MATH 566 LECTURE NOTES 6: NORMAL FAMILIES AND THE THEOREMS OF PICARD

TSOGTGEREL GANTUMUR

1. Introduction

Suppose that we want to solve the equation $f(z) = \beta$ where f is a nonconstant entire function and $\beta \in \mathbb{C}$. We know that if f is a polynomial, this equation has a solution for all $\beta \in \mathbb{C}$. On the other hand, the equation $\exp z = \beta$ has no solution if $\beta = 0$. It turns out that in some sense this is the worst that can happen, in that $f(z) = \beta$ is solvable for all β from $\mathbb{C}$ with a possible exception of one point.

Theorem 1 (Picard's little theorem). *If $f \in \mathcal{O}(\mathbb{C})$ and $f(\mathbb{C}) \subset U$ where $\mathbb{C} \setminus U$ contains at least two points, then f is constant.*

This is a dramatic strengthening of Liouville's theorem, as Liouville's theorem can be thought of as the preceding theorem where the condition on U is replaced by the condition $U = \mathbb{D}$. We will actually prove an even stronger result known as Picard's big theorem.

Definition 2. Let S and M be Riemann surfaces, and let $f \in \mathcal{O}(S \setminus \{\alpha\}, M)$ with $\alpha \in S$. We say $\beta \in M$ is an *omitted value* of f at α, if there is an open neighbourhood $U \subset S$ of α such that $f(U) \not\ni \beta$.

So in order for β *not* to be an omitted value, there must exist a sequence $\{z_n\} \subset S$ with $z_n \to \alpha$ such that $f(z_n) = \beta$ for all n.

Theorem 3 (Picard's big theorem). *If $f \in \mathcal{O}(S \setminus \{\alpha\}, \hat{\mathbb{C}})$ omits 3 values from $\hat{\mathbb{C}}$ at α then f extends to $f \in \mathcal{O}(S, \hat{\mathbb{C}})$.*

In certain sense, this theorem strengthens Riemann's removable singularity theorem the same way the little theorem strengthens Liouville's theorem. Another way to formulate the big theorem is to say that a holomorphic function can omit at most one value at its essential singularity. We will give a proof at the end of these notes. For now, let us see how the little theorem follows from the big theorem. It suffices to treat the case where f is not a polynomial, since a polynomial of degree n takes every value in $\mathbb{C}$ exactly n times, counting multiplicities. It turns out that we can derive a result that is much stronger than the little theorem, namely that not only entire transcendental (i.e., non-polynomial) functions assume every value in $\mathbb{C}$ with at most one exception, but each of those values are taken infinitely many times.

Corollary 4. *If $f \in \mathcal{O}(\mathbb{C})$ is not a polynomial, then it assumes every value in $\mathbb{C}$ infinitely many times with at most one exception.*

Indeed, since f has an essential singularity at $\infty \in \hat{\mathbb{C}}$, it cannot be extended to a meromorphic function on $\hat{\mathbb{C}}$. By the big theorem, at $\infty \in \hat{\mathbb{C}}$, f can omit at most 2 values in $\hat{\mathbb{C}}$, or, since ∞ is already omitted, at most 1 value in $\mathbb{C}$. For all other values β in $\mathbb{C}$, we have a sequence $\{z_n\} \subset \mathbb{C}$ with $z_n \to \infty$ such that $f(z_n) = \beta$ for all n.

In order to prove the big theorem, we will develop in the following sections a very precise theory on compactness properties of families of meromorphic functions.

Date: December 7, 2010.

2. Marty's theorem

Obviously, a proof of Picard's big theorem requires us in one way or another to consider holomorphic functions having values in $\hat{\mathbb{C}}$ (otherwise known as meromorphic functions). When working with $\hat{\mathbb{C}}$ one has to deal with the point $\infty \in \hat{\mathbb{C}}$ from time to time, and as a result discussions become spotted with applications of special transformations such as $1/z$, potentially obscuring the clarity of the arguments and giving the false impression that ∞ was somehow special. This can be dealt with by introducing general charts on $\hat{\mathbb{C}}$, but then its complexity becomes the same as considering a general Riemann surface. So for the sake of both clarity and generality we will be working with a general Riemann surface M instead of the Riemann sphere $\hat{\mathbb{C}}$. We equip M with a (metric space) metric ρ that is compatible with its topology. This metric induces a way to measure distances between functions: for $f, g : U \to M$ defined on some set $U \subseteq \mathbb{C}$, we define

$$\rho_U(f, g) = \sup_{z \in U} \rho(f(z), g(z)).$$

With $\Omega \subseteq \mathbb{C}$ an open connected set, we say that the sequence $\{f_n\}$ of functions $f_n : \Omega \to M$ converges to $f : \Omega \to M$ *uniformly* on Ω if $\rho_\Omega(f_n, f) \to 0$ as $n \to \infty$, and that $\{f_n\}$ converges to f *locally uniformly* on Ω if for any compact $K \subset \Omega$, $\rho_K(f_n, f) \to 0$ as $n \to \infty$.

For each coordinate chart (U, φ) of M, we require the metric ρ to satisfy

$$\rho(z, w) = C_z|\varphi(z) - \varphi(w)| + o(|\varphi(z) - \varphi(w)|), \qquad z, w \in U,$$

where the constant $C_z > 0$ depends continuously on z. Note that C_z is in general different for different coordinate charts. If $f : \Omega \to M$ is a holomorphic function, then for $z \in \Omega$ we have

$$\begin{aligned}\rho(f(z), f(z+h)) &= C_{f(z)}|\varphi(f(z+h)) - \varphi(f(z))| + o(|\varphi(f(z+h)) - \varphi(f(z))|) \\ &= C_{f(z)}|(\varphi \circ f)'(z)||h| + o(|h|),\end{aligned}$$

and therefore the following limit exists

$$f^\sharp(z) = \lim_{h \to 0} \frac{\rho(f(z), f(z+h))}{|h|}.$$

We call $f^\sharp : \Omega \to \mathbb{R}_+$ the *metric derivative* of f. It is clear that $f^\sharp$ has the local expression

$$f^\sharp(z) = C_{f(z)}|(\varphi \circ f)'(z)|, \tag{1}$$

which in particular shows that $f^\sharp$ is a continuous function.

Example 5. For the Riemann sphere $\hat{\mathbb{C}}$, an example of ρ satisfying the above assumptions is the cordial distance between $\alpha, \beta \in \hat{\mathbb{C}}$ considered as points on $S^2 \subset \mathbb{R}^3$. Another is the geodesic distance between two points $\alpha, \beta \in S^2$ with respect to the round metric on S^2.

The following is a generalization of the Weierstrass convergence theorem.

Theorem 6. *Let $\{f_n\} \subset \mathcal{O}(\Omega, M)$ be a sequence that converges locally uniformly on Ω to a function $f : \Omega \to M$. Then $f \in \mathcal{O}(\Omega, M)$, and $\{f_n^\sharp\}$ converges locally uniformly.*

Proof. We first prove that f is continuous. From the triangle inequality for ρ, we have

$$\rho(f(z), f(w)) \leq \rho(f(z), f_n(z)) + \rho(f_n(z), f_n(w)) + \rho(f_n(w), f(w)),$$

for $z, w \in \Omega$ and $n \in \mathbb{N}$. For any compact set $K \subset \Omega$, one can choose n so large that $\rho_K(f_n, f)$ is arbitrarily small, and then since f_n is continuous we conclude that f is continuous.

By continuity of f, for any $z \in \Omega$ there is an open neighbourhood $U \subset \Omega$ such that $f(U)$ is entirely in a single coordinate chart of M. Then for any open disk $D \subset U$ centred at z such that $\overline{D} \subset U$, the sets $f_n(D)$ will eventually be in the same coordinate chart, so that $\rho_D(f_n, f) \to 0$ implies $\|\varphi \circ f_n - \varphi \circ f\|_D \to 0$, where φ is the coordinate map. Consequently, we have $\varphi \circ f \in \mathcal{O}(D)$, and since D is an open disk centred at an arbitrary point $z \in \Omega$, we get $f \in \mathcal{O}(\Omega, M)$. At the same time, the uniform convergence of $\varphi \circ f_n$ to $\varphi \circ f$ on D implies locally uniform convergence of $(\varphi \circ f_n)'$, which by (1) means that $f_n^\sharp$ converges locally uniformly on Ω. □

We recall here a version of the Arzelà-Ascoli theorem.

Theorem 7 (Arzelà-Ascoli). *Assume that M is a compact Riemann surface. Let $f_n : \Omega \to M$ be a sequence that is equicontinuous on compact subsets of Ω. Then there is a subsequence of $\{f_n\}$ that converges locally uniformly.*

In complex analysis, a very important notion is that of relative compactness in the topology associated to locally uniform convergence, so important that it has a name.

Definition 8. A family $\mathfrak{F} \subset \mathcal{O}(\Omega, M)$ is called *normal* if every sequence in $\mathfrak{F}$ has a subsequence that converges locally uniformly on Ω.

Theorem 9 (Marty). *Assume that M is a compact Riemann surface. Then $\mathfrak{F} \subset \mathcal{O}(\Omega, M)$ is normal if and only if $\sup_{f\in\mathfrak{F}} \|f^\sharp\|_K < \infty$ for any compact set $K \subset \Omega$.*

Proof. Firstly, if there is a compact set $K \subset \Omega$ such that $\sup_{f\in\mathfrak{F}} \|f^\sharp\|_K = \infty$, then there is a sequence $\{f_n\} \subset \mathfrak{F}$ with $\|f_n^\sharp\|_K \to \infty$ as $n \to \infty$. Then by Theorem 6 a subsequence of $\{f_n\}$ cannot converge, meaning that the family $\mathfrak{F}$ cannot be normal.

As for the other direction, let $f \in \mathfrak{F}$. Also let $K \subset \Omega$ be a compact set, let $\alpha, \beta \in K$, and let h be such that $\beta = \alpha + nh$ for some large integer n. Then from the triangle inequality and the definition of $f^\sharp$, we have

$$\begin{aligned}\rho(f(\alpha), f(\beta)) &\leq \sum_{j=0}^{n-1} \rho(f(\alpha + jh), f(\alpha + (j + 1)h)) \\ &= \sum_{j=0}^{n-1} \Big(f^\sharp(\alpha + jh)|h| + o(|h|)\Big) \\ &= \sum_{j=0}^{n-1} f^\sharp(\alpha + jh)|h| + o(1),\end{aligned}$$

and by sending $n \to \infty$ and taking into account that $f^\sharp$ is continuous, we infer

$$\rho(f(\alpha), f(\beta)) \leq \int_{[\alpha\beta]} f^\sharp(z)|\mathrm{d}z| \leq \|f^\sharp\|_K|\beta - \alpha|.$$

Hence $\mathfrak{F}$ is equicontinuous on compact subsets, and an application of the Arzelà-Ascoli theorem finishes the proof. □

We end this section with a general version of Vitali's theorem, which will not used here, but is interesting in its own right.

Theorem 10 (Vitali). *If $\{f_n\} \subset \mathcal{O}(\Omega, M)$ is normal, and converges pointwise on a set that has an accumulation point in Ω, then $\{f_n\}$ converges locally uniformly on Ω.*

Proof. Suppose the opposite. Then there are two subsequences $\{g_m\}$ and $\{h_m\}$ of $\{f_n\}$, a sequence of points $\{z_m\} \subset K$ from a compact set $K \subset \Omega$, and a number $\delta > 0$, such that $\rho(g_m(z_m), h_m(z_m)) \geq \delta$ for all m. The sequences $\{g_m\}$ and $\{h_m\}$ are both normal, so up to subsequences they converge respectively to functions $g \in \mathcal{O}(\Omega, M)$ and $h \in \mathcal{O}(\Omega, M)$. Without loss of generality, we may assume $z_m \to z \in K$, so that $\rho(g(z), h(z)) \geq \delta$.

On the other hand, $\{f_n\}$ converges pointwise on a set that has an accumulation point in Ω, forcing $g = h$ on that set. Then the identity theorem concludes that $g = h$ on Ω. □

3. Zalcman's lemma

The following remarkable result gives a precise characterization of loss of normality.

Lemma 11 (Zalcman). *Suppose that $\mathfrak{F} \subset \mathcal{O}(\Omega, M)$ is not normal. Then there exist*

- $z_n \in \Omega$, *with* $z_n \to z \in \Omega$,
- $\varepsilon_n > 0$, *with* $\varepsilon_n \to 0$, *and*
- $f_n \in \mathfrak{F}$ *such that the sequence of functions* $\zeta \mapsto f_n(z_n+\varepsilon_n\zeta)$ *converges locally uniformly on* $\mathbb{C}$ *to a nonconstant function* $g \in \mathcal{O}(\mathbb{C}, M)$ *with* $\|g^\sharp\|_{\mathbb{C}} \leq g^\sharp(0) = 1$.

Proof. By Marty's theorem, there exist a sequence of points w_n in a compact set $K \subset \Omega$, and a sequence of functions $f_n \in \mathfrak{F}$ such that $f_n^\sharp(w_n) \to \infty$. Without loss of generality we may assume that $\overline{\mathbb{D}} \subset \Omega$, $w_n \to 0$, and $f_n^\sharp(w_n) > 0$. Let $z_n \in \mathbb{D}$ and consider the disk centred at z_n of radius $1 - |z_n|$. We blow up this disk $D_{1-|z_n|}(z_n)$ so that it becomes a disk of radius R_n with $R_n \to \infty$ as $n \to \infty$. More precisely, we define

$$h_n(\zeta) = f_n(z_n + \varepsilon_n\zeta), \quad \zeta \in D_{R_n}, \qquad \text{where} \qquad \varepsilon_n = \frac{1 - |z_n|}{R_n}.$$

Note that z_n and R_n are assumed to satisfy $z_n \in \mathbb{D}$ and $R_n \to \infty$; otherwise they are arbitrary. This freedom will be used shortly. Fix some large $R > 0$, and for all n large enough so that $R_n > R$, let us look at the normality of the sequence of functions h_n. We have

$$h_n^\sharp(\zeta) = \varepsilon_n f_n^\sharp(z_n + \varepsilon_n\zeta), \qquad \zeta \in D_{R_n},$$

since

$$\begin{aligned}\rho(h_n(\zeta), h_n(\zeta + h)) &= \rho(f_n(z_n + \varepsilon_n\zeta), f_n(z_n + \varepsilon_n\zeta + \varepsilon_n h)) \\ &= f_n^\sharp(z_n + \varepsilon_n\zeta) \cdot \varepsilon_n|h| + o(\varepsilon_n|h|).\end{aligned}$$

Trying to bound $h_n^\sharp(\zeta)$ from above for $\zeta \in D_R$, we get

$$\begin{aligned}\varepsilon_n f_n^\sharp(z_n + \varepsilon_n\zeta) &= (1 - |z_n|)f_n^\sharp(z_n + \varepsilon_n\zeta)/R_n \\ &\leq (1 - |z_n + \varepsilon_n\zeta| + \varepsilon_n|\zeta|)f_n^\sharp(z_n + \varepsilon_n\zeta)/R_n \\ &\leq (1 - |z_n + \varepsilon_n\zeta|)f_n^\sharp(z_n + \varepsilon_n\zeta)/R_n + (R/R_n)\varepsilon_n f_n^\sharp(z_n + \varepsilon_n\zeta),\end{aligned}$$

where we have used the definition of ε_n in the first line, the triangle inequality in the second, and $|\zeta| < R$ in the third. Since $R/R_n < 1$, we can write this as

$$\begin{aligned}(1 - R/R_n)h_n^\sharp(\zeta) &\leq (1 - |z_n + \varepsilon_n\zeta|)f_n^\sharp(z_n + \varepsilon_n\zeta)/R_n \\ &\leq \max_{|z|\leq 1}(1 - |z|)f_n^\sharp(z)/R_n.\end{aligned}$$

We have $1 - R/R_n \to 1$ as $n \to \infty$, therefore by choosing R_n so that the right hand side of the preceding inequality stays bounded, we can ensure the boundedness of $h_n^\sharp$ on D_R uniformly in n. In particular, the choice

$$R_n = \max_{|z|\leq 1} f_n^\sharp(z)(1 - |z|),$$

is convenient, as this implies

$$h_n^\sharp(\zeta) \leq \frac{1}{1 - R/R_n} \to 1 \qquad \text{as} \qquad n \to \infty.$$

We have $R_n \to \infty$ because $f_n^\sharp$ is unbounded near 0. Moreover, we have

$$h_n^\sharp(0) = \varepsilon_n f_n^\sharp(z_n) = \frac{(1-|z_n|)f_n^\sharp(z_n)}{R_n} = \frac{f_n^\sharp(z_n)(1-|z_n|)}{\max\limits_{|z|\le 1} f_n^\sharp(z)(1-|z|)}$$

So if we could choose $z_n \in \mathbb{D}$ to be a maximum of $z \mapsto f_n^\sharp(z)(1-|z|)$, that is,

$$f_n^\sharp(z_n)(1-|z_n|) = R_n,$$

then we would have $h_n^\sharp(0) = 1$ for all sufficiently large n. But since $z \mapsto f_n^\sharp(z)(1-|z|)$ is continuous, is positive at $z = w_n$ and vanishes for $|z| = 1$, it achieves its maximum in $\mathbb{D}$.

The outcome of the preceding paragraph is that for any $R > 0$, there is n_R such that the sequence $\mathfrak{H}_0 = \{h_n : n \ge n_R\} \subset \mathcal{O}(D_R, M)$ is normal. So there is a subsequence $\mathfrak{H}_1 = \{h_{11}, h_{12}, \dots\}$ of $\mathfrak{H}_0$ such that $h_{1k} \to g_1 \in \mathcal{O}(D_R, M)$ as $k \to \infty$. But a tail of $\mathfrak{H}_1$ is normal in $\mathcal{O}(D_{2R}, M)$, so there is a subsequence $\mathfrak{H}_2 = \{h_{21}, h_{22}, \dots\} \subset \mathcal{O}(D_{2R}, M)$ of $\mathfrak{H}_1$ such that $h_{2k} \to g_2 \in \mathcal{O}(D_{2R}, M)$ as $k \to \infty$. Continuing this process, we get a nested sequence $\mathfrak{H}_1 \supset \mathfrak{H}_2 \supset \dots$ of sequences $\mathfrak{H}_m = \{h_{m1}, h_{m2}, \dots\} \subset \mathcal{O}(D_{mR}, M)$, with the property that $h_{mk} \to g_m \in \mathcal{O}(D_{mR}, M)$ as $k \to \infty$. From the nestedness, g_m and g_{m+1} agree on D_{mR}, so that the sequence $g_1, g_2, \dots$ defines a function $g \in \mathcal{O}(\mathbb{C}, M)$. Then the diagonal sequence $\{h_{kk}\}$, which is obviously a subsequence of the original sequence $\{h_n\}$, is defined eventually on any disk D_R, and converges uniformly there to g. Finally, note that we have $g^\sharp(0) = 1$ and $\|g^\sharp\|_{\mathbb{C}} \le 1$ by construction. □

4. Montel's theorem

We have encountered a version of Montel's theorem during the proof of the Riemann mapping theorem, which is usually called the thesis version of Montel's theorem. The following strengthened version relates to the thesis version the same way Picard's little theorem relates to Liouville's theorem.

Theorem 12 (Montel)**.** *If $\mathfrak{F} \subset \mathcal{O}(\Omega, \hat{\mathbb{C}})$ omits 3 values, then $\mathfrak{F}$ is normal.*

Proof. Without loss of generality, we may assume that $\Omega = \mathbb{D}$, and that $\mathfrak{F}$ omits 0, 1, and ∞. In particular, $\mathfrak{F} \subset \mathcal{O}(\mathbb{D})$ and each $f \in \mathfrak{F}$ admits a holomorphic n-th root for any $n \in \mathbb{N}$. Let us collect all 2^n-th roots of elements of $\mathfrak{F}$ and form the family

$$\mathfrak{F}_n = \{g \in \mathcal{O}(\mathbb{D}) : g^{2^n} = f \text{ pointwise for some } f \in \mathfrak{F}\}.$$

It is obvious that $\mathfrak{F}_n$ omits 0, ∞, and the 2^n-th roots of unity.

Anticipating a contradiction, suppose that $\mathfrak{F}$ is not normal. Then $\mathfrak{F}_n$ is not normal, because convergence of a sequence implies convergence of the sequence composed of 2^n-th power of the elements from the original sequence. Let $g_n \in \mathcal{O}(\mathbb{C}, \hat{\mathbb{C}})$ be the limit function from Zalcman's lemma applied to $\mathfrak{F}_n$. We have $\|g_n^\sharp\| \le g_n^\sharp(0) = 1$ and g_n omits 0, ∞, and the 2^n-th roots of unity. In particular, $\{g_n\}$ is normal by Marty's theorem, and passing to a subsequence, the limit function $g = \lim g_n \in \mathcal{O}(\mathbb{C}, \hat{\mathbb{C}})$ omits 0, ∞, and the 2^n-th roots of unity for all n. Moreover, g is not constant since $g^\sharp(0) = 1$. It follows from the open mapping theorem that g omits $\partial\mathbb{D}$, hence either $g(\mathbb{C}) \subset \mathbb{D}$ or $g(\mathbb{C}) \subset \mathbb{C} \setminus \mathbb{D}$. Finally, Liouville's theorem applied to either g or $1/g$ implies that g is constant, reaching a contradiction. □

Now we are ready prove Picard's big theorem, which we rephrase here for convenience.

Theorem 13. *If $f \in \mathcal{O}(\mathbb{D}^\times)$ omits 2 values in $\mathbb{C}$, then f extends to $f \in \mathcal{O}(\mathbb{D}, \hat{\mathbb{C}})$.*

Proof. With a positive sequence $\varepsilon_n \to 0$, let $g_n(z) = f(\varepsilon_n z)$ for $z \in \mathbb{D}^\times$. Then $\{g_n\}$ omits 3 values in $\hat{\mathbb{C}}$, so it is normal. Passing to a subsequence, let $g = \lim g_n \in \mathcal{O}(\mathbb{D}^\times, \hat{\mathbb{C}})$. So either g is a meromorphic function on $\mathbb{D}^\times$, or $g \equiv \infty$.

If $g \not\equiv \infty$, then there is a circle ∂D_r that does not pass through any pole of g, i.e., such that $\|g\|_{\partial D_r} < M$ for some $M > 0$. This means that $\|g_n\|_{\partial D_r} < M$ for all large n, or in other words, that $|f(z)| < M$ for $|z| = \varepsilon_n r$ for all large n. By the maximum principle, $|f| < M$ on the annulus of inner and outer radii $\varepsilon_{n+1} r$ and $\varepsilon_n r$ respectively, and this is true for all large n, hence 0 is a removable singularity of f.

For the case $g \equiv \infty$, we have $|f(z)| \to \infty$ as $z \to 0$, so 0 is a pole of f. $\square$