutions of the equation

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

Case I: $c = 0$. Then T is entire and ∞ is a fixed point. Hence T has the form
$z \mapsto z + t$ (a translation which has no finite fixed points unless $t = 0$ i.e. T is the identity)
or
$z \mapsto kz + t$ $(k \neq 1, 0)$. This has a single finite fixed point $\dfrac{t}{1-k}$.
Case II. $c \neq 0$. Then the fixed point equation simplifies to the (genuine) quadratic equation

$$cz^2 + (d - a)z - b = 0$$

and so T has either 1 or 2 finite fixed points.
Hence there are three possibilities:
I. T has at least three fixed points. Then T is the identity and every point is a fixed point.
II. T has 2 fixed points $z_1 \neq z_2$ (one of which can be ∞);
III. T has a single fixed point z_1 (which can be at infinity).

We consider case II. Let S be a Möbius transformation with $0 \mapsto z_1$, $\infty \mapsto z_2$ and put $T_1 = S^{-1}TS$. Then T_1 has 0 and ∞ as fixed point and so $T_1 : z \mapsto kz \quad (k \neq 1)$.
Case III. Let S be a Möbius transformation with $\infty \mapsto z_1$. Then $T_1 = S^{-1}TS$ has ∞ as fixed point and so is entire. Hence $T_1 : z \mapsto z + t$ (otherwise it would have an additional fixed point).

Summarising we have:

Proposition. *Let T be a Möbius transformation. Then T is similar to a transformation of the form $z \mapsto kz$ or $z \mapsto z + t$.*

Exercise. Calculate σ for the two transformations above. Show that any two translations are similar and that $z \mapsto kz$ and $z \mapsto k_1 z$ are similar if and only if ($k = k_1$ or) $k = \dfrac{1}{k_1}$. Deduce that two Möbius transformations (both $\neq \mathrm{Id}$) are similar if and only if they have the same value of σ.

Hence we can say: T is

(1) **elliptic** if $|k| = 1$ but $k \neq 1$ (i.e. $-4 \leq \sigma \leq 0$);
(2) **proper hyperbolic** if $k > 0$ (i.e. $\sigma > 0$);
(3) **improper hyperbolic** if $k < 0$ (i.e. $\sigma < 4$);
(4) **loxodromic** if $|k| \neq 1$ but k is not real (i.e. σ is not real).

Exercise. Discuss the nature of the fixed points (i.e. attractive, repulsive or stable) of transformations of each of the above type.

For a general Möbius transformation we have four special points. Firstly the **pole** $z_\infty = -\dfrac{d}{c}$ i.e. the pre-image of ∞. Secondly $Z_\infty = T(\infty) = \dfrac{a}{c}$. We also have the two fixed points γ_1 and γ_2. If we consider the quadratic equation of which these are solutions we see that

$$\gamma_1 + \gamma_2 = \frac{-d + a}{c} = z_\infty + Z_\infty.$$

This means that

$$Z_\infty - \gamma_1 = \gamma_2 - z_\infty.$$

i.e. that the four points are the vertices of a parallelogram.

Conversely the parallelogram determines the Möbius transformation. For the Möbius transformation has the form $z \mapsto Z$ where

$$Z = \frac{z_\infty z - \gamma_1 \gamma_2}{Z - z_\infty}$$

(exercise) and so

$$Z - z_\infty = \frac{(\gamma_1 - Z_\infty)(\gamma_1 - z_\infty)}{z - z_\infty}$$

resp.

$$Z - \gamma_2 = \frac{(\gamma_1 - z_\infty)(z - \gamma_2)}{z - z_\infty}.$$

Hence

$$\frac{\gamma_1 - z_\infty}{z - z_\infty} = \frac{Z - Z_\infty}{\gamma_1 - Z_\infty} = \frac{Z - \gamma_2}{z - \gamma_2}$$

i.e. the triangles $\gamma_1 z_\infty z$, $Z Z_\infty \gamma_1$ and $Z \gamma_2 z$ are similar.

This fact can be used to give a ruler and compass construction of Z given z and the parallelogram.

The Poisson formula

We conclude by indicating briefly how Möbius transformations can be used to derive the Poisson formula for the solution of the Dirichlet problem on the circle. Let $w = \dfrac{z-a}{1-\bar{a}z}$. Then a routine calculation shows that

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

and

$$dw = \frac{1-a\bar{a}}{(1-\bar{a}z)^2}dz$$

so that

$$\frac{|dw|^2}{(1-|w|^2)^2} = \frac{|dz|^2}{(1-|z|^2)^2}$$

(we shall give a geometrical interpretation ot his equatio below). This suggests the relevance of the following invariant version of the Laplace operator

$$(1-|w|^2)^2\frac{\partial^2\Phi}{\partial w\partial\bar{w}} = (1-|z|^2)^2\frac{\partial^2\Phi}{\partial z\partial\bar{z}}$$

(N.B.

$$\frac{\partial^2\Phi}{\partial z\partial\bar{z}} = \frac{1}{4}\left(\frac{\partial^2\Phi}{\partial x^2}+\frac{\partial^2\Phi}{\partial y^2}\right)$$

and so the solutions of the corresponding homogeneous equation are as for the Laplace operator i.e. they are just the harmonic functions). Now consider the effect of the mapping $z \mapsto w$ on the circle. We write this as $e^{i\tau} \mapsto e^{i\psi}$. Then

$$e^{i\psi} = \frac{1-ae^{-i\tau}}{1-\bar{a}e^{i\tau}}e^{i\tau}.$$

Differentiating we get:

$$e^{i\psi}d\psi = \frac{1-a\bar{a}}{(1-\bar{a}e^{i\tau})^2}e^{i\tau}d\tau.$$

and dividing the latter two equations

$$d\psi = \frac{1-a\bar{a}}{|1-\bar{a}e^{i\tau}|^2}d\tau.$$

Hence if $a = \rho e^{i\theta} \quad (\rho < 1)$, then $d\psi = P(\rho, \theta-\tau)d\tau$ where

$$P(\rho,\theta-\tau) = \frac{1-|a|^2}{|1-\bar{a}e^{i\tau}|^2} = \frac{1-\rho^2}{1-2\rho\cos(\theta-\tau)+\rho^2}.$$

Then

$$\begin{aligned} P(\rho,\alpha) = \frac{1-\rho^2}{1-2\rho\cos\alpha+\rho^2} &= 1+2\Re\frac{e^{i\alpha}}{1-\rho e^{i\alpha}} \\ &= 1+2\sum\rho^n\cos n\alpha \end{aligned}$$

and so satisfies the Laplacian equation

$$\rho\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\rho}{\partial r}\right)+\frac{\partial^2 u}{\partial\theta^2}=0.$$

Now let u be continuous on $\overline{D}$ and harmonic on D. We shall now deduce Poisson's formula by means of the following computations:
Step 1.

$$\frac{1}{2\pi}\int_0^{2\pi}u(\rho e^{i\theta})d\theta=u(0).$$

For if we integrate the Laplace equation in polar form we get:

$$\begin{aligned}\rho\frac{\partial}{\partial\rho}&\left(\rho\frac{\partial}{\partial\rho}\frac{1}{2\pi}\int u(\rho e^{i\theta})\,d\theta\right)\\ &\qquad\qquad=-\frac{1}{2\pi}\int_0^{2\pi}\frac{\partial^2}{\partial\theta^2}u(\rho e^{i\theta})d\theta\\ &\qquad\qquad=-\frac{1}{2\pi}\frac{\partial}{\partial\theta}u(\rho e^{i\theta})\Big|_0^{2\pi}=0.\end{aligned}$$

Hence

$$\rho\frac{\partial}{\partial\rho}\frac{1}{2\pi}\int_0^{2\pi}u(\rho e^{i\theta})d\theta$$

is constant and letting $\rho\to 0$ we see that the constant is 0. Thus

$$\frac{1}{2\pi}\int_0^{2\pi}u(\rho e^{i\theta})d\theta=c.$$

Again letting $\rho\to 0$ shows that the constant c is $u(0)$.
Step 2. We consider the function v such that $v(z)=u(w)$ (so that $v(a)=u(0)$ and $v(e^{i\tau})=u(e^{i\psi})$). Then

$$v(a)=u(0)=\frac{1}{2\pi}\int_0^{2\pi}u(e^{i\psi})d\psi=\frac{1}{2\pi}\int_0^{2\pi}v(e^{i\tau})\frac{1-|a|^2}{|1-\bar{a}e^{i\tau}|^2}d\tau$$

and so if $a=a\rho^{i\theta}$, we have

$$v(\rho e^{i\theta})=\frac{1}{2\pi}\int_0^{2\pi}u(e^{i\tau})\frac{1-\rho^2}{1-2\rho\cos(\theta-\tau)+\rho^2}d\tau$$

which is Poisson's formula.

We conclude this section with some general exercises on Möbius transformations:

Exercise. Show that $T^2=\mathrm{Id}$ if and only if $a+d=0$. In this case, for any z the pair z and Tz are conjugate points. On the other hand show that if a Möbius transformation has two distinct conjugate points, then it is an involution (suppose that the conjugate points are 0 and ∞). (z_1 and z_2 are said to be **conjugate points** for T if $T(z_1)=z_2$ and $T(z_2)=z_1$. Determine the conjugate points for the transformations $z\mapsto z+t$ and $z\mapsto kz$).

Exercise. Show that every simple Möbius transformation is a product of two inversions and so that every transformation is a product of inversions. (By simple we mean one of the four basic types discussed at the beginning of the section).

Exercise. Show that inversion z^* of z in a circle C is characterised by the fact that it lies on each circle through z which is orthogonal to C.

Exercise. Show that a Möbius transformation $z \mapsto Z$ with fixed points γ_1 and γ_2 has the form

$$\frac{Z - \gamma_1}{Z - \gamma_2} = k \frac{z - \gamma_1}{z - \gamma_2}.$$

This implies $(Z_\infty, z_\infty; \gamma_1, \gamma_2) = k^2$. (This is the k of the characteristing theorem).

Exercise. Consider the Möbius transformation

$$w = k \frac{z - a}{z - b}.$$

This is the general form of a transformation with $a \mapsto 0$ and $b \mapsto \infty$. Consider the pre-image under this map of the polar coordinate lines in the plane. The lines through the origin go to circles through a and b and the circles with centre at the origin goe to the family of circles which is biorthogonal to the former. These are called the circles of Apollonius and have equations $\left| \frac{z - a}{z - b} \right| = c$. (i.e. they are the loci of the points whose distances to a and b are proportional).

Exercise. Show that the general form of the Möbius transformation $z \mapsto w$ with $a \mapsto a'$ and $b \mapsto b'$ is

$$\frac{w - a'}{w - b'} = k \frac{z - a}{z - b}.$$

Exercise. Describe those Möbius transformations which are periodic i.e. such that $T^n = \mathrm{Id}$ for some $n \in \mathbf{N}$.

Domains as Riemann manifolds

In this section we shall be motivated by the results and concepts from the course "Elementary Differential Geometry". We shall consider a domain Ω together with a positive C^2 function ρ defined on Ω. (Sometimes we shall allow ρ to have zeroes but the set of such zeroes will be discrete—we shall then only assume that ρ is C^2 on the complement of the zero set).

If $z \in \Omega$ and ξ is in $\mathbf{C} = \mathbf{R}^2$ (in reality in the tangent space to the manifold at z), we define

$$\|\xi\|_{\rho,z} = \rho(z)|\xi|.$$

Then we define the length $\ell_\rho(\gamma)$ of a path by the equation

$$\ell_\rho(\gamma) = \int_a^b \|\gamma'(s)\|_{\rho,\gamma(s)}\, ds.$$

($[a, b]$ is the parametrising interval). Note that we are not implying by the use of the symbols γ and s that γ has arc length parametrisation (as we did in DG).

If P and Q are points in U, then

$$\rho(P, Q) = \inf\{\ell_\rho(\gamma)\}$$

the infimum being taken over all paths in U from P to Q.

We remark that it follows from a general theorem on Riemann manifolds that the above infimum is attained (i.e. there is a path from P to Q with length $\rho(P, Q)$) for each pair P, Q if and only if U is complete under the metric ρ (it is relatively easy to see that the latter *is* a metric).

In the language of DG we are dealing with the Riemann manifold U with metric tensor (first fundamental form)

$$G = \begin{bmatrix} \rho^2(x, y) & 0 \\ 0 & \rho^2(x, y) \end{bmatrix}$$

(i.e. "$ds^2 = \rho^2(dx^2 + dy^2)$").
Then $g = \sqrt{\det G} = \rho^2$ and

$$G^{-1} = \begin{bmatrix} 1/\rho^2 & 0 \\ 0 & 1/\rho^2 \end{bmatrix}.$$

Our main examples will be
$\Omega = D$ and $\rho = \dfrac{1}{1 - |z|^2}$
resp.
$\Omega = H_+ = \{z \in \mathbf{C} : \Im z > 0\}$ with $\rho = \dfrac{1}{y^2}$.

In fact these spaces are essentially the same and are models for non-euclidean (hyperbolic) geometry.

Example. A routine calculation show that the length of the curve $\gamma(t) = t$ $(0 \leq t \leq 1-\epsilon)$ in the Poincaré metric $\dfrac{1}{1-|z|^2}$ is $\frac{1}{2}\ln\dfrac{2-\epsilon}{\epsilon}$. For the length is

$$\int_0^{1-\epsilon} \frac{1}{1-t^2}\,dt$$

which gives the above expression. If we put $R = 1-\epsilon$ then it takes the form

$$\frac{1}{2}\ln\frac{1+R}{1-R}.$$

This result can be reinterpreted as the fact that if $P = 0$ and $Q = (R, 0)$ where $0 < R < 1$, then $\rho(P,Q) = \frac{1}{2}\ln\left(\frac{1+R}{1-R}\right)$. (It is intuitively obvious and easy to demonstrate that the above path is the shortest route from P and Q. In fact if we consider a curve of the form $\gamma(t) = t + ib(t)$ with $b(0) = 0$, $b(1-\epsilon) = 1-\epsilon$, then it length is

$$\int_0^{1-\epsilon} \frac{(1+b'(t)^2)^{1/2}}{1-t^2-b(t)^2}\,dt$$

which is clearly larger than the above value).

We remark that the metrics on D and H_+ above define the usually topology on these subsets of $\mathbf{C}$. However, they are not *metrically* equivalent to the euclidean metrics there. In fact both of these metrics are complete (see below)

Definition. Suppose now that $f : \Omega_1 \to \Omega_2$ is a non-constant holomorphic function. If ρ is a metric on Ω_2, we define the **induced metric** $f^*\rho$ on Ω_1 by putting

$$f^*\rho(z) = \rho(f(z))\left|\frac{\partial f}{\partial z}\right|.$$

If f is a bijection and ρ_1 resp. ρ_2 are metrics on the above spaces, then f is **an isometry** if $f^*\rho_2 = \rho_1$. Then f^{-1} is also an isometry.
A simple calculation shows that then $\ell_{\rho_1}(\gamma) = \ell_{\rho_2}(f \circ \gamma)$ for each curve in Ω_1 and so that $\rho_1(P,Q) = \rho_2(f(P), f(Q))$ for P,Q in Ω_1.

For example one can show that if h is a conformal mapping of D, then h is an isometry for the Poincaré metric. For this it suffices to consider the two cases ρ_τ and ϕ_a discussed above:

Case 1) $w = e^{i\tau}z$. Then $h'(z) = e^{i\tau}$ and so its absolute value is 1. Hence

$$h^*\rho(z) = \rho(w) = \frac{1}{1-|w|^2} = \frac{1}{1-|z|^2} = \rho(z).$$

Case 2). $w = \dfrac{z-a}{1-\bar{a}z}$. Then

$$\begin{aligned} h^*\rho(z) &= \rho(w)\left|\frac{dw}{dz}\right| \\ &= \frac{1}{1-\frac{|a-a|^2}{|1-\bar{a}z|^2}}\frac{1-|a|^2}{|1-\bar{a}z|^2} \\ &= \frac{1-|a|^2}{1-\bar{a}z|^2-|z-a|^2} \\ &= \frac{1-|a|^2}{(1-|a|^2)(1-|z|^2)}. \end{aligned}$$

From this we can deduce the following formula:

Proposition. *If P and Q are points of D, then*

$$\rho(P,Q) = \frac{1}{2}\ln\left(\frac{1+\left|\frac{P-Q}{1-\bar{P}Q}\right|}{1-\left|\frac{P-Q}{1-\bar{P}Q}\right|}\right).$$

Proof. Firstly we have already seen that the formula is true for $P=0$ and $Q=(R,0)$. For the general case, we use the isometry ϕ_P (with the notation from above). Then

$$\rho(P,Q) = \rho(0,\phi_P(Q)) = \rho(0,|\phi_P(Q)|),$$

the last equality following from the fact that rotations about the origin are isometries. But

$$|\phi_P(Q)| = \left|\frac{P-Q}{1-\bar{P}Q}\right|.$$

The same calculation shows that the curve of shortest length from P to Q is

$$\gamma_{P,Q} : t \mapsto \frac{t\frac{Q-P}{1-Q\bar{P}}+P}{1+t\bar{P}\frac{Q-P}{1-Q\bar{P}}}.$$

This is the pre-image of the straight line from 0 to $\phi_P(Q)$ and so by the properties of Möbius transformations discussed above is (in general) an arc of a circle which cuts the unit circle at right angles.

Using the above formula one can check that $\rho(0,z) \le r$ if and only if $\dfrac{1}{2}\ln\left(\frac{1+|z|}{1-|z|}\right) < r$ i.e. if and only if $|z| \le \dfrac{e^{2r}-1}{e^{2r}-1}$. This allows us to show the equivalence of topologies mentioned above. For it follows that the discs with centre 0 form a basis there for both topologies. Further the above Möbius transformations are homemomorphisms for both topologies. From this it is easy to deduce that ρ induces the natural topology on the open ball. We can also see that (D, ρ) is complete. For let (z_n) be a ρ Cauchy sequence. Then it is bounded, say $\rho(z_n, 0) \leq M$ for some constant M. But then

$$|z_n| \leq \frac{e^{2M} - 1}{e^{2M} + 1}$$

and so the sequence lies in a compact subset of U. It follows easily from this that it converges (for both the natural and the ρ-topology).

Exercise. Calculate the Christoffel symbols and the geodetic equations for D.

We remark that it is not difficult to see that the above properties determine the Poincaré metric. In fact if $\bar{\rho}$ is a metric on D so that each conformal mapping of the disc is an isometry, then $\bar{\rho}$ is a multiple of the Poincaré metric. For suppose that $w = h(z) = \dfrac{z + z_0}{1 + \bar{z}_0 z}$. Then since

$$h^* \bar{\rho}(0) = \bar{\rho}(0)$$

we have

$$|h'(0)| \bar{\rho}(h(0)) = \bar{\rho}(0)$$

i.e.

$$\bar{\rho}(z_0) = \frac{1}{1 - |z_0|^2} \bar{\rho}(0) = \bar{\rho}(0) \rho(z_0)$$

which means that $\bar{\rho}$ is a multiple of the poincaré metric.

On the other hand, if $f : D \to D$ is an isometry for ρ then f is automatically holomorphic.

Proof. For by the usual reduction we can assume that $f(0) = 0$. Then as above the circle $C^R = \{z : |z| = R\}$ is mapped onto itself for each $0 < R < 1$. This means that for any P

$$\frac{|f(P) - f(0)|}{|P - 0|} = \frac{|f(P)|}{|P|}$$

and so f is conformal at 0 (in the sense that it preserves angles between curves). Once again by the homogeneity, this holds everywhere. But by a classical result (cf. Ahlfors) this implies that f is either holomorphic or anti-holomorphic. The latter case is impossible (since then $\left| \dfrac{\partial f}{\partial \bar{z}} \right| = 0$).

The Lemma of Schwarz-Pick can now be interpreted as follows: Suppose that f is a holomorphic function on the disc. Then f is a contraction i.e. $f^* \rho \leq \rho$ (and so $\ell_\rho(f \circ \gamma) \leq \ell_\rho(\gamma)$ and $\rho((f(P), f(Q)) \leq \rho(P, Q))$.

Using a refinement of the Banach fixed point theorem one can deduce that if f is a holomorphic mapping on the disc with relatively compact range, then it has a (unique) fixed point.

Proof. Under these conditions the function

$$g : z \mapsto f(z) + \epsilon(f(z) - f(0))$$

maps D into D for small enough ϵ. Then, by the above, g is a contraction in the weak sense and, since $f'(z) = \frac{1}{1+\epsilon} g'(z)$, f is a contraction in the sense of the Banach fixed point theorem.

It follows from the proof that the fixed point is obtained as the limit of the iterated sequence $z, f(z), f^2(z), f^3(z), \ldots$ for any z in the disc.

Curvature

Since the (Gaussian) curvature is an intrinsic quantity of a Riemann surface, it can be defined in terms of the metric tensor. In fact, the formula of the Theorema Egregium simplifies in this case to

$$-\frac{\Delta \ln \rho(z)}{\rho(z)^2}$$

and we denote this quantity by $\kappa_{U,\rho}(z)$ (or simply $\kappa(z)$). (For in the case where $F = 0$, the Theorema Egregium produces the following formula for the curvature:

$$4E^2G^2\kappa = E(E_2G_2 + G_1^2) + G(E_1G_1 + E_2^2) - 2EG(E_{22} + G_{11}).$$

A simple calculation shows that this yield the above expression when $E = \rho^2 = G$). It follows immediately from these remarks that this quantity is preserved by an isometry. (This can also be deduced directly by a simple computation).
It is easy to see that $\kappa = 0$ when ρ is the constant function 1 on $\mathbf{C}$, while $\kappa = -4$ for the Poincaré metric. On the other hand if $\rho = \dfrac{2}{1 + |z|^2}$, then $\kappa = 1$ (we shall see the geometrical reason for this shortly).

The curvature will play a crucial role in our versions of the Picard theorems. In order to see the connection note that the theorem of Liouville (in the form that each entire function with values in D is constant) can be regarded as a special case of Picard's little theorem. We shall show that the same result holds for entire functions with values in a domain U which allow a metric with certain curvature properties.

Proposition. *Let U be a domain with metric σ for which $\kappa \leq -4$. Then if $f : D \to U$ is holomorphic, $f^*\sigma \leq \rho$.*

Proof. We consider the smaller disc $U(0,r)$ (with $r<1$) and then let r go to 1. On this set we rescale the metric to $\rho_r = \dfrac{r}{r^2-|z|^2}$, for which we also have constant curvature $=-4$. Define the function $v = \dfrac{f^*\sigma}{\rho_r}$. This is continuous and non-negative, and converges to zero at the boundary. Hence $|v|$ attains its maximum M at a point $\tau \in U(0,r)$. We show that $M \leq 1$, from which our result follows. We can suppose that $f^*\sigma(\tau) > 0$. Then the curvature of $f^*\sigma$ is defined at τ (and is ≤ -4). Now since $\ln v$ has a maximum at τ, its Laplacian there is ≤ 0 (consider the Hessean matrix). Thus

$$\begin{aligned}
0 \geq \Delta \ln v(\tau) &= \Delta \ln f^*\sigma(\tau) - \Delta \ln \rho_r \\
&= -\kappa_{f^*\sigma}(\tau)(f^*\sigma(\tau))^2 + \kappa_{\rho_r}(\rho_r(\tau))^2 \\
&\geq 4(f^*\sigma(\tau))^2 - 4(\rho_r(\tau))^2
\end{aligned}$$

and so $v(\tau) \leq 1$ which implies that $M \leq 1$.

We remark that this can be regarded as a form of Schwarz' lemma (apply the above to an f which vanishes at 0).

Rescaling, we get:

Proposition. *Let U be a domain with metric σ for which $\kappa \leq -B < 0$. Then if ρ_A^α is the metric $\dfrac{2\alpha}{\sqrt{A}(\alpha^2-|z|^2)}$ on $U(0,\alpha)$, we have*

$$f^*\sigma(z) \leq \frac{\sqrt{A}}{\sqrt{B}}\rho_A^\alpha(z)$$

for each holomorphic mapping f from $U(0,\alpha) \to U$.

Proof. Exercise.

As an application we have

Proposition. *Let U be a domain with a metric σ so that $\kappa_\sigma \leq -B < 0$. Then each holomorphic function f from* **C** *into U is constant.*

Proof. We consider f as a mapping from $U(0,\alpha)$ with metric ρ_A^α for positive A. Then

$$f^*\sigma(z) \leq \frac{\sqrt{A}}{\sqrt{B}}\rho_A^\alpha(z)$$

for $|z| < \alpha$. Letting α go to infinity gives $f^*\sigma(z) \leq 0$ and so $f^*\sigma = 0$. But this can only happen if f' vanishes identically.

This result contains Liouville's theorem as a special case.

In order to motivate our proof of Picard's little theorem, consider how we can prove the following generalisation of Liouville. We show that every entire function with values in $\mathbf{C} \setminus [0,1]$ is constant. For it is a standard exercise in conformal mappings to map the above range space conformally into D and so we can deduce it from the usual version of Liouville. (Use successively the mappings $z \mapsto \dfrac{z}{z-1}$, $z \mapsto z^{1/2}$ and $z \mapsto \dfrac{z-1}{z+1}$).

Proposition. *Let U be a subset of $\mathbf{C}$ whose complement contains at least two distinct points. Then there is a metric ρ on U with $\kappa_\rho \leq -B < 0$.*

Proof. It is no real loss of generality to assume that the omitted points are 0 and 1. Then we define

$$\rho(z) = \left[\frac{\left(1 + |z|^{1/3}\right)^{1/3}}{|z|^{5/6}} \right] \left[\frac{\left(1 + |z-1|^{1/3}\right)^{1/3}}{|z-1|^{5/6}} \right].$$

ρ is a smooth positive function on U and a tedious calculation shows that the curvature is

$$\kappa(z) = -\frac{1}{18} \left[\frac{\left(|z-1|^{1/3}\right)^{5/3}}{(1+|z|^{1/3})^2(1+|z-1|^{1/3})} + \frac{|z|^{5/3}}{(1+|z|)^{1/3}(1+|z-1|^{2/3}} \right].$$

Then $\kappa < 0$ and it follows from the fact that $\lim_{z\to 0} \kappa(z) = -\frac{1}{36} = \lim_{z \to 1} \kappa(z)$ and $\lim_{z\to\infty} \kappa = -\infty$ that it is bounded away from zero.

This, combined with the above result, immediately implies Picard's little theorem.

We now turn to the great theorem. We shall be interested in functions which take their values in the extended plane $\hat{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$. We identify this with the sphere S^2 in $\mathbf{R}^3$ via stereographic projection. More precisely, if we denote by (α, β, γ) the coordinates of a point P on the sphere (which is distinct from the north pole $N = (0,0,1)$) then an elementary exercise in analytic geometry shows that the point $z = x + iy$ on the plane which is the intersection of the latter with the line through N and P is given by the formula $z = \dfrac{\alpha + i\beta}{1-\gamma}$. On the other hand, if we are given a point z in the plane, then the corresponding point $P = (\alpha, \beta, \gamma)$ on the unit sphere is given by the equations

$$\alpha = \frac{2\Re z}{1+|z|^2}, \quad \beta = \frac{2\Im z}{1+|z|^2}, \quad \gamma = \frac{-1+|z|^2}{1+|z|^2}.$$

The north pole is mapped onto ∞ by convention.

If we calculate the first fundamental form of the corresponding parametrisation

$$\phi(x,y) = \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{-1+x^2+y^2}{1+x^2+y^2} \right)$$

of the sphere, we get:

$$E = \frac{4}{(1+|z|^2)^2} = G, \quad F = 0.$$

This implies, amongst other facts, that stereographic projection is conformal and that the metric $\sigma(z) = \dfrac{2}{1+|z|^2}$ in the plane corresponds to the usual metric on S^2 as a surface in $\mathbf{R}^3$ (i.e. our correspondence is an isometry for these metrics).

We can compute that if P_1 and P_2 are the points on the sphere which correspond to the complex numbers z_1 and z_2 in the plane, then the chordal distance from P_1 to P_2 (i.e. the distance in $\mathbf{R}^3$, not the geodetic distance on the sphere) is

$$\frac{2|z_1 - z_2|}{\sqrt{1+|z_1|^2}\sqrt{1+|z_2|^2}}$$

resp. $\dfrac{2}{\sqrt{1+|z_1|^2}}$ if $z_2 = \infty$.

For if P_1 has coordinates $(\alpha_1, \beta_1, \gamma_1)$ and P_2 $(\alpha_2, \beta_2, \gamma_2)$, then this distance is

$$2 - 2(\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2)$$

and if we substitute the above expressions for the α's, β's and γ's, then this reduces to the above formula.

We denote this quantity by $\chi(z_1, z_2)$. An easy computation shows that

$$\chi\left(\frac{1}{z_1}, \frac{1}{z_2}\right) = \chi(z_1, z_2).$$

The length of a curve in the plane (using the metric σ) is then $\int_\gamma \dfrac{2|dz|}{1+|z|^2}$. $\sigma(z_1, z_2)$, the spherical metric, is the infimum of the lengths of the paths joining z_1 and z_2.

If f is a holomorphic mapping from U into the sphere, we define the quantity $f^\#(z)$ as the limit

$$\lim_{z' \to z} \frac{\chi(f(z), f(z'))}{|z - z'|}$$

and this can be computed to be $\dfrac{|f'(z)|}{1+|f(z)|^2}$. It follows immediately from this definition that this quantity coincides for f and $\dfrac{1}{f}$.

If we compare the above formula for $f^\#$ with the definition of the induced metric then we see that the length of a curve is given by the formula $\int_\gamma 2f^\#(z)\,|dz|$. (i.e. this is the length of the curve in the plane, using the above metric or the length of its image on the sphere using the spherical metric).

We now consider the space of continuous functions from U into $\hat{\mathbf{C}}$. In particular, each meromorphic function can be so regarded—we set the value of such a function at a pole to be ∞. (For the definition and elementary properties of meromorphic functions we refer to any standard text on function theory e.g. Ahlfors). We are now in the situation described in the appendix to the first section. Hence we can define:

Definition. Let (f_n) be a sequence of meromorphic functions. Then we say that (f_n) converges **normally** if it converges in the sense of the metric defined above to a meromorphic function or to the constant function ∞.

Under this definition we see that both the sequence (n) and $\left(\dfrac{n}{z}\right)$ converge normally.

A family $\mathcal{F}$ of meromorphic functions on U is then said to be **normal** if each sequence in $\mathcal{F}$ contains a subsequence which converges normally.

Using the version of Ascoli's theorem quoted in the above appendix, we see that a family of meromorphic functions is normal if and only if it is spherically equicontinuous on compacta (i.e. equicontinuous as a family of functions with values in the sphere under its geodetic metric).

Normal convergence can be characterised as follows:

Proposition. *We have $f_n \to f$ normally if and only if each z_0 has a neighbourhood on which either $f_n \to f$ or $\dfrac{1}{f_n} \to \dfrac{1}{f}$ uniformly.*

Marty's theorem. *Let $\mathcal{F}$ be a family of meromorphic functions on U. $\mathcal{F}$ is normal if and only if the family $\{f^{\#}\sigma : f \in \mathcal{F}\}$ is uniformly bounded on compact subsets of U (σ is the natural metric on the sphere).*

Remark. Using the definition of the induced metric this means that for each compact subset K of U there is a positive constant M so that for each $z \in K$ and each $f \in \mathcal{F}$, we have

$$\frac{|f'(z)|}{1 + |f(z)|^2} \leq M.$$

Proof. First suppose that the above condition is satisfied. We fix z_0 and consider each z in a suitable compact disc around z_0. Then if we choose the appropriate M for this disc we have, for any path from z_0 to z within this disc and any $f \in \mathcal{F}$,

$$\chi(f(z_0), f(z)) \leq \int_\gamma f^{\#}(\zeta)\,|d\zeta| \leq C|z - z_0|.$$

This implies that $\mathcal{F}$ is equicontinuous on K.

On other hand suppose that there is a compact subset of U, a sequence (z_n) in K and a sequence (f_n) in $\mathcal{F}$ with $f_n^{\#}(z_n) \to \infty$. By the normality we can suppose that f_n is convergent. Let z_0 be a limit point of (z_n). Then there is a disc around this point for which either $f_n \to f$ or $\frac{1}{f_n} \to \frac{1}{f}$ uniformly. In either case $f_n^{\#} \to f^{\#}$ uniformly on this disc (we use here the fact mentioned above that $\left(\frac{1}{f}\right)^{\#}$ coincides with $f^{\#}$). But this implies that $f_n^{\#}$ is uniformly bounded on the disc and this contradicts the assumptions.

Proposition. *Let U be a domain in $\mathbf{C}$ and P, Q and R three distinct points in the extended plane. Then if $\mathcal{F}$ is a family of holomorphic functions taking values in $\hat{\mathbf{C}} \setminus \{P, Q, R\}$, $\mathcal{F}$ is normal.*

Proof. We make the customary reduction to the case where the three exceptional points are 0, 1 and ∞. Then it suffices to show that the family is normal on any disc $U(z_0, \alpha)$. It is no loss of generality to suppose that $z_0 = 0$. We use the special metric constructed above on $\mathbf{C} \setminus \{0, 1\}$. By rescaling we can assume that it is ≤ -4. We denote this metric by μ. Then by the version of the Schwarz' Lemma (with $A = B = 4$) we have $f^*\mu(z) \leq \rho_\alpha^A(z)$ for z in $U(0, \alpha)$. We now compare μ with the spherical metric σ. One sees easily that $\frac{\sigma}{\mu}$ goes to zero near the critical points 0, 1 and ∞. Hence there is an $M > 0$ so that $\sigma \leq M.\mu$ and so

$$f^{\#} = f^*\sigma \leq M.f^*\mu \leq M.f^*\rho_\alpha^A$$

on $U(0, \alpha)$. By Marty's theorem, $\mathcal{F}$ is normal.

In particular, we can deduce as a corollary that a family of holomorphic functions on U which omit two (finite) values is normal.

The second Corollary of the above result is Picard's great theorem.

Proof. We suppose that we have a function on the punctured unit disc D' which omits the values 0 and 1. We prove that 0 is either a pole or a removable singularity. We consider the family of functions $f_n : z \mapsto f\left(\frac{z}{n}\right)$. This family also omits the values 0 and 1 and so is normal. Hence it has a subsequence which converges normally and so either to a holomorphic function on the punctured disc or to the constant function ∞. It is easy to see that f has in the first case a removable singularity and in the second case a pole at 0.

Literature

This course is based on sections of the following books:

Yaglom: Complex numbers in geometry;
Schwerdtfeger: Geometry of complex numbers;
Krantz: Funcion theory: a geometric approach.

A useful and beautiful introduction to function theory is the classical text of Ahlfors.