Postgraduate notes on complex analysis

J.K. Langley

For Hong, Helen and Natasha

Preface

These notes originated from a set of lectures on basic results in Nevanlinna theory and their application to ordinary differential equations in the complex domain, given at the Christian-Albrechts-Universität zu Kiel in December 1998. Over the years additional topics have been added, such as some elements of potential theory which are of use in value distribution theory, including the important technique of harmonic measure. Analytic continuation and singularities of the inverse function are also discussed, and the various themes are brought together in the Denjoy-Carleman-Ahlfors theorem and a recent theorem of Bergweiler and Eremenko concerning asymptotic values of entire and meromorphic functions.

The aim has been to develop in a single set of notes some of the key concepts and methods of function theory, in a form suitable for a postgraduate student starting out in the area. The notes have drawn on many sources, and these are indicated in the course of the development.

I would like to thank several people for drawing my attention to numerous obscurities and typos in earlier versions of these notes. These include my PhD students James Hinchliffe, Guy Kendall, Eleanor Lingham, Abdullah Alotaibi, Rob Trickey, Dan Nicks, Matt Buck and Asim Asiri, as well as Professor Christian Berg of the University of Copenhagen, who used parts of these notes in a graduate lecture course.

Last updated 27-1-2014

Contents

Chapter 1

Some topics from real analysis

This chapter contains a number of topics from real analysis. They have nothing in particular in common except that they all play a useful role in various aspects of function theory.

1.1 Convex functions

The property of convexity plays an important role in function theory because a number of key quantities associated with entire, meromorphic and subharmonic functions turn out to be convex functions of $\log r$. A good reference for this section is Chapter 5 of Royden's book [63], which along with Rudin's classic text [64] will be the main source for measure theoretic results used in these notes.

The real-valued function f is convex on the open interval $I = (p, q)$, $-\infty \leq p < q \leq \infty$, if

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \quad \text{for} \quad p < a < x < b < q.$$

This says that the graph of f over the closed interval $[a, b]$ lies on or below the straight line from $(a, f(a))$ to $(b, f(b))$. Rearranging, we find that

$$C(x,a) = \frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x} \quad \text{for} \quad a < x < b. \tag{1.1}$$

Keeping a fixed in (1.1) we get $C(x, a) \leq C(b, a)$ for $a < x < b$. So $C(x, a)$ is non-decreasing on (a, q) and the right derivative

$$f'_R(a) = \lim_{x \to a+} C(x,a) \leq C(b,a)$$

exists, with $f'_R(a) < \infty$ for every $a \in I$. Next, keeping b fixed in (1.1) we find that $C(b, x)$ is non-decreasing on (p, b) and the left derivative

$$f'_L(b) = \lim_{x \to b-} C(x,b) = \lim_{x \to b-} C(b,x) \geq C(b,a)$$

exists and satisfies $f'_L(b) > -\infty$ for all $b \in I$. Moreover, (1.1) gives $C(x, a) \leq C(b, x)$ for $a < x < b$ and so $f'_R(a) \leq f'_L(b)$ for $a < b$. Now let $a \to x-, b \to x+$ in (1.1), which gives

$$f'_L(x) \leq f'_R(x).$$

So

$$f'_L(a) \leq f'_R(a) \leq f'_L(b) \leq f'_R(b) \quad \text{for} \quad a < b.$$

Thus both left and right derivatives are real-valued non-decreasing functions, and f is continuous on I.

Fix $n \in \mathbb{N}$. If

$$f'_L(x) < f'_R(x) - 1/n \tag{1.2}$$

then for $y > x$ we have $f'_L(x) < f'_L(y) - 1/n$. Hence on any closed interval $[a,b] \subseteq I$ there are finitely many points x satisfying (1.2), because if $x_1, \ldots, x_m$ are such points with $a \leq x_1 < x_2 < \ldots < x_m \leq b$ then

$$f'_L(b) - f'_L(a) \geq \sum_{j=2}^{m} (f'_L(x_j) - f'_L(x_{j-1})) \geq \frac{m-1}{n}.$$

Thus there exists a countable set J such that on the complement $I \setminus J$ we have $f'_L = f'_R$. It follows that f is differentiable on $I \setminus J$, and f' is non-decreasing on $I \setminus J$.

1.2 The growth of real functions

1.2.1 O and o notation

Let $s(r), g(r)$ be functions defined on $[a, \infty)$, with $s(r)$ complex-valued and $g(r)$ real and positive. We say that $s(r) = O(g(r))$ as $r \to \infty$ if there exist constants K, L such that $|s(r)| \leq Kg(r)$ for all $r \geq L$. Thus, for example, $(r^2 + 3)\sin r = O(r^2)$ as $r \to \infty$. We write $s(r) = o(g(r))$ as $r \to \infty$ if $s(r)/g(r) \to 0$: for example $\log r = o(r)$.

We can also use this notation when r tends to a finite limit, for example, $r^2 + 3r = O(r)$ as $r \to 0+$, and for sequences, such as $2^n = o(n!)$ as $n \to \infty$.

1.2.2 lim sup and lim inf

Let $s(r)$ be a real-valued function defined on $[a, \infty)$. For each $r \geq a$, define

$$T_r = \{s(t) : t \geq r\}.$$

Obviously $T_r \subseteq T_u$ if $r \geq u \geq a$. Next define, for each $r \geq a$,

$$p(r) = p_s(r) = \inf T_r, \quad q(r) = q_s(r) = \sup T_r.$$

Here we use the convention that if a set is not bounded above then its sup is $+\infty$, while if a set is not bounded below then its inf is $-\infty$. We obviously now have

$$p(r) \leq s(r) \leq q(r). \tag{1.3}$$

Also $p(r)$ is a non-decreasing function, and $q(r)$ is a non-increasing function.

We define the "limsup" and "liminf" of $s(r)$ by

$$\tau = \limsup_{r \to \infty} s(r) = \lim_{r \to \infty} q(r), \quad \mu = \liminf_{r \to \infty} s(r) = \lim_{r \to \infty} p(r).$$

Obviously $\mu \leq \tau$, and we obtain the following properties of τ and μ.

(i) The limit $\lim_{r \to \infty} s(r)$ exists, with value L (possibly $\pm\infty$), if and only if $\tau = \mu = L$.

Proof. Suppose $s(r)$ has limit L. Assume first that $-\infty \leq y < L$. Then for large t we have $s(t) > y$. So $q(r) \geq p(r) \geq y$ for all large r, and so $\tau \geq \mu \geq y$. Similarly, if $y > L$ we get $y \geq \tau \geq \mu$.

Conversely, suppose that $\tau = \mu = L$. Then $p(r)$ and $q(r)$ tend to L as $r \to \infty$ and, by (1.3), so does $s(r)$.

(ii) Suppose $h < \tau$. Then for all large r we have $h < q(r)$ and so we can find $t \geq r$ with $s(t) > h$. Hence there exists a sequence $r_n \to \infty$ with $s(r_n) > h$.

(iii) Suppose $H > \tau$. Then $s(r) \leq q(r) < H$ for all large r.

Obviously properties (ii) and (iii) determine τ uniquely.

(iv) If $h < \mu$ then $s(r) > h$ for all large r. If $H > \mu$ there exists a sequence $r_n \to \infty$ with $s(r_n) < H$. These are proved in the same way as (ii), (iii), or using:

(v) We have
$$\limsup(-s(r)) = -\liminf s(r).$$
This is easy, since $q_{-s}(r) = -p_s(r)$ etc.

1.2.3 The order of a function

Let $s(r)$ be a non-negative real-valued function defined on $[a, \infty)$. The *order* of $s(r)$ is
$$\rho_s = \limsup_{r \to \infty} \frac{\log^+ s(r)}{\log r},$$
in which
$$\log^+ x = \max\{\log x, 0\}. \tag{1.4}$$
If $\rho_s < K < \infty$ then for all large enough r we have $\log^+ s(r) < K \log r$ and so $s(r) < r^K$.

1.2.4 Lemma

Suppose that $s(r), S(r)$ are non-negative real-valued functions defined on $[a, \infty)$ and that there exist $A, B, C, D \geq 1$ such that
$$S(r) < As(Br)(\log r)^C$$
for $r > D$. Then $\rho_S \leq \rho_s$.

Proof. Assume $\rho_s < K < \infty$, since if $\rho_s = \infty$ there is nothing to prove. For large r we then have
$$\log^+ S(r) \leq \log^+ A + \log^+ s(Br) + C \log\log r < K \log Br + o(\log r)$$
and so $\rho_S \leq K$.

1.2.5 Borel's lemma

Let $A > 1$. Let the function $T : [r_0, \infty) \to [1, \infty)$ be continuous from the right and non-decreasing. Then
$$T(r + 1/T(r)) \leq AT(r) \tag{1.5}$$
for all $r > r_0$ outside a set E of linear measure at most $\frac{A}{A-1}$.

Proof. Let r_1 be the infimum of those $r > r_0$ (if any) for which (1.5) is false, and set $r_1' = r_1 + 1/T(r_1)$. Continue this as follows: if $r_1, \ldots, r_n$ have been defined, put $r_n' = r_n + 1/T(r_n)$ and let r_{n+1} be the infimum of $r > r_n'$ for which (1.5) fails.

If $n \geq 1$ and r_n exists then, by the definition of r_n as an infimum, there exists a sequence $s_j \to r_n+$ such that (1.5) fails, i.e.

$$T(s_j + 1/(T(s_j)) > AT(s_j).$$

Since $T(r)$ is non-decreasing and continuous from the right, while $s_j \to r_n+$, this gives

$$T(s_j + 1/T(r_n)) \geq T(s_j + 1/(T(s_j)) > AT(s_j), \quad T(r_n') = T(r_n + 1/T(r_n)) \geq AT(r_n).$$

If, in addition, r_{n+1} exists then $T(r_{n+1}) \geq T(r_n') \geq AT(r_n)$.

We identify three cases. The first is that r_1 does not exist, in which case E is empty and there is nothing more to prove. The second is that $r_1, \ldots, r_n$ exist, but (1.5) holds for all $r > r_n'$. In this case, E is contained in the union of the intervals $[r_m, r_m']$ $(m = 1, \ldots, n)$ since, by the definition of the r_m, (1.5) holds for $r_m' < r < r_{m+1}$. Thus

$$\int_E dr \leq \sum_{m=1}^{n} (r_m' - r_m) = \sum_{m=1}^{n} T(r_m)^{-1} \leq \sum_{m=1}^{n} A^{1-m} T(r_1)^{-1} \leq \frac{A}{A-1}.$$

The final case is that in which the sequence r_n is infinite. In this case $r_n \to \infty$, for otherwise

$$r_n \to r^* \in (r_0, \infty), \quad r_n < r_n' \leq r_{n+1}, \quad r_n' \to r^*,$$

and

$$1/T(r^*) \leq 1/T(r_n) = r_n' - r_n \to 0,$$

which is impossible. As in the second case we get

$$\int_E dr \leq \sum_{m=1}^{\infty} (r_m' - r_m) \leq \sum_{m=1}^{\infty} A^{1-m} T(r_1)^{-1} \leq \frac{A}{A-1}.$$

1.3 Some results on certain integrals

1.3.1 The Riemann-Stieltjes integral

See Apostol's book [3, Ch. 7] for details of the Riemann-Stieltjes integral. Let f and h be real-valued functions on the interval $I = [a, b]$. Let $P = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[a, b]$. This means that $a = t_0 < t_1 < \ldots < t_n = b$; the t_j are then called vertices of P. By a Riemann-Stieltjes sum, we mean

$$S(P, f, h) = \sum_{k=1}^{n} f(s_k)(h(t_k) - h(t_{k-1})),$$

in which $t_{k-1} \leq s_k \leq t_k$. The case $h(x) = x$ gives the standard Riemann sums of ordinary integration.

We say that the Riemann-Stieltjes integral

$$\int_a^b f(x) dh(x)$$

exists and equals $L \in \mathbb{R}$ if the following is true. To each $\varepsilon > 0$ corresponds a partition P_0 of I such that $|S(P, f, h) - L| < \varepsilon$ for every refinement P of P_0 (this means that each vertex of P_0 is a vertex of P), regardless of how the s_k are chosen.

In particular, the integral exists if f is continuous and h is monotone [3, p.159]. Further, if

$$\int_a^b f(x)dh(x)$$

exists then so does

$$\int_a^b h(x)df(x)$$

and they satisfy the integration by parts formula [3, p.144]

$$\int_a^b f(x)dh(x) = f(b)h(b) - f(a)h(a) - \int_a^b h(x)df(x). \tag{1.6}$$

The following lemma concerning the interplay between sums and Riemann-Stieltjes integrals is useful in Nevanlinna theory.

1.3.2 Lemma

Let $-\infty < a = t_0 < t_1, \ldots < t_m = b < \infty$. Let the real-valued functions f and h be such that:
(i) f is continuous on $[a, b]$;
(ii) $h(x)$ is non-decreasing on $[a, b]$ and constant on each interval $[t_{j-1}, t_j)$, $j = 1, \ldots, m$.
Then

$$\int_a^b f dh = I = \sum_{j=1}^{m} f(t_j)(h(t_j) - h(t_{j-1})).$$

Proof. Let $\varepsilon > 0$ and choose $\delta > 0$ such that $\delta(h(b) - h(a)) < \varepsilon$. Next, choose $\eta > 0$ such that $|f(x) - f(y)| < \delta$ for $a \leq x < y \leq b$, $y - x < \eta$, which is possible since f is uniformly continuous on $[a, b]$. Fix a partition P_0 of $[a, b]$ such that (a) each t_j is a vertex of P_0 and (b) the distance between successive vertices of P_0 is less than η.

Now let P be any refinement of P_0. Then properties (a) and (b) holds with P_0 replaced by P. For $j = 1, \ldots, m$ let x_j be the greatest vertex of P in $[a, t_j)$. Then $t_{j-1} \leq x_j < t_j$. By property (ii), any Riemann-Stieltjes sum using the partition P has the form

$$S(P, f, h) = \sum_{j=1}^{m} f(s_j)(h(t_j) - h(x_j)) = \sum_{j=1}^{m} f(s_j)(h(t_j) - h(t_{j-1})),$$

where $x_j \leq s_j \leq t_j$, because all other subintervals contribute nothing to $S(P, f, h)$. But then, since h is non-decreasing,

$$|S(P, f, h) - I| \leq \sum_{j=1}^{m} |f(s_j) - f(t_j)|(h(t_j) - h(t_{j-1})) < \delta \sum_{j=1}^{m} (h(t_j) - h(t_{j-1})) = \delta(h(b) - h(a)) < \varepsilon.$$

1.3.3 Lemma

Let $g(r)$ be a non-negative measurable function on $[0, \infty)$, with $\int_0^r g(t)dt < \infty$ for every finite $r > 0$. Let h be the non-decreasing function

$$h(r) = \int_0^r g(t)dt.$$

Let f be real-valued and continuous on $[a,\infty)$. Then for each real $r > a$ the Riemann-Stieltjes integral

$$\int_a^r f(t)dh(t)$$

and the Lebesgue integral

$$\int_a^r f(t)g(t)dt$$

are equal.

Proof. Let $r > a$ and $\varepsilon > 0$ and take $\delta > 0$ with

$$\delta\int_a^r g(t)dt < \varepsilon.$$

Pick $\eta > 0$ so that $|f(x) - f(y)| < \delta$ for $a \leq x < y \leq r$, $y - x < \eta$. Fix a partition P_0 of $[a,r]$ such that the distance between successive vertices of P_0 is less than η. Let P be a refinement of P_0, with vertices $a = t_0 < t_1 < \ldots < t_n = r$. Let $t_{k-1} \leq s_k \leq t_k$. The corresponding Riemann-Stieltjes sum $S(P,f,h)$ is given by

$$S(P,f,h) = \sum_{k=1}^n f(s_k)(h(t_k) - h(t_{k-1})) = \sum_{k=1}^n f(s_k)\int_{t_{k-1}}^{t_k} g(t)dt.$$

Hence

$$S(P,f,h) - \int_a^r f(t)g(t)dt = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (f(s_k) - f(t))g(t)dt$$

has modulus at most

$$\int_a^r \delta g(t)dt < \varepsilon.$$

1.3.4 Lemma

If $h > 0$ on $[0,2\pi]$ and h and $\log h$ are integrable,

$$\frac{1}{2\pi}\int_0^{2\pi}\log h(t)dt \leq \log\left(\frac{1}{2\pi}\int_0^{2\pi} h(t)dt\right).$$

This says that the average of $\log h$ is not more than the log of the average of h. To prove the lemma we set

$$m = \frac{1}{2\pi}\int_0^{2\pi} h(t)dt, \quad g(t) = h(t) - m > -m.$$

Then

$$\frac{1}{2\pi}\int_0^{2\pi} g(t)dt = m - m = 0.$$

Also

$$h = m(1 + g/m), \quad \log h(t) = \log m + \log(1 + g(t)/m) \leq \log m + g(t)/m,$$

using the fact that $\log(1+x) \leq x$ for $x > -1$, which holds since $p(x) = \log(1+x) - x$ has $p'(x) < 0$ for $x > 0$ and $p'(x) > 0$ for $-1 < x < 0$.

We now get

$$\frac{1}{2\pi}\int_0^{2\pi}\log h(t)dt \leq \log m + \frac{1}{2\pi}\int_0^{2\pi}(g(t)/m)dt = \log m.$$

This proves the lemma, which is a special case of *Jensen's inequality*.

1.4 The density of sets

Let E be a measurable subset of $[0, \infty)$. The following quantities give some idea of how large and widely spread the set E is [13, 38]. First, we set $\chi_E(t)$ to be 1 if t is in E, and 0 otherwise, and χ is then a measurable function. We define the upper and lower linear density of E by

$$D_E = \overline{\mathrm{dens}}(E) = \limsup_{r\to\infty} \frac{\int_0^r \chi_E(t)dt}{r}, \quad d_E = \underline{\mathrm{dens}}(E) = \liminf_{r\to\infty} \frac{\int_0^r \chi_E(t)dt}{r}.$$

Obviously $0 \leq d_E \leq D_E \leq 1$, and if E has finite measure then $D_E = 0$. It is also easy to see that

$$D_E = 1 - d_F, \quad d_E = 1 - D_F, \quad \text{where} \quad F = [0, \infty) \setminus E.$$

Next we define the upper and lower logarithmic densities, by

$$LD_E = \overline{\log \mathrm{dens}}(E) = \limsup_{r\to\infty} \frac{\int_1^r \chi_E(t)\frac{dt}{t}}{\log r}, \quad ld_E = \underline{\log \mathrm{dens}}(E) = \liminf_{r\to\infty} \frac{\int_1^r \chi_E(t)\frac{dt}{t}}{\log r}.$$

Again, it is obvious that $0 \leq ld_E \leq LD_E \leq 1$.

1.4.1 Example

Let $r_n = e^{e^n}, n \geq 1$, and let E be the union of the intervals $[r_n, er_n]$. Then $d_E = 0, D_E > 0$, $ld_E = LD_E = 0$.

Proof. Let $s_n = er_n$. Then

$$\int_0^{s_n} \chi_E(t)dt \geq \int_{r_n}^{s_n} dt = (e-1)r_n = (1 - 1/e)s_n,$$

and so $D_E \geq 1 - 1/e$. However,

$$\int_0^{r_n} \chi_E(t)dt \leq \int_0^{s_{n-1}} \chi_E(t)dt \leq s_{n-1} = o(r_n),$$

which gives $d_E = 0$.

Suppose now that r is large, with $r_n \leq r < r_{n+1}$. Then

$$\int_1^r \chi_E(t)\frac{dt}{t} \leq \sum_{j=1}^n \int_{r_j}^{s_j} \chi_E(t)\frac{dt}{t} = \sum_{j=1}^n \int_{r_j}^{s_j} \frac{dt}{t} = n = \log\log r_n \leq \log\log r.$$

So $LD_E = 0$.

1.4.2 Theorem

Let E be a measurable subset of $[0, \infty)$. Then $0 \leq d_E \leq ld_E \leq LD_E \leq D_E$.

Proof. We only need to prove that $LD_E \leq D_E$, because with $F = [0, \infty) \setminus E$ we get

$$d_E = 1 - D_F \leq 1 - LD_F = ld_E.$$

There is nothing to prove if $D_E = 1$ so assume that $D_E < K < 1$. Then

$$h(r) = \int_1^r \chi_E(t)dt \leq \int_0^r \chi_E(t)dt < Kr$$

for all large r. So there exists $C > 0$ with $h(r) < C + Kr$ for all $r \geq 1$. Lemma 1.3.3 and the integration by parts formula (1.6) for Riemann-Stieltjes integrals now give, for large r,

$$\int_1^r \chi_E(t)\frac{dt}{t} = \int_1^r \frac{1}{t}dh(t) = \frac{h(r)}{r} + \int_1^r \frac{h(t)}{t^2}dt,$$

which is at most

$$\frac{C}{r} + K + \int_1^r \frac{C}{t^2} + \frac{K}{t}dt \leq K\log r + O(1).$$

Thus $LD_E \leq K$.

1.5 Upper semi-continuity

Let X be a metric space. A function $u : X \to [-\infty, \infty)$ is called *upper semi-continuous* if the following is true: for every real t the set $\{x \in X : u(x) < t\}$ is open. Obviously if $X = \mathbb{R}^n$ then every upper semi-continuous function u is (Borel) measurable.

1.5.1 Theorem

Let X be a metric space, with metric d, and suppose that $u : X \to [-\infty, M]$ is upper semi-continuous for some $M \in \mathbb{R}$. Then there exist continuous functions $u_n : X \to \mathbb{R}$ with $u_1 \geq u_2 \geq u_3 \geq \ldots \geq u$, such that $u_n(x) \to u$ pointwise on X.

Proof. This proof is from [61]. If $u \equiv -\infty$ just take $u_n = -n$. Now assume that $u \not\equiv -\infty$ and for $x \in X$ and $n \in \mathbb{N}$ put

$$u_n(x) = \sup\{u(y) - nd(x,y) : y \in X\}.$$

Then clearly

$$u_n(x) \in (-\infty, M].$$

To prove that u_n is continuous we must estimate $u_n(x) - u_n(x')$, so assume without loss of generality that $u_n(x) \geq u_n(x')$. Take $\delta > 0$. Then the definition of u_n gives y with $u(y) - nd(x,y) > u_n(x) - \delta$. Then

$$u_n(x) - \delta - u_n(x') < u(y) - nd(x,y) - (u(y) - nd(x',y)) = nd(x',y) - nd(x,y) \leq nd(x,x').$$

Since δ may be chosen arbitrarily small it follows that $u_n(x) - u_n(x') \leq nd(x,x')$, and so each u_n is continuous. Clearly $u_1 \geq u_2 \geq \ldots$...., and choosing $y = x$ shows that $u_n \geq u$. Note that we have not yet used the fact that u is upper semi-continuous.

To show that $u_n(x) \to u(x)$, take $t \in \mathbb{R}$ with $u(x) < t$, and using the fact that u is upper semi-continuous take $r > 0$ such that

$$\sup\{u(y) : y \in D(x,r)\} < t.$$

Now

$$u_n(x) \leq \max\{\sup\{u(y) : y \in D(x,r)\}, \sup\{u(y) : y \in X\} - nr\} \leq \max\{t, M - nr\}.$$

We thus have $u_n(x) \leq t$ for large n.

Exercise: if $u(0) = 1$ and $u(x) = 0$ for all real $x \neq 0$, determine $u_n(x)$ for each x. Do the same for $v = -u$ (which is not upper semi-continuous).

1.5.2 Lemma

Let the function u be upper semi-continuous on a domain containing the compact subset K of $\mathbb{C}$. Then u has a maximum on K.

Proof. Let S be the supremum of $u(z)$ on K, and take $z_n \in K$ such that $u(z_n) \to S$. We may assume that z_n converges, and the limit w is in K since K is closed and bounded. But then $u(w) \geq S$, because if $u(w) < t < S$ then we get $u(z) < t$ near w and hence $u(z_n) < t$ for all large n. We also have $u(w) \leq S$, by the definition of S, and so $u(w) = S$.

Chapter 2

Entire functions

2.1 The growth of entire functions

2.1.1 Notation

For $z_0 \in \mathbb{C}$ and $r > 0$ the open Euclidean disc and circle of centre z_0 and radius r will be denoted by

$$D(z_0, r) = \{z \in \mathbb{C} : |z - z_0| < r\}, \quad S(z_0, r) = \{z \in \mathbb{C} : |z - z_0| = r\},$$

respectively. If $z_0 \in \mathbb{C}^* = \mathbb{C} \cup \{\infty\}$ then $D_q(z_0, r)$ is the spherical disc

$$D_q(z_0, r) = \{z \in \mathbb{C}^* : q(z, z_0) < r\}.$$

2.1.2 The maximum modulus

Let f be entire (i.e. an analytic function from the complex plane into itself). Let $r > 0$ and define

$$M(r, f) = \max\{|f(z)| : |z| = r\}. \tag{2.1}$$

By the maximum principle, we have

$$M(r, f) = \max\{|f(z)| : |z| \leq r\},$$

from which it follows immediately that $M(r, f)$ is non-decreasing. Note also that if $0 < r < s$ and $M(r, f) = M(s, f)$ we can choose z with $|z| = r$ and $|f(z)| = M(r, f)$. Thus $|f(w)| \leq |f(z)|$ for all w in $D(0, s)$ and so f is constant, again by the maximum principle, since $|f|$ has a local maximum. Hence $M(r, f)$ is strictly increasing if f is non-constant.

For an entire function f, we now define the order (of growth) ρ of f by

$$\rho = \rho(f) = \limsup_{r \to \infty} \frac{\log^+ \log^+ M(r, f)}{\log r},$$

in which $\log^+ x$ is defined by (1.4).

Example 1:
Let $f(z) = a_n z^n + \ldots + a_0$ be a polynomial in z. For $|z| \geq 1$ we have $|f(z)| \leq c|z|^n, c = \sum_{j=0}^n |a_j|$. Thus $\log M(r, f) \leq n \log r + \log c \leq (n + 1) \log r$ for $r \geq 1 + c$, and so $\log^+ \log^+ M(r, f) \leq \log \log r + O(1)$ as $r \to \infty$, and $\rho = 0$.

Example 2:
Let $f(z) = \exp(z^n)$, with n a positive integer. Then $\log M(r, f) = r^n$ and $\rho = n$.

Example 3:
Let $f(z) = \exp(\exp(z))$. Then $\log M(r, f) = e^r$ and $\rho = \infty$.

2.2 Wiman-Valiron theory

The Wiman-Valiron theory is concerned with determining the local behaviour of an entire function from its power series. The main references for this subject are [36], from which this chapter will draw extensively, and [71]. First, if

$$P(z) = a_n z^n + \ldots + a_0, \quad a_n \neq 0,$$

is a polynomial of positive degree n, and if z and z_0 are large, then we have

$$P(z) \sim \left(\frac{z}{z_0}\right)^n P(z_0) \quad \text{and} \quad \frac{P'(z)}{P(z)} \sim \frac{n}{z}.$$

If P is replaced by a non-polynomial entire function f then it is clear from Picard's theorem that no such asymptotic relation can hold for all large z and z_0, but the aim of the Wiman-Valiron theory is to obtain comparable estimates when z is close to z_0 and $|f(z_0)|$ is close to $M(|z_0|, f)$. Let

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \tag{2.2}$$

be a transcendental entire function (here "transcendental" means "not a rational function"). Thus $a_k \neq 0$ for infinitely many k.

2.2.1 The maximum term

We define the *maximum term* $\mu(r, f)$ as follows. For each $r \geq 0$ let

$$\mu(r) = \mu(r, f) = \max\{|a_k| r^k : k = 0, 1, 2, \ldots\}. \tag{2.3}$$

This $\mu(r, f)$ is well-defined, because for fixed r the terms $|a_k| r^k$ tend to 0 as $k \to \infty$. Obviously $\mu(0) = |a_0|$. Since f is non-constant there exists $k > 0$ with $a_k \neq 0$ and so we have $\mu(r) \geq |a_k| r^k > 0$ for $r > 0$, as well as

$$\lim_{r \to \infty} \mu(r, f) = \infty.$$

The first step is an initial comparison between the growth rates of $M(r, f)$ and $\mu(r, f)$.

2.2.2 Lemma

For $r > 0$ we have

$$\mu(r, f) \leq M(r, f) \leq 2\mu(2r, f). \tag{2.4}$$

Further, the orders of the functions $\log^+ M(r, f)$ and $\log^+ \mu(r, f)$ are equal, these being defined by

$$\rho_f = \limsup_{r \to \infty} \frac{\log^+ \log^+ M(r, f)}{\log r}, \quad \rho_\mu = \limsup_{r \to \infty} \frac{\log^+ \log^+ \mu(r, f)}{\log r}.$$

Proof. The first inequality of (2.4) comes from Cauchy's integral formula, since for $k \geq 0$ we have

$$|a_k| = \left| \frac{f^{(k)}(0)}{k!} \right| = \left| \frac{1}{2\pi i} \int_{|z|=r} \frac{f(z)}{z^{k+1}} \, dz \right| \leq \frac{1}{2\pi}(2\pi r) \frac{M(r,f)}{r^{k+1}} = \frac{M(r,f)}{r^k}.$$

The second inequality is proved as follows. For $k \geq 0$ we have

$$|a_k|(2r)^k \leq \mu(2r,f), \quad |a_k|r^k \leq 2^{-k}\mu(2r,f),$$

and so

$$M(r,f) \leq \sum_{k=0}^{\infty} |a_k|r^k \leq \sum_{k=0}^{\infty} 2^{-k}\mu(2r,f) = 2\mu(2r,f).$$

The last assertion of the lemma now follows from (2.4) and Lemma 1.2.4.

2.2.3 Lemma

$\mu(r,f)$ is continuous and non-decreasing on $[0,\infty)$, and there exists $R \geq 0$ such that $\mu(r)$ is strictly increasing on $[R,\infty)$.

Proof. By the definition (2.3) of μ and Lemma 2.2.2 we have $\mu(0) = |a_0| \leq \mu(r) \leq M(r,f) \to |a_0|$ as $r \to 0+$, and so $\mu(r)$ is continuous as $r \to 0+$. Now choose $m > 0$ with $a_m \neq 0$. If $r_0 > 0$ then there exists $k_0 > m$ such that $|a_k|(2r_0)^k < |a_m|r_0^m$ for $k > k_0$. So for $r_0 \leq r \leq 2r_0$ we have

$$\mu(r) \geq |a_m|r^m \geq |a_m|r_0^m$$

and so

$$\mu(r) = \max\{|a_k|r^k : 0 \leq k \leq k_0\}.$$

So on $[r_0, 2r_0]$ our $\mu(r)$ is the maximum of finitely many continuous functions and so continuous.

If $0 \leq r < s < \infty$ take n such that $\mu(r) = |a_n|r^n$. Then

$$\mu(s) \geq |a_n|s^n \geq |a_n|r^n = \mu(r), \tag{2.5}$$

so $\mu(r)$ is non-decreasing. Now take $R \geq 0$, so large that $|a_m|R^m \geq |a_0|$ for some $m > 0$ with $a_m \neq 0$. Then for $R \leq r < s < \infty$ we have $|a_m|r^m \geq |a_0|$ and so $\mu(r) = |a_n|r^n$ for some $n > 0$ with $a_n \neq 0$, which gives strict inequality in (2.5).

2.2.4 The central index

For $r > 0$ and $\mu(r)$ as above, we define the central index $\nu(r) = \nu(r,f)$ (also called $N(r)$) to be the largest k for which $|a_k|r^k = \mu(r,f)$. Note that if $a_0 = 0$ then $\nu(0)$ is not defined, whereas if $a_0 \neq 0$ then $\nu(0) = 0$.

Observe further that if $r > 0$ then $\mu(r) > 0$, and that if $k \neq n$ with $a_k a_n \neq 0$ then $|a_k|r^k = |a_n|r^n$ for exactly one positive value of r. Thus there are only countably many values of r for which there does not exist a unique n with $|a_n|r^n = \mu(r)$.

2.2.5 Example

For $f(z) = e^z$ and $f(z) = \sin z$, determine $\mu(r)$ and $\nu(r)$ (hint for e^z: consider those r for which $|a_k|r^k = |a_{k+1}|r^{k+1}$). Use Stirling's formula to compare $M(r,f)$ with $\mu(r)$.

2.2.6 Lemma

The central index $\nu(r)$ is non-decreasing on $(0,\infty)$, and $\nu(r) \to \infty$ as $r \to \infty$. Also, $\nu(r)$ is continuous from the right, i.e., for each $s > 0$,

$$\lim_{r\to s+} \nu(r) = \nu(s).$$

Proof. Suppose first that $0 < r < s$ and $\nu(r) = N$. If $N = 0$ then obviously $\nu(s) \geq \nu(r)$. Now suppose that $N > M \in \{0, 1, 2, \ldots\}$. Then we have

$$|a_N|r^N \geq |a_M|r^M, \quad |a_N|s^N = |a_N|r^N \left(\frac{s}{r}\right)^N \geq |a_M|r^M \left(\frac{s}{r}\right)^N \geq |a_M|s^M,$$

and so $\nu(s) \geq N$.

Now let $P > 0$ and choose $k \geq P$ be such that $a_k \neq 0$. Then if $m < k$ we have $|a_m|r^m < |a_k|r^k$ for all large r, and so $\nu(r) \geq k \geq P$ for all large r. This says precisely that $\nu(r)$ tends to ∞.

Now we prove that $\nu(r)$ is continuous from the right. Let $s > 0$ and put $N = \nu(s)$. Take $k_0 > N$ such that $|a_k|(2s)^k < \mu(s)$ for $k > k_0$ (this is possible since the terms $|a_k|(2s)^k$ tend to 0). Then

$$\mu(s) \leq \mu(r, f) = \max\{|a_k|r^k : 0 \leq k \leq k_0\}$$

for $s \leq r \leq 2s$. But N is the largest k for which $|a_k|s^k = \mu(s)$, so that $|a_k|s^k < \mu(s)$ for $k > N$. By continuity there exists δ with $0 < \delta < s$ such that

$$|a_k|r^k < |a_N|s^N = \mu(s)$$

for $s \leq r \leq s + \delta$ and for $N < k \leq k_0$. By the choice of k_0, we now have $|a_k|r^k < |a_N|r^N$ for $s \leq r \leq s + \delta$ and for all $k > N$. Hence $\nu(r) = N$ for $s \leq r \leq s + \delta$. A similar argument shows that $\nu(r)$ is continuous as $r \to 0+$ if $a_0 \neq 0$.

2.2.7 Lemma

The unbounded integer-valued function $\nu(r)$ has the following property. There exists a strictly increasing sequence $r_n \to \infty$, with $r_0 = 0$, such that $\nu(r)$ is constant on (r_0, r_1) and on $[r_n, r_{n+1})$, for each $n \geq 1$. Also if $0 < s < r$ then

$$\log \mu(r) = \log \mu(s) + \int_s^r \frac{\nu(t)\, dt}{t}. \tag{2.6}$$

For large r we have

$$\log^+ \mu(r) < \nu(r) \log r + O(1) \tag{2.7}$$

and

$$\nu(r) \log 2 \leq \log^+ \mu(2r), \quad \nu(r) \log r \leq \log^+ \mu(r^2). \tag{2.8}$$

The orders of growth of $\log^+ \mu(r)$ and $\nu(r)$ are the same i.e.

$$\limsup_{r\to\infty} \frac{\log^+ \log^+ \mu(r)}{\log r} = \limsup_{r\to\infty} \frac{\log^+ \nu(r)}{\log r}. \tag{2.9}$$

Proof. We just set $r_0 = 0$, and let r_n, $n \geq 1$, be the points in $(0,\infty)$ at which $\nu(r)$ is discontinuous. Here we note that if $r > 0$ and $\nu(r) = N$, the function ν cannot have more than N discontinuities in

$(0,r)$. Since $\nu(r)$ is continuous from the right and integer-valued, it must be constant on (r_0,r_1) and $[r_n,r_{n+1})$.

Now suppose that $\nu(r)=N$ for $r_n<r<r_{n+1}$. Then on this interval we have $\mu(r)=|a_N|r^N$ and so

$$\frac{d\log\mu(r)}{d\log r}=N. \tag{2.10}$$

Since $\mu(r)$ is continuous we get

$$\log\mu(b)-\log\mu(a)=\int_a^b\frac{\nu(t)\,dt}{t}$$

for $r_n\le a\le b\le r_{n+1}$. Adding these gives (2.6).

To prove (2.7) and (2.8), choose $s\ge 1$, so large that $\mu(s)\ge 1$. Then for $r\ge s$ we have, since $\nu(t)$ is non-decreasing,

$$\log\mu(r)\le\log\mu(s)+\nu(r)\int_s^r\frac{dt}{t}\le\log\mu(s)+\nu(r)\log r,$$

which gives (2.7). We also have

$$\log\mu(2r)\ge\int_r^{2r}\frac{\nu(t)\,dt}{t}\ge\nu(r)\int_r^{2r}\frac{dt}{t}=\nu(r)\log 2$$

and

$$\log\mu(r^2)\ge\int_r^{r^2}\frac{\nu(t)\,dt}{t}\ge\nu(r)\int_r^{r^2}\frac{dt}{t}=\nu(r)\log r.$$

This proves (2.8), the second inequality of which gives

$$\frac{\log\mu(r)}{\log r}\ge\frac{\nu(r^{1/2})}{2}\to\infty$$

as $r\to\infty$. Finally, (2.9) follows from (2.7), (2.8) and Lemma 1.2.4.

2.2.8 Lemma

Let $\varepsilon>0$. Then

$$N(r)=\nu(r)\le(\log\mu(r))^{1+\varepsilon}\le(\log M(r,f))^{1+\varepsilon} \tag{2.11}$$

for all $r\ge 1$ outside a set E of finite logarithmic measure, i.e.

$$\int_{[1,\infty)\cap E}\frac{dt}{t}<\infty.$$

Proof. Choose $s\ge 1$ with $\mu(s)>1$ and let F be the set of $r\ge s$ for which (2.11) fails. Then, for $R>s$, integration of (2.10) gives

$$\int_{[1,R]\cap F}\frac{dt}{t}\le\int_s^R\frac{N(t)\,dt}{t(\log\mu(t))^{1+\varepsilon}}=\frac{1}{\varepsilon}\left(\frac{1}{(\log\mu(s))^{\varepsilon}}-\frac{1}{(\log\mu(R))^{\varepsilon}}\right).$$

Letting $R\to\infty$ then shows that F has finite logarithmic measure, and so has E, since $E\backslash F$ is bounded.

2.2.9 The comparison sequences

Let (α_n) and (ρ_n) be sequences such that

$$\alpha_n > 0, \quad 0 < \rho_0 < \frac{\alpha_0}{\alpha_1}, \quad \frac{\alpha_{n-1}}{\alpha_n} < \rho_n < \frac{\alpha_n}{\alpha_{n+1}} \quad \text{for} \quad n \geq 1. \tag{2.12}$$

Note that suitable sequences (α_n) and (ρ_n) will be constructed subsequently.

2.2.10 Lemma

Let f be a transcendental entire function with $a_0 \neq 0$ in (2.2), and assume that the sequence (ρ_n) is bounded above in §2.2.9. A real number $r > 0$ will be called normal for f with respect to the sequences (α_n) and (ρ_n) if there exists an integer $N \geq 0$ with

$$|a_n|r^n \leq |a_N|r^N \frac{\alpha_n(\rho_N)^n}{\alpha_N(\rho_N)^N} \quad \text{for all } n \geq 0. \tag{2.13}$$

Then there exists an exceptional set E_0 of finite logarithmic measure such that every $r \geq 1$ with $r \notin E_0$ is normal, and satisfies (2.13) with $N = N(r)$.

Proof. It follows from (2.12) that

$$\frac{\alpha_n}{\alpha_N} < (\rho_N)^{N-n} \quad \text{for } n, N \geq 0,\ n \neq N. \tag{2.14}$$

For if $n < N$ then

$$\frac{\alpha_n}{\alpha_N} = \frac{\alpha_n}{\alpha_{n+1}} \cdots \frac{\alpha_{N-1}}{\alpha_N} < \rho_{n+1} \cdots \rho_N \leq (\rho_N)^{N-n},$$

while $n > N$ gives

$$\frac{\alpha_n}{\alpha_N} = \frac{\alpha_{N+1}}{\alpha_N} \cdots \frac{\alpha_n}{\alpha_{n-1}} < \frac{1}{\rho_N} \cdots \frac{1}{\rho_{n-1}} \leq \frac{1}{(\rho_N)^{n-N}}.$$

This proves (2.14), which now implies in particular that if (2.13) holds then

$$|a_n|r^n < |a_N|r^N \quad \text{for } n \neq N$$

and so $N = N(r) = \nu(r)$.

We assert that there exists a non-decreasing sequence (s_n) with limit ∞ and with the following properties: (i) we have $s_0 = 0$; (ii) if $s_n < s_{n+1}$ then $N(r) = n$ on $[s_n, s_{n+1})$. To see this, observe first that $N(0) = 0$ (because $a_0 \neq 0$) and that $N(r)$ is non-decreasing and continuous from the right, and integer-valued. So let

$$0 = p_0 < p_1 < \ldots$$

be the values taken by $N(r)$, and let $t_k = \min\{t \geq 0 : N(t) = p_k\}$, which exists because $N(r)$ is continuous from the right. So we set $s_0 = 0$ and then $s_1 = \ldots = s_{p_1} = t_1$, and $s_{p_1+1} = \ldots = s_{p_2} = t_2$ and so on.

Now we claim that

$$\left|\frac{a_n}{a_0}\right| \leq \frac{1}{s_1 \ldots s_n} \quad \text{for } n > 0. \tag{2.15}$$

We prove (2.15) by induction. For $0 \leq t < s_1$ we have $N(t) = 0$ and so $|a_1|t \leq |a_0|$, which gives (2.15) for $n = 1$ on letting $t \to s_1-$. Now let $n > 1$ and let m be the largest integer such that $s_m < s_n$. Then on $[s_m, s_n)$ we have

$$N(r) = m \quad \text{and} \quad |a_n|r^n \leq |a_m|r^m.$$

Let $r \to s_n-$. If $m = 0$ then $s_1 = \ldots = s_n$ and we have

$$|a_n| \le \frac{|a_0|}{(s_n)^n} = \frac{|a_0|}{s_1 \ldots s_n}$$

as required. On the other hand if $m > 0$ then we may assume by the induction hypothesis that (2.15) holds with n replaced by m and we get

$$|a_n| \le |a_m|(s_n)^{m-n} \le \frac{|a_0|}{s_1 \ldots s_m} \frac{1}{(s_n)^{n-m}} = \frac{|a_0|}{s_1 \ldots s_n}.$$

This proves (2.15).

It follows from (2.12) that, for $n \ge 1$,

$$\frac{\alpha_n}{\alpha_0} = \frac{\alpha_n}{\alpha_{n-1}} \ldots \frac{\alpha_1}{\alpha_0} > \frac{1}{\rho_n \ldots \rho_1}. \tag{2.16}$$

Combining (2.15) and (2.16) then gives, for $n \ge 1$,

$$\left| \frac{a_n}{\alpha_n} \right|^{1/n} \le \left(\frac{|a_0|}{\alpha_0} \frac{\rho_1}{s_1} \ldots \frac{\rho_n}{s_n} \right)^{1/n}. \tag{2.17}$$

Now we use the fact that (ρ_m) is assumed to be bounded above, from which it follows that if $T > 1$ then $s_m > T\rho_m$ for all $m > M$, say. This in turn gives, by (2.17),

$$\left| \frac{a_n}{\alpha_n} \right|^{1/n} \le \left(\frac{|a_0|}{\alpha_0} \frac{\rho_1}{s_1} \ldots \frac{\rho_M}{s_M} \right)^{1/n} \frac{1}{T^{(n-M)/n}} \le \frac{2}{\sqrt{T}}$$

for all large enough n. Hence

$$\lim_{n \to \infty} \left| \frac{a_n}{\alpha_n} \right|^{1/n} = 0,$$

and so if we set

$$F(z) = \sum_{n=0}^{\infty} A_n z^n, \quad A_n = \left| \frac{a_n}{\alpha_n} \right| \tag{2.18}$$

then F is an entire function.

The point now is to deduce properties of f from those of F. Suppose that $\rho > 0$ and that $M = \nu(\rho, F)$. Then for all $n \ne M$ we have, by (2.14) and (2.18),

$$\frac{|a_n|(\rho\rho_M)^n}{|a_M|(\rho\rho_M)^M} = \frac{\alpha_n A_n \rho^n (\rho_M)^n}{\alpha_M A_M \rho^M (\rho_M)^M} \le \left(\frac{\alpha_n}{\alpha_M} \right) (\rho_M)^{n-M} < 1. \tag{2.19}$$

This implies that $N(r) = \nu(r, f) = M$ for $r = \rho\rho_M$, and also that r is normal for f (with M taking the role of N in (2.13)).

Since $A_0 \ne 0$ we can define a sequence (S_n) for the function F, exactly as we defined (s_n) for f. If we now have $\nu(\rho, F) = n$ on (S_n, S_{n+1}) then we have $\nu(r, f) = n$ on $I_n = (S_n\rho_n, S_{n+1}\rho_n)$, and every r in the interval I_n is normal for f. We also have $S_{n+1}\rho_n < S_{n+1}\rho_{n+1}$, by (2.12). Hence all non-normal r for f lie in the union of the intervals $[S_{n+1}\rho_n, S_{n+1}\rho_{n+1}]$, each of which has logarithmic measure

$$\log \frac{\rho_{n+1}}{\rho_n}.$$

Since (ρ_n) is bounded above, these logarithmic measures have finite sum, and so the lemma is proved.

2.2.11 Construction of the sequences (α_n) and (ρ_n)

Choose $\sigma \in (1,2)$, and set

$$\alpha(t) = \int_0^t \beta(s)\,ds, \tag{2.20}$$

where

$$\beta(s) = -1 \quad (0 \le s \le 1), \quad \beta(s) = -\frac{1}{s^\sigma} \quad (1 \le s < \infty). \tag{2.21}$$

Then $\alpha(t)$ is a negative, strictly decreasing function on $(0,\infty)$, with a finite limit as $t \to \infty$. Set

$$\alpha_n = \exp\left(\int_0^n \alpha(t)\,dt\right), \quad \rho_n = \exp(-\alpha(n)). \tag{2.22}$$

Since $\alpha(t)$ is bounded below on $(0,\infty)$, the sequence (ρ_n) is bounded above. It is obvious that $\alpha_n > 0$.

To check the remaining conditions of (2.12) we note that, for $n \ge 1$,

$$\log\frac{\alpha_n}{\alpha_{n-1}} = \int_{n-1}^n \alpha(t)\,dt > \int_{n-1}^n \alpha(n)\,dt = \alpha(n) = \log\frac{1}{\rho_n},$$

and also that, this time for $n \ge 0$,

$$\log\frac{1}{\rho_n} = \alpha(n) > \int_n^{n+1} \alpha(t)\,dt = \log\frac{\alpha_{n+1}}{\alpha_n}.$$

This shows that sequences with the required properties do exist.

2.2.12 Lemma

The construction of §2.2.11 gives, for $n, N \ge 0$, and $k = n - N \ne 0$,

$$\frac{\alpha_n(\rho_N)^n}{\alpha_N(\rho_N)^N} \le \exp\left(-\frac{k^2}{2(N+|k|)^\sigma}\right). \tag{2.23}$$

Proof. For $n \ne N$ we have, on integrating by parts,

$$\begin{aligned}
\frac{\alpha_n(\rho_N)^n}{\alpha_N(\rho_N)^N} &= \exp\left(\int_N^n \alpha(t)\,dt\right)\exp(-\alpha(N)(n-N)) \\
&= \exp\left(\int_N^n (\alpha(t)-\alpha(N))\,dt\right) \\
&= \exp\left(\int_N^n (n-t)\beta(t)\,dt\right).
\end{aligned}$$

If $n > N$ then, since $-\beta(t)$ is positive and non-increasing,

$$-\int_N^n (n-t)\beta(t)\,dt \ge -\beta(n)\int_N^n (n-t)\,dt = \frac{(n-N)^2}{2n^\sigma} = \frac{k^2}{2(N+|k|)^\sigma}.$$

On the other hand, if $n < N$ then, again since $-\beta(t)$ is positive and non-increasing,

$$-\int_N^n (n-t)\beta(t)\,dt = \int_n^N (t-n)(-\beta(t))\,dt \ge -\beta(N)\int_n^N (t-n)\,dt = \frac{(N-n)^2}{2N^\sigma} \ge \frac{k^2}{2(N+|k|)^\sigma}.$$

2.2.13 Lemma

Let $1 < \sigma < 2$ and let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be a transcendental entire function with central index $N(r)$ and maximum term $\mu(r)$. Then for all large r outside a set of finite logarithmic measure we have, with $N = N(r)$,

$$\frac{|a_{N+k}|r^{N+k}}{\mu(r)} \leq \exp\left(-\frac{k^2}{2(N+|k|)^\sigma}\right). \tag{2.24}$$

Proof. Obviously there is nothing to prove if $k = 0$. Suppose first that $a_0 \neq 0$. Then we take the sequences (α_n) and (ρ_n) and the set of non-normal r has finite logarithmic measure. Moreover if r is normal then combining (2.13) with (2.23) gives, with $n = N + k$ and $k \neq 0$,

$$|a_n|r^n \leq \mu(r)\frac{\alpha_n(\rho_N)^n}{\alpha_N(\rho_N)^N} \leq \mu(r)\exp\left(-\frac{k^2}{2(N+|k|)^\sigma}\right). \tag{2.25}$$

Now suppose that $a_0 = 0$. Then we may write $f(z) = z^p g(z)$ for some $p > 0$, where $g(z) = \sum_{k=0}^{\infty} c_k z^k$ is entire and $g(0) = c_0 \neq 0$. It is then easy to see that $c_n = a_{n+p}$ and $\mu(r) = r^p \mu(r, g)$, while $N(r) = \nu(r, g) + p$. Hence, for all large r outside a set of finite logarithmic measure, writing $\nu = \nu(r, g)$ and using (2.25) with f replaced by g gives

$$\frac{|a_{N+k}|r^{N+k}}{\mu(r)} = \frac{|c_{\nu+k}|r^{\nu+k}}{\mu(r,g)} \leq \exp\left(-\frac{k^2}{2(\nu+|k|)^\sigma}\right) \leq \exp\left(-\frac{k^2}{2(N+|k|)^\sigma}\right).$$

2.2.14 Comparison between $\nu(r, f)$ and $\nu(r, f')$

It is convenient to consider $g(z) = zf'(z) = \sum_{k=1}^{\infty} k a_k z^k$, and obviously $\nu(r, g) = \nu(r, f') + 1$. Now fix $\varepsilon > 0$, and suppose that r is large and lies outside the exceptional set E of Lemma 2.2.13, and set $N = \nu(r, f)$. Then for $n \leq N$ we have

$$n|a_n|r^n \leq N|a_n|r^n \leq N|a_N|r^N$$

and so $\nu(r, g) \geq N = \nu(r, f)$. Now take $n = N + k$ with $k \geq \varepsilon N$. Then $N + k \leq k(1 + 1/\varepsilon)$ and Lemma 2.2.13 gives

$$n|a_n|r^n \leq (N+k)|a_N|r^N \exp\left(-\frac{k^2}{2(N+|k|)^\sigma}\right) \leq c_1 k \exp(-c_2 k^{2-\sigma}) N|a_N|r^N,$$

where the positive constants c_1 and c_2 are independent of r. If N is large then so is k, and thus $n|a_n|r^n < (1/2)N|a_N|r^N$ for $n \geq (1+\varepsilon)N$, which forces $\nu(r, g) \leq (1+\varepsilon)N$.

We conclude that

$$\nu(r, f') \sim \nu(r, f) \quad \text{as } r \to \infty \text{ with } r \notin E, \text{ where } \int_E dt/t < \infty. \tag{2.26}$$

2.2.15 Lemma

Let $\alpha > 0$. Then

$$N(r\exp(N(r)^{-\alpha})) < (1+\alpha)N(r) \tag{2.27}$$

for all $r \geq 1$ outside a set of finite logarithmic measure.

Proof. Choose $R \geq 1$ with $N(R) \geq 1$ and set

$$s = \log r, \quad M(s) = N(r)^\alpha = N(e^s)^\alpha$$

for $s \geq S = \log R$. Then $M(s)$ is non-decreasing and continuous from the right, and $M(s) \geq 1$ for $s \geq S$. Choose $A > 1$ with $A^{1/\alpha} < 1 + \alpha$. The Borel lemma 1.2.5 gives

$$N(r\exp(N(r)^{-\alpha}))^\alpha = M(\log(r\exp(N(r)^{-\alpha}))) = M(s + 1/M(s)) \leq AM(s) = AN(r)^\alpha$$

for $s \geq S$ outside a set E_0 of finite measure. The corresponding exceptional set of r is just

$$F_0 = \{e^s : s \in E_0\}$$

and satisfies

$$\int_{F_0} \frac{dr}{r} = \int_{E_0} ds < \infty.$$

2.2.16 Estimates for sums of terms in the power series

Let $1 < \sigma < 2$ and let $\sigma < 2\tau < 2$. Let q be a non-negative integer, and let $f(z) = \sum_{k=0}^\infty a_k z^k$ be a transcendental entire function with maximum term $\mu(r)$ and central index $N(r)$. We will estimate

$$\sum_{|n-N(r)| \geq N(r)^\tau} n^q |a_n| \rho^n$$

for ρ close to r.

In order to do this, let r lie outside the exceptional sets of Lemmas 2.2.13 and 2.2.15, taking $\alpha = 1/4$ in the latter. Note that the union E^* of these exceptional sets has finite logarithmic measure (and does not depend on q). Write

$$N = N(r), \quad \mu_0(\rho) = |a_N|\rho^N, \tag{2.28}$$

where

$$|\log(\rho/r)| \leq N^{-\tau}. \tag{2.29}$$

We use $c_1, c_2, \ldots$ to denote positive constants which do not depend on r or ρ (although in general they will depend on f, σ, τ and q). Write

$$\rho_1 = r\exp(N(r)^{-1/4}), \quad M = N(\rho_1), \quad N \leq M \leq \frac{5N}{4}, \tag{2.30}$$

in which the last inequality follows from Lemma 2.2.15.

Then for r large enough, not in E^*, and $n > 2N$ we have, by (2.28), (2.29) and (2.30), the inequality $n - M \geq n - 5N/4 \geq c_1 n$ and the estimates

$$\begin{aligned}
\frac{|a_n|\rho^n}{\mu_0(\rho)} &= \frac{|a_n|\rho^n}{|a_N|\rho^N} = \frac{|a_n|\rho^n}{|a_M|\rho^M} \frac{|a_M|\rho^M}{|a_N|\rho^N} \\
&= \frac{|a_n|\rho_1^n}{|a_M|\rho_1^M} \left(\frac{\rho}{\rho_1}\right)^{n-M} \frac{|a_M|r^M}{|a_N|r^N} \left(\frac{\rho}{r}\right)^{M-N} \\
&\leq \left(\frac{\rho}{\rho_1}\right)^{n-M} \left(\frac{\rho}{r}\right)^{M-N} = \left(\frac{\rho_1}{r}\right)^{M-n} \left(\frac{\rho}{r}\right)^{n-N} \\
&\leq \exp((M-n)N^{-1/4} + (n-N)|\log(\rho/r)|) \\
&\leq \exp(-c_1 n N^{-1/4} + nN^{-\tau}) \leq \exp(-c_2 n N^{-1/4}),
\end{aligned} \tag{2.31}$$

using the fact that $\tau > 1/2$. Thus we have

$$\sum_{n>2N} n^q |a_n| \rho^n \leq \mu_0(\rho) \sum_{n>2N} n^q t^n, \quad t = \exp(-c_2 N^{-1/4}) < 1, \tag{2.32}$$

for ρ satisfying (2.29).

Now, since N is large,

$$\sum_{n>2N} n^q t^n < \sum_{n>N} n^q t^n = t^N \sum_{k=1}^{\infty} (N+k)^q t^k = t^N \sum_{k=1}^{\infty} (1+N/k)^q k^q t^k \le 2t^N N^q \sum_{k=1}^{\infty} k^q t^k.$$

But repeated differentiation of the geometric series shows that the power series $\sum_{k=1}^{\infty} k^q t^k$ may be written as a linear combination of

$$1, \frac{1}{1-t}, \ldots, \frac{1}{(1-t)^{q+1}},$$

with constant coefficients, independent of r and ρ. Since $0 < t < 1$ this gives

$$\frac{1}{1-t} = \frac{\exp(c_2 N^{-1/4})}{\exp(c_2 N^{-1/4}) - 1} \le c_3 N^{1/4}$$

and

$$\sum_{n>2N} n^q t^n \le \frac{c_4 t^N N^q}{(1-t)^{q+1}} \le c_5 t^N N^q N^{(q+1)/4}.$$

On recalling (2.32) we therefore have, for $r \notin E^*$ large enough,

$$\begin{aligned}\sum_{n>2N} n^q |a_n| \rho^n &\le \mu_0(\rho) c_5 t^N N^q N^{(q+1)/4} = c_5 \mu_0(\rho) \exp(-c_2 N^{3/4} + c_6 \log N) \\ &\le \mu_0(\rho) \exp(-c_7 N^{3/4}). \end{aligned} \tag{2.33}$$

We consider next those n satisfying

$$0 \le n = N + p \le 2N, \quad |p| \ge N^\tau.$$

For these n and for ρ satisfying (2.29) we have, by Lemma 2.2.13 and the fact that $2\tau > \sigma$ gives $\sigma - \tau < \tau$,

$$\begin{aligned} n^q \frac{|a_n|\rho^n}{\mu_0(\rho)} &= n^q \left(\frac{\rho}{r}\right)^p \frac{|a_n| r^n}{a_N |r^N} \\ &\le (2N)^q \left(\frac{\rho}{r}\right)^p \exp(-p^2/2(N+|p|)^\sigma) \\ &\le (2N)^q \exp(|p| N^{-\tau} - p^2/2(2N)^\sigma) \\ &= (2N)^q \exp(|p| N^{-\sigma}(N^{\sigma-\tau} - c_8|p|)) \\ &\le (2N)^q \exp(|p| N^{-\sigma}(o(N^\tau) - c_8|p|)) \\ &\le (2N)^q \exp(|p| N^{-\sigma}(o(|p|) - c_8|p|)) \\ &\le (2N)^q \exp(-c_9 p^2 N^{-\sigma}) \\ &\le (2N)^q \exp(-c_9 N^{2\tau-\sigma}) = (2N)^q \exp(-c_9 N^{2\varepsilon}), \end{aligned}$$

where $2\varepsilon = 2\tau - \sigma > 0$. Hence we get, for $r \notin E^*$ large, and for ρ satisfying (2.29),

$$\begin{aligned} \sum_{n \le 2N, |n-N| \ge N^\tau} n^q \frac{|a_n|\rho^n}{\mu_0(\rho)} &\le (2N)^{q+1} \exp(-c_9 N^{2\varepsilon}) \\ &= \exp(-c_9 N^{2\varepsilon} + c_{10} \log N) \le \exp(-N^{3\varepsilon/2}). \end{aligned} \tag{2.34}$$

Combining (2.33) with (2.34) then gives the following fundamental lemma.

2.2.17 Lemma

Let $1 < \sigma < 2\tau < 2$. Then there exists $\delta > 0$ with the following property. Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be a transcendental entire function and let q be a non-negative integer. Then for all $r \geq 1$ outside a set E_1 of finite logarithmic measure we have, with the notation

$$N = \nu(r, f), \quad \mu_0(\rho) = |a_N|\rho^N, \tag{2.35}$$

the estimate

$$\sum_{|n-N|\geq N^\tau} n^q |a_n| \rho^n \leq \mu_0(\rho) \exp(-N^\delta) \quad \text{for} \quad |\log(\rho/r)| \leq N^{-\tau}. \tag{2.36}$$

We also obtain another comparison between the maximum modulus and the maximum term. It follows using Lemma 2.2.8 that, for r outside a perhaps larger set of finite logarithmic measure,

$$\sum_{|n-N|\geq N^\tau} n^q |a_n| r^n \leq \mu_0(r) = \mu(r, f), \quad N = N(r) \leq (\log \mu(r, f))^2$$

and so

$$\mu(r, f) \leq M(r, f) \leq \sum_{k=0}^{\infty} |a_k| r^k \leq 3N(r)^\tau \mu(r, f) \leq 3\mu(r, f)(\log \mu(r, f))^2. \tag{2.37}$$

2.2.18 A lemma concerning polynomials

Let λ, δ and ε be positive real numbers, and let $j \in \{0, 1\}$. Let

$$P(z) = \alpha_m z^m + \ldots + \alpha_0$$

be a polynomial of degree at most m. Then for $R \geq r > 0$ we have

$$|P^{(j)}(z)| \leq e^j \left(\frac{m}{r}\right)^j \left(\frac{R}{r}\right)^{m-j} M(r, P) \tag{2.38}$$

for $|z| \leq R$. Further, if $m^\varepsilon > e^2/(\delta\lambda)$ and $|z_0| = r > 0$ and $|P(z_0)| \geq \lambda M(r, P)$, then

$$|P(z) - P(z_0)| < \delta |P(z_0)| \quad \text{for} \quad |z - z_0| \leq \frac{r}{m^{1+\varepsilon}}. \tag{2.39}$$

Proof. We first prove (2.38) for $j = 0$. Let

$$M = M(r, P), \quad Q(z) = \frac{P(z) r^m}{z^m} = r^m \alpha_m + \ldots.$$

Then $Q(z)$ is analytic for $r \leq |z| \leq \infty$, with $Q(\infty) = r^m \alpha_m$. We also have $|Q(z)| \leq M$ on $|z| = r$, and so the maximum principle implies that $|Q(z)| \leq M$ for $|z| \geq r$. In particular, $M(R, P) \leq (R/r)^m M$, which gives (2.38) for $j = 0$, using the maximum principle again.

Next, we consider the case $j = 1$. Let $|z| \leq R$, and put $h = R/m$. Then (2.38) for $j = 0$ and Cauchy's integral formula lead to

$$\begin{aligned}
|P'(z)| &= \left| \frac{1}{2\pi i} \int_{|u-z|=h} \frac{P(u)}{(u-z)^2}\, du \right| \\
&\leq \frac{1}{h} \max\{|P(u)| : |u - z| = h\} \leq \frac{1}{h} M(R + h, P) \\
&\leq \frac{M}{h} \frac{(R + h)^m}{r^m} = \frac{M}{h} \frac{R^m}{r^m} \left(1 + \frac{1}{m}\right)^m \\
&= \frac{mMR^{m-1}}{r^m} \left(1 + \frac{1}{m}\right)^m \leq e \left(\frac{m}{r}\right) \left(\frac{R}{r}\right)^{m-1},
\end{aligned}$$

since $1+1/m < e^{1/m}$. This proves (2.38) for $j=1$.

To prove (2.39) let $S = r(1+m^{-1-\varepsilon})$. Then for z as in (2.39) we have

$$|P'(z)| \leq M(S,P') \leq M(r,P)e\left(\frac{m}{r}\right)\left(\frac{S}{r}\right)^{m-1} = \frac{emM(r,P)}{r}\left(1+\frac{1}{m^{1+\varepsilon}}\right)^{m-1} \leq \frac{e^2 mM(r,P)}{r}.$$

Hence we obtain, for such z,

$$|P(z)-P(z_0)| = \left|\int_{z_0}^{z} P'(t)\,dt\right| \leq \frac{r}{m^{1+\varepsilon}}\,\frac{e^2 mM(r,P)}{r} \leq \frac{e^2 M(r,P)}{m^{\varepsilon}} \leq \frac{e^2|P(z_0)|}{\lambda m^{\varepsilon}} < \delta|P(z_0)|$$

by the lower bound on m.

2.2.19 The main estimates at points near to the maximum modulus

Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be a transcendental entire function, let $1 < \sigma < 2\tau < 2$ and let δ and the exceptional set E_1, which has finite logarithmic measure, be as in Lemma 2.2.17. Choose $\lambda \in (0,1/2]$ and let ε be small and positive. In addition let $\tau < \gamma < 1$.

Let $r \notin E_1$ be large and set

$$N = \nu(r,f), \quad k = [N^{\tau}], \quad \mu_0(\rho) = |a_N|\rho^N, \tag{2.40}$$

where $[x]$ denotes the greatest integer not exceeding x. Then Lemma 2.2.17 implies that

$$\sum_{|n-N|>k} n|a_n|\rho^n \leq \mu_0(\rho)\exp(-N^{\delta}) \tag{2.41}$$

for

$$|\log(\rho/r)| \leq N^{-\gamma}. \tag{2.42}$$

Note that for ρ satisfying (2.42) we have

$$|k\log(\rho/r)| \leq N^{\tau-\gamma} = o(1), \quad \left(\frac{\rho}{r}\right)^k \sim 1, \tag{2.43}$$

as $r \to \infty$ with $r \notin E_1$.

Write

$$f(z) = \sum_{n=N-k}^{N+k} a_n z^n + \phi(z) = z^{N-k}P(z) + \phi(z), \tag{2.44}$$

where P is a polynomial of degree at most $m = 2k$. The aim will be to show that, for appropriate choice of z, the remainder term $\phi(z)$ is relatively small and the polynomial $P(z)$ does not vary too much, so that $f(z)$ is essentially controlled by the monomial z^{N-k}. To this end we apply Lemma 2.2.18 to P and P', with

$$R = r\exp(N^{-\gamma}),$$

to get

$$\begin{aligned} M(R,P) &\leq \left(\frac{R}{r}\right)^m M(r,P) \leq M(r,P)\exp(2N^{\tau-\gamma}) \sim M(r,P), \\ M(R,P') &\leq e\left(\frac{2k}{r}\right)\left(\frac{R}{r}\right)^{m-1} M(r,P) < \frac{12kM(r,P)}{r}. \end{aligned} \tag{2.45}$$

For $|z| = \rho$ satisfying (2.42), the estimate (2.41) and the relation (2.44) imply that

$$f(z) = z^{N-k}P(z) + o(\mu_0(\rho)) = z^{N-k}P(z) + o(M(\rho, f)), \tag{2.46}$$

from which it follows easily that

$$M(r, f) \sim r^{N-k}M(r, P). \tag{2.47}$$

Now choose z_0 with

$$|z_0| = r, \quad |f(z_0)| \geq 2\lambda M(r, f). \tag{2.48}$$

Then (2.46) gives

$$f(z_0) \sim z_0^{N-k}P(z_0), \quad |f(z_0)| \sim r^{N-k}|P(z_0)|, \tag{2.49}$$

and hence, using (2.47),

$$|P(z_0)| \sim r^{k-N}|f(z_0)| \geq 2\lambda r^{k-N}M(r, f) \geq \lambda M(r, P). \tag{2.50}$$

For $|z| = \rho$ satisfying (2.42) we may now write, using the first relation of (2.46), as well as (2.43) and (2.49),

$$\begin{aligned} \frac{f(z)}{z^N} &= z^{-k}P(z) + o(|a_N|) = z^{-k}P(z) + o(r^{-N}M(r, f)) \\ &= z^{-k}P(z) + o(r^{-N}|f(z_0)|) = z^{-k}P(z) + o(r^{-k}|P(z_0)|) \\ &= z^{-k}(P(z) + o(|P(z_0)|)). \end{aligned} \tag{2.51}$$

For ρ satisfying (2.42) we deduce, using (2.43) and (2.45), that

$$\begin{aligned} M(\rho, f) &\leq \rho^{N-k}(M(\rho, P) + o(|P(z_0)|)) = (1 + o(1))\rho^{N-k}M(r, P) \\ &\sim \left(\frac{\rho}{r}\right)^{N-k} M(r, f) \sim \left(\frac{\rho}{r}\right)^{N} M(r, f). \end{aligned} \tag{2.52}$$

Next, consider z satisfying

$$|\log(z/z_0)| \leq N^{-\gamma}. \tag{2.53}$$

For such z we have

$$|k \log(z/z_0)| = o(1), \quad \left(\frac{z}{z_0}\right)^k \sim 1, \quad |z - z_0| = O(rN^{-\gamma}) = o\left(\frac{r}{m^{1+\varepsilon}}\right), \tag{2.54}$$

since ε is small. Thus for z satisfying (2.53) we have $P(z) \sim P(z_0)$ by Lemma 2.2.18 and so (2.49), (2.51) and (2.54) give

$$f(z) \sim z^{N-k}P(z_0) \sim \left(\frac{z}{z_0}\right)^{N-k} f(z_0) \sim \left(\frac{z}{z_0}\right)^{N} f(z_0), \tag{2.55}$$

which is the main estimate of the Wiman-Valiron theory.

In particular, if we choose z_0 such that $|z_0| = r$ and $|f(z_0)| = M(r, f)$ then for z satisfying (2.53) and $|z| = \rho$ we get

$$|f(z)| \geq (1 - o(1)) \left(\frac{\rho}{r}\right)^N M(r, f)$$

and so (2.52) now becomes, for ρ satisfying (2.42),

$$M(\rho, f) \sim \left(\frac{\rho}{r}\right)^N M(r, f). \tag{2.56}$$

The next step is to estimate $f'(z)$. For $|z| = \rho$ as in (2.42), the function $\phi(z)$ of (2.44) satisfies, by (2.41),

$$|\phi'(z)| = \left| \sum_{|n-N|>k} n a_n z^{n-1} \right| \leq \frac{\mu_0(\rho)\exp(-N^\delta)}{\rho}. \tag{2.57}$$

Differentiating (2.44) thus gives, for $|z| = \rho$ satisfying (2.42),

$$\begin{aligned} f'(z) &= (N-k)z^{N-k-1}P(z) + z^{N-k}P'(z) + \phi'(z) \\ &= (N-k)z^{N-k-1}P(z) + z^{N-k}P'(z) + o(\rho^{-1}|a_N|\rho^N) \end{aligned}$$

and hence, using (2.45) and (2.50),

$$\begin{aligned} \frac{f'(z)}{z^N} &= (N-k)z^{-k-1}P(z) + z^{-k}P'(z) + o(\rho^{-1}r^{-N}M(r,f)) \\ &= (N-k)z^{-k-1}P(z) + z^{-k}P'(z) + o(\rho^{-1}r^{-k}|P(z_0)|) \\ &= z^{-k-1}\left[(N-k)P(z) + zP'(z) + o(|P(z_0)|)\right] \\ &= z^{-k-1}\left[(N-k)P(z) + O(kM(r,P))\right] \\ &= z^{-k-1}\left[(N-k)P(z) + O(k|P(z_0)|)\right]. \end{aligned} \tag{2.58}$$

In particular, we obtain an upper bound for $M(\rho, f')$ as follows. For $|z| = \rho$ satisfying (2.42), applying (2.43) and (2.45) again, as well as (2.47) and (2.58), gives, since $k = o(N)$,

$$M(\rho, f') \leq (1+o(1))N\rho^{N-k-1}M(r,P) \sim N\rho^{N-k-1}r^{k-N}M(r,f) \sim \frac{N}{\rho}\left(\frac{\rho}{r}\right)^N M(r,f). \tag{2.59}$$

Next, we estimate $f'(z)$ for z satisfying (2.53). Again we have $P(z) \sim P(z_0)$ and so (2.55) and (2.58) lead to

$$f'(z) \sim z^{N-k-1}NP(z_0) \sim \frac{N}{z}f(z) \sim \frac{N}{z}\left(\frac{z}{z_0}\right)^N f(z_0). \tag{2.60}$$

Again, if we choose z_0 such that $|z_0| = r$ and $|f(z_0)| = M(r,f)$ then we obtain a lower bound for $M(\rho, f')$ and (2.59) becomes, for $|z| = \rho$ satisfying (2.42), using (2.56),

$$M(\rho, f') \sim \frac{N}{\rho}\left(\frac{\rho}{r}\right)^N M(r,f) \sim \frac{N}{\rho}M(\rho,f). \tag{2.61}$$

It follows from (2.61) that the method may be extended to handle a finite number of higher derivatives as follows. Since z_0 satisfies (2.48), we obtain, using (2.60) and (2.61),

$$|f'(z_0)| \geq (1-o(1))\left(\frac{N}{r}\right)2\lambda M(r,f) \geq (2\lambda - o(1))M(r,f').$$

If $\tau < \gamma' < \gamma$ then, provided r lies outside a set of finite logarithmic measure, we have $\nu(r,f') \sim N$ by (2.26) and

$$\frac{f''(z)}{f'(z)} \sim \frac{\nu(r,f')}{z} \sim \frac{N}{z}$$

for $|\log(z/z_0)| \leq \nu(r,f')^{-\gamma'}$ and hence for $|\log(z/z_0)| \leq N^{-\gamma}$. Similarly, for these r and for ρ satisfying (2.42) we get $M(\rho, f'') \sim (N/\rho)M(\rho, f')$, and the whole process may be repeated a finite number of times.

Thus we have proved:

2.2.20 The main theorem of the Wiman-Valiron theory

Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be a transcendental entire function, and let $1/2 < \gamma < 1$ and $0 < \kappa \leq 1$. Let q be a positive integer. Then there exists a set $E_2 \subseteq [1\infty)$, of finite logarithmic measure, such that, if $|z_0| = r \in [1, \infty) \setminus E_2$ and $|f(z_0)| \geq \kappa M(r, f)$ then

$$f(z) \sim \left(\frac{z}{z_0}\right)^N f(z_0) \quad \text{and} \quad \frac{f^{(j)}(z)}{f(z)} \sim \frac{N^j}{z^j} \qquad \text{for } |\log(z/z_0)| \leq N^{-\gamma}$$

and $j = 1, \ldots, q$, where $N = \nu(r, f)$. Furthermore, for $|\log(\rho/r)| \leq N^{-\gamma}$ we have

$$M(\rho, f^{(j)}) \sim \frac{N^j}{\rho^j} M(\rho, f), \quad M(\rho, f) \sim \left(\frac{\rho}{r}\right)^N M(r, f)$$

for $j = 1, \ldots, q$.

The condition on γ is essentially best-possible. The Weierstrass σ-function has zeros at the points $m + n\omega$, where ω is a fixed non-real complex number and m and n are any integers. This function has order 2, and therefore so has $N(r)$. Now on the region $|\log(z/z_0)| \leq N^{-\gamma}$ we may write

$$z = z_0 e^{\zeta}, \quad |\zeta| \leq N^{-\gamma}, \quad |z - z_0| = |z_0|\,|e^{\zeta} - 1| \sim |z_0|\,|\zeta|,$$

and so this region has diameter roughly $rN(r)^{-\gamma}$. If it were possible to take $\gamma < 1/2$ then this diameter would be large, and our Wiman-Valiron region would contain a disc of centre z_0 and large radius compared to $1 + |\omega|$. But such a disc must contain a zero of the σ-function.

2.3 Exercises

1. Let f be a transcendental entire function. Prove that

$$\max\{\mathrm{Re}\, f(z) : |z| = r\} \sim M(r, f)$$

as $r \to \infty$ outside a set of finite logarithmic measure.

2. Prove that every non-constant solution of

$$y^{(4)} + zy' - z^4 y = 0$$

has order 2 (every solution is entire: see the chapter on differential equations).

3. Let P and Q be non-constant polynomials. Prove that the differential equation

$$2yy'' - (y')^2 + P(z)y + Q(z) = 0$$

has no transcendental entire solutions.

2.4 Coefficients and the order of growth

Let $g(z) = \sum_{n=0}^{\infty} b_n z^n$ be a transcendental entire function. We may then prove that the order of g is

$$\rho = \frac{1}{\sigma}, \quad \text{where} \quad \sigma = \liminf_{n\to\infty} \frac{-\log|b_n|}{n \log n},$$

with the convention that $1/0 = \infty$.

(i) To prove that $\rho \le 1/\sigma$ assume WLOG that $\sigma > \beta > 0$. Hence

$$-\log|b_n| \ge \beta n \log n, \quad |b_n| \le n^{-\beta n},$$

for all sufficiently large n.

Let r be large. Evidently if $n^\beta \ge r$ then n is large and

$$|b_n| r^n \le n^{-\beta n} r^n \le 1.$$

Hence

$$\mu(r) \le \max_{n \le r^{1/\beta}} |b_n| r^n.$$

But we can assume WLOG that $|b_n| \le 1$ for all n (why?) and so

$$\mu(r) \le r^{r^{1/\beta}} = \exp(r^{1/\beta} \log r),$$

which gives $\rho \le 1/\beta$. Fill in the details.

(ii) To prove that $\rho \ge 1/\sigma$ assume WLOG that $\rho < \tau < \infty$. Let n be large and $r = n^{1/\tau}$. Then r is large and

$$|b_n| r^n \le \mu(r) \le \exp(r^\tau) = e^n$$

which gives

$$\log|b_n| \le n - n\log r = n - \frac{n\log n}{\tau} = -(1 + o(1))\frac{n\log n}{\tau}$$

and so

$$\frac{-\log|b_n|}{n\log n} \ge \frac{1}{\tau} - o(1)$$

as $n \to \infty$. Again fill in the details.

Chapter 3

Nevanlinna theory

3.1 Introduction

The standard reference for this is Hayman's text [33], but this chapter will borrow several ideas from the excellent book by Jank and Volkmann [48].

A meromorphic function is one analytic function divided by another i.e. $f = g/h$, where g and h are analytic, and $h \not\equiv 0$. A good example is $f(z) = \tan z$, which has poles (i.e. $f(z) = \infty$) wherever $\cos z = 0$.

The *multiplicity* (or order) is defined as follows. Suppose g is analytic at a, with $g(a) = 0$. If $g \not\equiv 0$, then the Taylor series of g about a has a first non-zero coefficient, say

$$g(z) = a_m(z-a)^m + a_{m+1}(z-a)^{m+1} + ..., \quad a_j = \frac{g^{(j)}(a)}{j!}, \quad a_m \neq 0.$$

We say that g has a zero of multiplicity m at a. If $g(a) \neq 0$, we can think of this as a zero of multiplicity 0. Now consider g/h. If

$$g(z) = a_m(z-a)^m + \ldots, \quad h(z) = b_n(z-a)^n + \ldots,$$

as $z \to a$, with $a_m b_n \neq 0$, then

$$f(z) = \frac{g(z)}{h(z)} = (z-a)^{m-n}\left(\frac{a_m + \ldots}{b_n + \ldots}\right) = (z-a)^{m-n}H(z), \quad H(a) = \frac{a_m}{b_n}$$

near a. Here H is analytic at a. If $m > n$ then $f(a) = 0$ (zero of multiplicity $m-n$). If $m < n$ then $f(a) = \infty$ (pole of multiplicity $n-m$).

Example: show that

$$f(z) = \frac{z}{\sin^2 z}$$

has a simple pole at 0 and double poles at $z = k\pi$, $k \in \mathbb{Z} \setminus \{0\}$.

We have seen that the non-decreasing function $\log^+ M(r,f)$ measures the growth of an entire function f. The central idea of Nevanlinna theory is to develop an analogue for meromorphic functions, and to this end Nevanlinna introduced his characteristic function $T(r,f)$.

3.2 Nevanlinna theory: the first steps

We begin with:

3.2.1 Poisson's formula for the logarithm

Let $0 < R < \infty$ and let $E = \{z \in \mathbb{C} : |z| \leq R\}$. Let g be meromorphic on a domain containing E, with no zeros or poles in $D(0,R)$. Let the distinct zeros and poles of g on the circle $S(0,R)$ be $\zeta_1, \ldots, \zeta_q$. Then an analytic branch U of $\log g$ may be defined on a simply connected domain containing $E \setminus \{\zeta_1, \ldots, \zeta_q\}$ and, for $|a| < R$,

$$U(a) = \frac{1}{2\pi} \int_0^{2\pi} U(Re^{i\phi}) \frac{R^2 - |a|^2}{|Re^{i\phi} - a|^2} d\phi. \tag{3.1}$$

Proof. The first assertion is true since there exists $R' > R$ such that g is meromorphic in $D(0,R')$ with no zeros or poles in $R < |z| < R'$. Now let $|a| < R$. Let δ be small and positive and let Γ_δ be the circle $S(0,R)$ described once counter-clockwise, except that each ζ_j (if there are any) is avoided by instead describing clockwise an arc ω_j of the circle $S(\zeta_j, \delta)$. The resulting curve Γ_δ then goes once counter-clockwise around a, since δ is small. Set

$$V(w) = U(w) \left(\frac{R^2 - |a|^2}{R^2 - \overline{a}w} \right).$$

Then V is analytic on and inside Γ_δ and so Cauchy's integral formula gives

$$\begin{aligned} U(a) &= V(a) = \frac{1}{2\pi i} \int_{\Gamma_\delta} \frac{V(w)}{w-a} \, dw \\ &= \frac{1}{2\pi} \int_{\Gamma_\delta} U(w) \left(\frac{w}{w-a} \right) \left(\frac{R^2 - |a|^2}{R^2 - \overline{a}w} \right) \frac{dw}{iw}. \end{aligned} \tag{3.2}$$

But there exist non-zero constants a_j and integers m_j such that

$$g(w) \sim a_j (w - \zeta_j)^{m_j} \quad \text{and} \quad U(w) = \pm m_j \log \frac{1}{|w - \zeta_j|} + O(1) \quad \text{as} \quad w \to \zeta_j, \, w \in D(0,R).$$

In particular the argument of $g(w)$ remains bounded as $w \to \zeta_j$ in $D(0,R)$, and

$$U(w) = O\left(\log \frac{1}{\delta} \right)$$

on ω_j, for small δ. Hence the contribution to the integral in (3.2) from each circular arc ω_j tends to 0 as $\delta \to 0$, so that writing $w = Re^{i\phi}$ gives

$$\begin{aligned} U(a) &= \frac{1}{2\pi} \int_{S(0,R)} U(w) \left(\frac{w}{(w-a)} \right) \left(\frac{R^2 - |a|^2}{R^2 - \overline{a}w} \right) \frac{dw}{iw} \\ &= \frac{1}{2\pi} \int_0^{2\pi} U(w) \left(\frac{w}{w-a} \right) \left(\frac{R^2 - |a|^2}{\overline{w}w - \overline{a}w} \right) d\phi \\ &= \frac{1}{2\pi} \int_0^{2\pi} U(w) \left(\frac{1}{w-a} \right) \left(\frac{R^2 - |a|^2}{\overline{w} - \overline{a}} \right) d\phi, \end{aligned}$$

and (3.1) follows.

3.2.2 The Poisson-Jensen formula

Let R be finite and positive and let f be meromorphic and not identically zero in $|z| \leq R$. Let the zeros and poles of f in $0 < |z| < R$ be $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ respectively, in each case with repetition according to multiplicity. Assume that near the origin $f(z)$ is given by

$$f(z) = cz^d(1 + o(1)) \quad \text{as} \quad z \to 0,$$

with d an integer and c a non-zero constant: this says that cz^d is the first term of the Laurent series of f valid in some annulus $0 < |z| < s_0$. Then

$$g(z) = f(z)\frac{R^d}{z^d}\prod_{j=1}^{m}\left(\frac{R(z-a_j)}{R^2-\overline{a_j}z}\right)^{-1}\prod_{k=1}^{n}\left(\frac{R(z-b_k)}{R^2-\overline{b_k}z}\right) \tag{3.3}$$

is meromorphic on $|z| \leq R$, and analytic and non-zero in $|z| < R$. Moreover, $|g(z)| = |f(z)|$ on $|z| = R$. Taking real parts in the Poisson formula 3.2.1 gives, for $u = \log|g|$ and $z = re^{i\theta}$ with θ real and $0 \leq r < R$,

$$u(re^{i\theta}) = \frac{1}{2\pi}\int_0^{2\pi} u(Re^{i\phi})\frac{R^2-r^2}{R^2+r^2-2Rr\cos(\theta-\phi)}d\phi. \tag{3.4}$$

But for $|w| = R$ we have $u(w) = \log|g(w)| = \log|f(w)|$, and using (3.3) this gives the *Poisson-Jensen formula*: if $z = re^{i\theta}, |z| < R$ and $f(z) \neq 0, \infty$ then

$$\log|f(z)| = \frac{1}{2\pi}\int_0^{2\pi}\log|f(Re^{i\phi})|\frac{R^2-r^2}{R^2+r^2-2Rr\cos(\theta-\phi)}d\phi + d\log|z/R| \quad +$$

$$+\sum_{j=1}^{m}\log\left|\frac{R(z-a_j)}{R^2-\overline{a_j}z}\right| - \sum_{k=1}^{n}\log\left|\frac{R(z-b_k)}{R^2-\overline{b_k}z}\right|. \tag{3.5}$$

Here the a_j and b_k are the zeros and poles of f in $0 < |z| < R$. In particular, letting $z \to 0$ we have Jensen's formula

$$\log|c| = \frac{1}{2\pi}\int_0^{2\pi}\log|f(Re^{i\phi})|d\phi + \sum_{j=1}^{m}\log\frac{|a_j|}{R} - \sum_{k=1}^{n}\log\frac{|b_k|}{R} - d\log R. \tag{3.6}$$

Of course, $c = f(0)$ if $f(0) \neq 0, \infty$.

3.2.3 The Nevanlinna functionals

We retain the notation used in the Poisson-Jensen formula. Let $n(r) = n(r, f)$ denote the number of poles of f in $|z| \leq r$, counting multiplicity, and let $\mu(t) = n(t) - n(0)$. Then, using Lemma 1.3.2 and the integration by parts formula (1.6) for Riemann-Stieltjes integrals we obtain

$$\sum_{k=1}^{n}\log\frac{R}{|b_k|} = \int_0^R \log\frac{R}{t}\,d\mu(t) = -\int_0^R (n(t)-n(0))\,d\left(\log\frac{R}{t}\right) = \int_0^R (n(t)-n(0))\,\frac{dt}{t}. \tag{3.7}$$

Here the first formula follows by writing the sum as a Riemann-Stieltjes integral as in Lemma 1.3.2. Alternatively, we can prove by elementary means that

$$\sum_{k=1}^{n}\log\frac{R}{|b_k|} = \int_0^R (n(t)-n(0))\frac{dt}{t}. \tag{3.8}$$

Indeed, if f has p poles on $|z| = \rho$ these contribute p to $n(t) - n(0)$ for $\rho \leq t \leq R$ and so $p \log R/\rho$ to the integral and this gives us (3.8).

Now write

$$N(R, f) = \int_0^R (n(t, f) - n(0, f)) \frac{dt}{t} + n(0, f) \log R \tag{3.9}$$

and

$$m(R, f) = \frac{1}{2\pi} \int_0^{2\pi} \log^+ |f(Re^{i\phi})| d\phi, \tag{3.10}$$

where $\log^+ x$ is defined by (1.4) and satisfies

$$\log x = \log^+ x - \log^+ \frac{1}{x}, \quad x > 0. \tag{3.11}$$

Using (3.7), (3.9), (3.10) and (3.11), Jensen's formula (3.6) becomes

$$\log |c| = m(R, f) + N(R, f) - m(R, 1/f) - N(R, 1/f). \tag{3.12}$$

Here $m(R, f)$ is called the proximity function (Schmiegungsfunktion) and $N(R, f)$ the (integrated) counting function (Anzahlfunktion). The Nevanlinna characteristic is

$$T(R, f) = m(R, f) + N(R, f), \tag{3.13}$$

and the Jensen formula (3.12) can now be written

$$\log |c| = T(R, f) - T(R, 1/f). \tag{3.14}$$

3.2.4 Examples

(i) Let $F(z) = P(z)/Q(z)$ be a rational function, in which P and Q are polynomials, of degrees p, q respectively, and with $Q \not\equiv 0$. We can assume that P and Q have no common zeros. Then $Q(z) = 0$ has q roots, counting multiplicities, and so

$$N(r, F) = q \log r + O(1)$$

for large r. Also, as $z \to \infty$ we have $F(z) = dz^{p-q}(1 + o(1))$ for some constant $d \neq 0$, and so

$$\log |F(z)| = (p - q) \log |z| + O(1), \quad z \to \infty,$$

from which

$$m(r, F) = \max\{(p - q), 0\} \log r + O(1), \quad r \to \infty.$$

This gives

$$T(r, F) = \max\{p, q\} \log r + O(1), \quad r \to \infty.$$

(ii) Let $f(z) = e^z$. Show that $T(r, f) = m(r, f) = r/\pi$ for $r > 0$.

(iii) Show that

$$\log^+ |\cos z| = |\mathrm{Im}\, z| + O(1),$$

by considering separately the cases where $|\mathrm{Im}\, z|$ is or is not at least 100. Deduce that

$$T(r, \cos z) = 2r/\pi + O(1)$$

as $r \to \infty$. Illustrate Jensen's formula by estimating $m(r, \sec z)$ and $N(r, \sec z)$.

(iv) Let f be meromorphic in the plane and, with k a positive integer, define $g(z) = f(z^k)$. Prove that

$$n(r,g) = kn(r^k, f), \quad N(r,g) = N(r^k, f), \quad m(r,g) = m(r^k, f), \quad T(r,g) = T(r^k, f).$$

Show also that $T(r, f^k) = kT(r,f)$ and that, if a is a non-zero constant and $f(0) \neq \infty$, then $T(r, f(az)) = T(|a|r, f)$.

(v) Show that if $P(z) = az^k + \ldots$ is a polynomial of degree k then $T(r, e^P) \sim |a|r^k/\pi$ as $r \to \infty$ (hint: consider first the case $P(z) = z^k$).

3.2.5 Properties of the characteristic

Suppose that f, f_1, f_2 are meromorphic and non-constant. Then

$$T(R, f_1 f_2) \leq T(R, f_1) + T(R, f_2), \quad T(R, f_1 + f_2) \leq T(R, f_1) + T(R, f_2) + \log 2. \tag{3.15}$$

These follow easily from the inequalities

$$\log^+ xy \leq \log^+ x + \log^+ y, \quad \log^+(x+y) \leq \log^+(2\max\{x,y\}) \leq \log^+ x + \log^+ y + \log 2, \quad x, y > 0,$$

and the fact that a pole of $f_1 f_2$ or $f_1 + f_2$ can only arise at a pole of f_1 or f_2, and has multiplicity not greater than the sum of the multiplicities for f_1 and f_2.

3.2.6 Comparing $T(r,f)$ and $\log M(r,f)$

Let f be analytic in $|z| \leq R$. If $0 < r < R$ then

$$T(R,f) \leq \log^+ M(R,f), \quad \log M(r,f) \leq \left(\frac{R+r}{R-r}\right) T(R,f).$$

The first inequality is obvious, since $\log^+ |f(z)| \leq \log^+ M(R,f)$ on $|z| = R$. To prove the second, we take z with $|z| = r$ and $|f(z)| = M(r,f)$, and we apply the Poisson-Jensen formula, using the fact that the contribution from the zeros of f is non-positive, and the inequality

$$R^2 + r^2 - 2Rr\cos t \geq (R-r)^2, \quad R > r \geq 0, \quad t \in \mathbb{R}.$$

This relation shows that for entire functions $T(r,f)$ and $\log^+ M(r,f)$ are comparable.

3.2.7 A useful inequality

If $0 < r < R$ then

$$\begin{aligned} N(R,f) &= \int_0^R (n(t,f) - n(0,f)) \frac{dt}{t} + n(0,f) \log R \\ &\geq \int_r^R (n(t,f) - n(0,f)) \frac{dt}{t} + n(0,f) \log R \\ &\geq \int_r^R (n(r,f) - n(0,f)) \frac{dt}{t} + n(0,f) \log R \\ &= (n(r,f) - n(0,f)) \log \frac{R}{r} + n(0,f) \log R \\ &= n(r,f) \log \frac{R}{r} + n(0,f) \log r. \end{aligned}$$

3.2.8 Lemma

Let f be meromorphic in $\mathbb{C}$, and not a rational function. Then

$$\frac{T(r,f)}{\log r} \to \infty$$

as $r \to \infty$.

Proof. Note that we saw in Examples 3.2.4, ((i) that if f is a rational function then $T(r,f) = O(\log r)$ as $r \to \infty$.

Suppose then that f is meromorphic and non-constant in the plane, and that $T(r_n, f) = O(\log r_n)$ through some sequence $r_n \to \infty$. Now the inequality 3.2.7 gives, with $r^2 = r_n$ and r_n large,

$$C \log r > T(r^2, f) \geq N(r^2, f) \geq n(r, f) \log r$$

so f has finitely many poles. Hence there exists a polynomial P such that $g = Pf$ is entire, and $T(r_n, g) = O(\log r_n)$. Hence §3.2.6 gives

$$\log M(s_n, g) \leq 3T(2s_n, g) \leq C_1 \log r_n, \quad s_n = r_n/2,$$

so there exists an integer $M > 0$ such that $|g(z)| \leq (s_n)^M$ on the circles $|z| = s_n \to \infty$. Thus Cauchy's integral formula shows us that $g^{(M)}$ is bounded and so constant, and g is a polynomial.

3.2.9 An alternative proof of Jensen's formula

Let the function f be meromorphic in $|z| < R$ and for simplicity assume that $f(0) \neq 0, \infty$. For $0 \leq r < R$ set

$$I(r) = \frac{1}{2\pi} \int_0^{2\pi} \log |f(re^{i\theta}| \, d\theta.$$

Here $I(0) = \log |f(0)|$ and it is not hard to see that $I(r)$ is continuous. Now suppose that f has neither zeros nor poles on the circle $|z| = s \in (0, R)$. Then setting $\tau = \log |z|$ and writing $\log f$ locally as a function of $\tau + i\theta$ gives

$$sI'(s) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial \log |f|}{\partial \tau} (se^{i\theta}) \, d\theta = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial \arg f}{\partial \theta} (se^{i\theta}) \, d\theta = n(s, 1/f) - n(s, f).$$

Dividing by s and integrating from 0 to r then yields

$$m(r,f) - m(r,1/f) = I(r) = I(0) + N(r,1/f) - N(r,f) = \log |f(0)| + N(r,1/f) - N(r,f).$$

3.3 Nevanlinna's first fundamental theorem

3.3.1 First fundamental theorem

For non-constant meromorphic f and $a \in \mathbb{C}$ we have

$$m(R, 1/(f-a)) + N(R, 1/(f-a)) = T(R, f) + O(1). \tag{3.16}$$

For if a is a finite complex number we have by (3.15), as $R \to \infty$,

$$T(R, f-a) \leq T(R,f) + O(1), \quad T(R,f) \leq T(R, f-a) + O(1)$$

and so $T(R,f) = T(R,f-a) + O(1)$.

This is an equidistribution theorem: if f is meromorphic and non-constant in $\mathbb{C}$ then by Example 3.2.4 (i) and Lemma 3.2.8 the characteristic $T(R,f)$ tends to infinity as R tends to infinity. Hence either f takes the value a very often (so that N is large) or f is close to a on part of the circle $|z| = R$.

A good example is $f(z) = e^z$. Then $m(r,f) = m(r,1/f) = r/\pi$, while $N(r,1/f) = N(r,f) = 0$. Also $m(r,1/(f-1))$ is small, but f has a lot of 1-points.

For brevity we write

$$m(r,1/(f-a)) = m(r,a,f) = m(r,a), \quad N(r,1/(f-a)) = N(r,a,f) = N(r,a). \tag{3.17}$$

Also $m(r,f) = m(r,\infty), N(r,f) = N(r,\infty)$.

3.3.2 More examples

(i) Show that if T is a Möbius transformation and $g = T(f)$ then

$$T(r,g) = T(r,f) + O(1), \quad r \to \infty.$$

Deduce that $T(r,\tan z) = (2r/\pi) + O(1)$ (Hint: write $\tan z$ in terms of e^{2iz}).

Illustrate the first fundamental theorem by looking at $m(r,\tan z), N(r,\tan z), N(r,1/\tan z)$.

(ii) Show that $f(z) = e^{2z} - e^z$ has, as $r \to \infty$,

$$\begin{aligned} N(r,\infty) &= N(r,f) = 0, \\ m(r,\infty) &= m(r,f) \sim m(r,e^{2z}) = \frac{2r}{\pi}, \\ N(r,0) &= N(r,0,e^z-1) = \frac{r}{\pi} + O(1), \\ m(r,0) &\sim m(r,e^{-z}) = \frac{r}{\pi}, \\ m(r,a) &= O(1), \quad N(r,a) = \frac{2r}{\pi} + O(1), \quad (a \in \mathbb{C} \setminus \{0\}). \end{aligned}$$

3.3.3 An application of the first fundamental theorem: a lemma of Clunie

Let f be transcendental meromorphic and let g be entire. Then

$$T(r,g) = o(T(r,f\circ g)) \quad \text{as} \quad r \to \infty.$$

Proof. Choose $a \in \mathbb{C}$ such that a is not a critical value of $h = f \circ g$ and f has infinitely many a-points $w_1, w_2, \ldots$. Fix $N \in \mathbb{N}$ and choose $C, \delta > 0$ such that

$$|w - w_j| < \delta \quad \text{implies that} \quad |f(w) - a| \le C|w - w_j| \quad (j = 1, \ldots, N).$$

This gives

$$\sum_{j=1}^{N} m(r,w_j,g) \le m(r,a,h) + O(1)$$

and

$$\sum_{j=1}^{N} N(r,w_j,g) \le N(r,a,h).$$

Adding and applying the first fundamental theorem yields

$$\sum_{j=1}^{N} T(r, w_j, g) \leq T(r, a, h) + O(1), \quad NT(r, g) \leq T(r, h) + O(1),$$

so that $T(r, g) = o(T(r, f \circ g))$.

3.4 Cartan's formula and the growth of the characteristic function

3.4.1 Cartan's formula

We saw earlier that if f is entire then $\log^+ M(r, f)$ is a non-decreasing function, and the aim of this section is to show that $T(r, f)$ is non-decreasing.

Let f be non-constant and meromorphic in $|z| < R$, with $f(0)$ finite. Let $r \in (0, R)$ and assume that the number of points on $|z| = r$ at which $|f(z)| = 1$ is finite (in particular, this will always be true unless f is a rational function: see Lemma 3.9.1). Now Jensen's formula applied to the function $a - z$ gives

$$\frac{1}{2\pi} \int_0^{2\pi} \log |a - e^{is}| ds = \log^+ |a| \tag{3.18}$$

for a complex number a. Thus

$$m(r, f) = \frac{1}{2\pi} \int_0^{2\pi} \log^+ |f(re^{it})| dt = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \log |f(re^{it}) - e^{is}| ds dt. \tag{3.19}$$

Let

$$\phi(s, t) = \log |f(re^{it}) - e^{is}|, \quad \phi^+(s, t) = \max\{\phi(s, t), 0\}, \quad \phi^-(s, t) = \max\{-\phi(s, t), 0\}.$$

Then $\phi = \phi^+ - \phi^-$. Also, the Fubini-Tonelli theorem gives

$$I_1 = \int_0^{2\pi} \int_0^{2\pi} \phi^+(s, t) ds dt = \int_0^{2\pi} \int_0^{2\pi} \phi^+(s, t) dt ds$$

and

$$I_2 = \int_0^{2\pi} \int_0^{2\pi} \phi^-(s, t) ds dt = \int_0^{2\pi} \int_0^{2\pi} \phi^-(s, t) dt ds.$$

But, by (3.15),

$$0 \leq I_1 \leq \int_0^{2\pi} 2\pi (\log^+ |f(re^{it})| + \log 2) dt \leq 4\pi^2 (m(r, f) + \log 2).$$

Thus I_1 is finite. Also Jensen's formula gives, since $m(r, g) \leq T(r, g)$ and $f(0) - e^{is} \neq 0$ for almost all s,

$$\begin{aligned} I_2 &\leq 2\pi \int_0^{2\pi} m\left(r, \frac{1}{f - e^{is}}\right) ds \\ &\leq \int_0^{2\pi} 2\pi (T(r, f - e^{is}) - \log |f(0) - e^{is}|) ds \\ &\leq 4\pi^2 T(r, f) + 4\pi^2 \log 2 - 4\pi^2 \log^+ |f(0)| < \infty, \end{aligned}$$

using (3.15) and (3.18) again. Thus (3.19) and Jensen's formula give

$$\begin{aligned} m(r,f) &= \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \phi^+(s,t) - \phi^-(s,t) ds dt \\ &= \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log|f(re^{it}) - e^{is}| dt ds \\ &= \frac{1}{2\pi}\int_0^{2\pi} N(r,e^{is}) - N(r,f) + \log|f(0) - e^{is}| ds \end{aligned}$$

and so

$$m(r,f) = \log^+|f(0)| + \frac{1}{2\pi}\int_0^{2\pi} N(r,e^{is}) ds - N(r,f). \tag{3.20}$$

We thus obtain Cartan's formula: for $f(0)$ finite we have

$$T(r,f) = \log^+|f(0)| + \frac{1}{2\pi}\int_0^{2\pi} N(r,e^{is}) ds, \tag{3.21}$$

for r in $(0,R)$. To obtain an analogue of (3.21) when $f(0) = \infty$ we just apply (3.20) to $1/f$.

We proceed to differentiate (3.21). Let r be such that the equation $|f(z)| = 1$ has finitely many solutions z on $|z| = \rho$, for all ρ close to r (this is true for all but at most one r in $(0,R)$). Let $0 < s < r$. Then there exists a constant C_1 such that, for all ρ close to r we have

$$T(\rho,f) = C_1 + \frac{1}{2\pi}\int_0^{2\pi} N(\rho,e^{it}) - N(s,e^{it}) dt$$

and so

$$T(\rho,f) = C_1 + \frac{1}{2\pi}\int_0^{2\pi}\int_s^{\rho} n(r,e^{it})\frac{dr}{r} dt.$$

Since the integrand is non-negative we can reverse the order of integration to get

$$T(r,f) = C_1 + \frac{1}{2\pi}\int_s^{\rho}\int_0^{2\pi} n(r,e^{it}) dt \frac{dr}{r}.$$

But we saw above that

$$\frac{1}{2\pi}\int_0^{2\pi} n(s,e^{it}) dt$$

is continuous at r, and so

$$r\frac{dT}{dr} = \frac{1}{2\pi}\int_0^{2\pi} n(r,e^{it}) dt,$$

which is the differentiated Cartan formula.

In particular $T(r,f)$ is an increasing convex function of $\log r$ i.e.

$$P(s) = T(e^s,f)$$

satisfies

$$P(a) \le P(s) \le P(a)\frac{(b-s)}{(b-a)} + P(b)\frac{(s-a)}{(b-a)}, \quad a < s < b.$$

This is because $P'(t)$ is non-decreasing, so that

$$\frac{P(s) - P(a)}{s-a} = \frac{1}{(s-a)}\int_a^s P'(t) dt \le \frac{1}{(b-s)}\int_s^b P'(t) dt = \frac{P(b) - P(s)}{b-s}.$$

3.4.2 The order of a meromorphic function

If f is meromorphic on $\mathbb{C}$ we define the order $\rho(f)$ and lower order $\mu(f)$ by

$$\rho(f) = \limsup_{r\to\infty} \frac{\log^+ T(r,f)}{\log r}, \quad \mu(f) = \liminf_{r\to\infty} \frac{\log^+ T(r,f)}{\log r}.$$

We now have two definitions for the order of growth of an entire function h. However, since §3.2.6 gives

$$T(r,h) \leq \log^+ M(r,h) \leq 3T(2r,h),$$

Lemma 1.2.4 tells us that both give the same value ρ.

3.5 The logarithmic derivative

The key to Nevanlinna's methods is an estimate for $m(r, f'/f)$ when f is meromorphic. This leads to the second fundamental theorem, which is a strong generalization of Picard's theorem, and to a host of further results. The treatment here will follow the approach of Jank and Volkmann [48].

The Poisson formula (3.4) may be differentiated to give a formula for the derivative g'/g of $\log g$. Here we write $u(z) = \log|g(z)|$ as the real part of

$$I(z) = \left(\frac{1}{2\pi} \int_0^{2\pi} u(Re^{i\phi}) \frac{Re^{i\phi}+z}{Re^{i\phi}-z} d\phi \right).$$

Hence $\log g - I$ is constant on $|z| < R$. Writing f'/f in terms of g'/g and using the fact that $|f| = |g|$ on $|z| = R$ we obtain, for $|z| = r < R$,

$$\frac{f'(z)}{f(z)} = \frac{g'(z)}{g(z)} + \sum_{j=1}^{m} \left(\frac{\overline{a_j}}{R^2 - \overline{a_j}z} + \frac{1}{z-a_j} \right) - \sum_{k=1}^{n} \left(\frac{\overline{b_k}}{R^2 - \overline{b_k}z} + \frac{1}{z-b_k} \right) + \frac{d}{z},$$

and so

$$\frac{f'(z)}{f(z)} = \frac{1}{2\pi} \int_0^{2\pi} \log|f(Re^{i\phi})| \frac{2Re^{i\phi}}{(Re^{i\phi}-z)^2} d\phi \quad +$$
$$+ \sum_{j=1}^{m} \left(\frac{\overline{a_j}}{R^2 - \overline{a_j}z} + \frac{1}{z-a_j} \right) - \sum_{k=1}^{n} \left(\frac{\overline{b_k}}{R^2 - \overline{b_k}z} + \frac{1}{z-b_k} \right) + \frac{d}{z}. \tag{3.22}$$

Now for $|z| = r < R$ and $|A| \leq R$ we have

$$\frac{1}{z-A} + \frac{\overline{A}}{R^2 - \overline{A}z} = \frac{1}{z-A} \left(1 + \frac{\overline{A}(z-A)}{R^2 - \overline{A}z} \right). \tag{3.23}$$

Since $|A| \leq R$ and since

$$w = \frac{R(z-A)}{R^2 - \overline{A}z}$$

has modulus 1 when $|z| = R$, the term in parentheses in (3.23) has modulus at most 2. Using

$$|\log x| = \log^+ x + \log^+ \frac{1}{x}$$

we now get, for $|z| = r < R$,

$$\left| \frac{f'(z)}{f(z)} \right| \leq (m(R,f) + m(R,1/f)) \frac{2R}{(R-r)^2} + 2 \sum \left(\frac{1}{|z-A|} \right) + \frac{|d|}{r}, \tag{3.24}$$

with the sum over all zeros and poles A of f in $0 < |\zeta| < R$, repeated according to multiplicity.

This formula can be used to give pointwise estimates for f'/f (see §3.7). We will show that it leads to a very strong estimate for $m(r, f'/f)$.

3.5.1 Estimates for the proximity function of a logarithmic derivative

Let f be non-constant and meromorphic in $|z| \leq R$, and let $0 < r < R$, such that f has no zeros or poles on $|z| = r$. Set $S = (R+r)/2$. Assume for now that $f(0) \neq 0, \infty$, and replace R by S in (3.24), to give

$$1 + \left|\frac{f'(z)}{f(z)}\right| \leq 1 + \left[\frac{2S}{(S-r)^2}(m(S,f) + m(S,1/f))\right] + \sum\left(\frac{1}{|z-A|}\right) + \sum\left(\frac{1}{|z-A|}\right).$$

Here each sum is over all zeros and poles A of f in $0 < |\zeta| < S$, repeated according to multiplicity. Using the formula

$$\left(\sum_{k=1}^{n} x_k\right)^{1/2} \leq \sum_{k=1}^{n} x_k^{1/2}, \quad x_k \geq 0,$$

which is proved simply by squaring both sides, then yields

$$\left(1 + \left|\frac{f'(z)}{f(z)}\right|\right)^{1/2} \leq I(z,S) = 1 + \left[\frac{2S}{(S-r)^2}(m(S,f) + m(S,1/f))\right]^{1/2} + 2\sum \frac{1}{|z-A|^{1/2}}. \quad (3.25)$$

But, in view of the fact that $\log^+ x \leq \log(1+x)$ for $x \geq 0$, (3.25) gives

$$m(r, f'/f) \leq \frac{1}{2\pi}\int_0^{2\pi} \log\left(1 + \left|\frac{f'(re^{it})}{f(re^{it})}\right|\right) dt \leq \frac{2}{2\pi}\int_0^{2\pi} \log I(re^{it}, S)\, dt. \quad (3.26)$$

Now Lemma 1.3.4 and (3.26) lead to

$$m(r, f'/f) \leq 2\log X, \quad X = \frac{1}{2\pi}\int_0^{2\pi} I(re^{it}, S)\, dt. \quad (3.27)$$

Recalling (3.25) delivers next

$$X \leq 1 + \left[\frac{2S}{(S-r)^2}\left(2T(S,f) + \log^+\frac{1}{|f(0)|}\right)\right]^{1/2} + 2\sum I_A, \quad I_A = \frac{1}{2\pi}\int_0^{2\pi} |re^{it} - A|^{-1/2} dt. \quad (3.28)$$

To estimate I_A, we write

$$I_A = r^{-1/2} J_D, \quad J_D = \frac{1}{2\pi}\int_0^{2\pi} |e^{it} - D|^{-1/2} dt, \quad D = A/r. \quad (3.29)$$

To obtain an upper bound for J_D, there is no loss of generality in assuming that D is real and positive. Thus

$$J_D = \frac{1}{2\pi}\int_0^{2\pi} (1 + D^2 - 2D\cos t)^{-1/4} dt = (1+D^2)^{-1/4}\frac{1}{2\pi}\int_0^{2\pi} (1 - u\cos t)^{-1/4} dt,$$

in which

$$u = \frac{2D}{1+D^2} \leq 1.$$

This gives

$$J_D \leq \frac{1}{2\pi}\int_0^{2\pi} (1 - u\cos t)^{-1/4} dt \leq \frac{1}{2\pi}\int_0^{2\pi} (1 - |\cos t|)^{-1/4} dt = \gamma,$$

in which γ is some fixed positive number, independent of r, R and f.

Thus (3.28) and (3.29) combine to deliver

$$X \leq 1 + \left[\frac{2S}{(S-r)^2}\left(2T(S,f) + \log^+ \frac{1}{|f(0)|}\right)\right]^{1/2} + 2r^{-1/2}\gamma(n(S,f) + n(S,1/f)). \tag{3.30}$$

But the inequality from §3.2.7 gives

$$N(R,f) \geq n(S,f) \log \frac{R}{S}.$$

Since

$$\log \frac{R}{S} = \log\left(1 + \frac{R-S}{S}\right) \geq \frac{(R-S)/S}{1+(R-S)/S} = \frac{R-S}{R},$$

we get

$$n(S,f) + n(S,1/f) \leq \frac{R}{R-S}(2T(R,f) + \log^+ |1/f(0)|).$$

Thus (3.27) and (3.30) and the inequality

$$\log^+ \sum_{k=1}^{n} x_k \leq \log^+(n \max\{x_k\}) \leq \log n + \log^+(\max\{x_k\}) \leq \log n + \sum_{k=1}^{n} \log^+ x_k, \quad x_k > 0,$$

imply that there are positive absolute constants C_j such that

$$\begin{aligned} m(r,f'/f) \quad &\leq \quad C_1 + C_2 \log^+ T(R,f) + C_3 \log^+ \log^+ \frac{1}{|f(0)|} + \\ &\qquad + C_4 \log^+ R + C_5 \log^+ \frac{1}{r} + C_6 \log^+ \frac{1}{R-r}. \end{aligned} \tag{3.31}$$

An analogous formula when $f(0) = 0, \infty$ is easy to obtain. If $f(z) = cz^d(1 + o(1))$ as $z \to 0$, we write $f(z) = cz^d h(z)$ so that $h(0) = 1$. Now we need only use the fact that

$$\frac{f'(z)}{f(z)} = \frac{d}{z} + \frac{h'(z)}{h(z)}, \quad \left|\frac{f'(z)}{f(z)}\right| \leq \left|\frac{h'(z)}{h(z)}\right| + \left|\frac{d}{z}\right|$$

and

$$T(r,h) \leq T(r,f) + T(r,1/cz^d) \leq T(r,f) + d\log r + O(1).$$

3.5.2 The lemma of the logarithmic derivative

Let f be non-constant and meromorphic in the plane. Then there are positive constants C_j such that we have

$$m(r,f'/f) \leq C_1 \log r + C_2 \log T(r,f) \tag{3.32}$$

as r tends to ∞ outside a set of finite measure.

To prove this, choose $R = r + 1/T(r)$ in (3.31), and apply the Borel lemma 1.2.5.

Note that this estimate is only needed for transcendental f. If f is a rational function then $f'(z)/f(z) \to 0$ as $z \to \infty$ so $m(r,f'/f) = 0$ for large r.

If f has finite order we have $m(r,f'/f) = O(\log rT(2r,f)) = O(\log r)$ with no exceptional set (just take $R = 2r$).

We write $S(r,f)$ for any term which is $O(\log^+(rT(r,f)))$ outside some set E^* of finite measure. Note that if f is not a rational function then $S(r,f) = o(T(r,f))$ as $r \to \infty$ with $r \notin E^*$.

3.5.3 Theorem

We have $T(r, f') \leq T(r, f) + \overline{N}(r, f) + S(r, f)$.

Here $\overline{N}(r, f)$ counts poles of f, but without regard to multiplicity. The proof is easy. We have

$$N(r, f') \leq N(r, f) + \overline{N}(r, f), \quad m(r, f') \leq m(r, f) + m(r, f'/f).$$

In particular, if f has finite order, then

$$T(r, f') \leq 2T(r, f) + O(\log r). \tag{3.33}$$

3.5.4 Lemma

If f is non-constant and meromorphic in the plane, then $\rho(f') \leq \rho(f)$.

If $\rho(f) = \infty$, this is obvious. If $\rho(f) < \infty$, then we just use Lemma 1.2.4 and (3.33). In fact, the two orders are the same, but it is much harder to prove that $\rho(f) \leq \rho(f')$.

3.6 The second fundamental theorem

Let f be again non-constant and meromorphic in the plane, and let $a_1, \ldots, a_q$ be q distinct finite complex numbers. Let

$$H = \sum_{j=1}^{q} \frac{1}{f - a_j}. \tag{3.34}$$

Take a small positive ε, so small that $|w - a_j| < \varepsilon$ implies that $|w - a_k| > \varepsilon$ for $j \neq k$. If $|f(z) - a_j| < \varepsilon$ we then have

$$\frac{1}{|f(z) - a_j|} \leq |H(z)| + \frac{q-1}{\varepsilon}$$

and so

$$\log^+ \frac{1}{|f(z) - a_j|} \leq \log^+ |H(z)| + O(1),$$

while if $|f(z) - a_j| \geq \varepsilon$ then obviously

$$\log^+ \frac{1}{|f(z) - a_j|} \leq \log \frac{1}{\varepsilon}.$$

Since the sets $E_j = \{z : |f(z) - a_j| < \varepsilon\}$ are pairwise disjoint it follows that

$$\begin{aligned}
\sum_{j=1}^{q} m(r, a_j, f) &\leq \sum_{j=1}^{q} \left[\frac{1}{2\pi} \int_{[0,2\pi] \cap E_j} \log^+ \frac{1}{|f(re^{i\phi}) - a_j|} \, d\phi + \log \frac{1}{\varepsilon} \right] \\
&\leq \sum_{j=1}^{q} \frac{1}{2\pi} \int_{[0,2\pi] \cap E_j} \log^+ |H(re^{i\phi})| \, d\phi + O(1) \\
&\leq m(r, H) + O(1) = m(r, f'H/f') + O(1) \\
&\leq m(r, f'H) + m(r, 1/f') + O(1) \\
&\leq m(r, 1/f') + S(r, f),
\end{aligned}$$

since $f'H$ is a sum of logarithmic derivatives and $T(r, f - a_j) \leq T(r, f) + O(1)$. Here

$$\begin{aligned} m(r, 1/f') &= T(r, 1/f') - N(r, 1/f') \\ &= T(r, f') - N(r, 1/f') + O(1) \quad \text{(by Jensen's formula)} \\ &= m(r, f') + N(r, f') - N(r, 1/f') + O(1). \end{aligned}$$

Moreover,

$$\begin{aligned} m(r, f') &= m(r, f \cdot f'/f) \leq m(r, f) + m(r, f'/f) \\ &\leq m(r, f) + S(r, f). \end{aligned}$$

Also, since $\overline{N}(r, f)$ counts each pole exactly once, we have

$$\begin{aligned} N(r, f') &= N(r, f) + \overline{N}(r, f) \\ &= 2N(r, f) - [N(r, f) - \overline{N}(r, f)]. \end{aligned}$$

Thus

$$\begin{aligned} m(r, f) + \sum_{j=1}^{q} m(r, a_j, f) &\leq m(r, f) + m(r, 1/f') + S(r, f) \\ &\leq 2m(r, f) + N(r, f') - N(r, 1/f') + S(r, f) \\ &\leq 2m(r, f) + 2N(r, f) - [N(r, f) - \overline{N}(r, f)] - N(r, 1/f') + S(r, f) \\ &= 2T(r, f) + S(r, f) - N_1(r, f), \end{aligned} \tag{3.35}$$

in which

$$N_1(r, f) = N(r, f) - \overline{N}(r, f) + N(r, 1/f') \geq 0.$$

This term $N_1(r, f)$ counts the multiple points of f in the following sense. The function f is one-one on some neighbourhood of z_0 if and only if either $f(z_0)$ is finite and $f'(z_0) \neq 0$, or z_0 is a simple pole of f. Indeed, if $f(z)$ has an a-point (a finite or infinite) of multiplicity p at z_0 then by Rouché's theorem all values w which are sufficiently close to a are taken p times near to z_0. Thus z_0 is a multiple point of order $p - 1$, and contributes $p - 1$ to $n_1(r, f)$.

3.6.1 Statement of the second fundamental theorem

From (3.35) and the fact that $m(r, f) \geq 0$ we obtain the second fundamental theorem: given any s distinct values b_j in $\mathbb{C}^*$ (one of them is allowed to be ∞ here), we have

$$\sum_{j=1}^{s} m(r, b_j, f) \leq 2T(r, f) - N_1(r, f) + S(r, f). \tag{3.36}$$

Adding the terms $N(r, b_j, f)$ to both sides of (3.36) we get, by the first fundamental theorem,

$$(s - 2)T(r, f) \leq \sum_{j=1}^{s} N(r, b_j, f) - N_1(r, f) + S(r, f).$$

But if f has a b_j-point at a, of multiplicity p, then a contributes p to $n(r, b_j, f)$ and $p - 1$ to $n_1(r, f)$. Thus we get

$$(s - 2)T(r, f) \leq \sum_{j=1}^{s} \overline{N}(r, b_j, f) + S(r, f).$$

Picard's theorem is an immediate corollary. If f is transcendental and meromorphic in the plane and takes three values b_j each only finitely often, then $N(r, b_j, f) = O(\log r)$ for these b_j. Since $S(r, f) = o(T(r, f))$ as r tends to infinity outside a set E of finite measure, we deduce that $T(r, f) = O(\log r)$ for r not in E, a contradiction. This proves the "great" Picard theorem. It remains only to prove that if f omits three values then f is constant (this is the "little" theorem). However, this is easy: if f is a non-constant rational function $f = P/Q$ with P, Q polynomials having no common zero, then $Q = 0$ gives $f = \infty$, while the equation $P(z) = bQ(z)$ has solutions in $\mathbb{C}$, for all but at most one finite b.

3.6.2 The defect relation

Nevanlinna defined, for $a \in \mathbb{C}^*$, the *deficiency*

$$\delta(a, f) = \liminf_{r\to\infty} \frac{m(r, a, f)}{T(r, f)} = 1 - \limsup_{r\to\infty} \frac{N(r, a, f)}{T(r, f)} \tag{3.37}$$

as a measure of the extent to which the value a is taken rarely. The equality in (3.37) follows from the first fundamental theorem. From (3.37), we have $0 \leq \delta(a, f) \leq 1$. Also (3.36) gives the *defect relation*

$$\sum_{a\in\mathbb{C}^*} \delta(a, f) \leq 2. \tag{3.38}$$

3.6.3 Examples

(i) If a is an omitted value of f then $\delta(a, f) = 1$. Thus the defect relation (3.38) implies Picard's theorem.

(ii) A meromorphic function f can take a value a infinitely often, but still have $\delta(a, f) = 1$. For example,

$$f(z) = e^{z^2} \tan z$$

has $\delta(0, f) = \delta(\infty, f) = 1$, since

$$T(r, \tan z) = O(r), \quad T(r, e^{z^2}) = \frac{r^2}{\pi} \leq T(r, f) + T(r, \cot z).$$

(iii) Determine the Nevanlinna deficiencies of $e^{2z} - e^z$ (see Examples 3.3.2).

(iv) Here we give an example of an entire function f having two finite deficient values, each with deficiency $\frac{1}{2}$, and so sum of all deficiencies equal to 2. Set

$$f(z) = \int_0^z e^{-t^2} dt, \quad I = \int_0^\infty e^{-t^2} dt.$$

Here the integral I is over $[0, \infty)$, and in fact equals $\frac{1}{2}\sqrt{\pi}$, although all we require here is that $I \neq 0$. Suppose first that $|\arg z| < \pi/4$. Then Cauchy's theorem gives

$$f(z) = I - \int_{\gamma_z} e^{-t^2} dt,$$

in which γ_z follows the (shorter) circular arc from z to $r = |z|$, followed by the straight line from r to infinity. On γ_z we have

$$|e^{-t^2}| = e^{-|t|^2 \cos(2\arg t)} \leq |e^{-z^2}|.$$

We write

$$e^{-t^2} = \frac{2te^{-t^2}}{2t}$$

and integrate by parts. This gives

$$\int_{\gamma_z} e^{-t^2} dt = \frac{e^{-z^2}}{2z} - \int_{\gamma_z} \frac{e^{-t^2}}{2t^2} dt$$

and so, as $r = |z| \to \infty$,

$$\left| \int_{\gamma_z} e^{-t^2} dt \right| \leq |e^{-z^2}| \left(o(1) + \int_{\gamma_z} \frac{1}{2|t|^2} |dt| \right) \leq |e^{-z^2}|.$$

Thus

$$m(r, 1/(f - I)) \geq \frac{1}{2\pi} \int_{-\pi/4}^{\pi/4} r^2 \cos 2\theta \, d\theta = \frac{r^2}{2\pi}.$$

Since Taylor's theorem gives $f(z) = -f(-z)$ we also have

$$m(r, 1/(f + I)) \geq \frac{r^2}{2\pi}.$$

We now estimate $T(r, f) = m(r, f)$. For $|\arg z| \leq \pi/4$ or $|\arg(-z)| \leq \pi/4$ we have $f(z) = O(1)$. On the other hand if $\pi/4 < \arg z < 3\pi/4$ we have $|e^{-t^2}| \leq |e^{-z^2}|$ on the straight line from 0 to z, and so $|f(z)| \leq |ze^{-z^2}|$. Thus

$$T(r, f) \leq O(\log r) + \frac{1}{\pi} \int_{\pi/4}^{3\pi/4} (-r^2 \cos 2\theta) \, d\theta = \frac{r^2}{\pi}.$$

Exercise: generalize this to $g(z) = \int_0^z e^{-t^q} dt$, using the fact that $g(ze^{2\pi i/q}) = e^{2\pi i/q} g(z)$.

3.7 Pointwise estimates for logarithmic derivatives

3.7.1 Definition

By an R-set we mean a countable union U of discs $D(z_j, r_j)$ such that $z_j \to \infty$ as $j \to \infty$ and $\sum r_j < \infty$.

3.7.2 Lemma

Let U be an R-set. Let E be the set of $r > 0$ for which the circle $|z| = r$ meets at least one disc of U, and let H be the set of $\theta \in [0, 2\pi]$ such that the ray $\arg z = \theta$ meets infinitely many discs of U.

Then E has finite Lebesgue measure, and H has zero Lebesgue measure.

Proof. The first assertion is easy, since the set of $r > 0$ for which the circle $|z| = r$ meets $D(z_j, r_j)$ has measure at most $2r_j$.

Now suppose that j_0 is large, and $j \geq j_0$. Then z_j is large, and r_j is small, and the disc $D(z_j, r_j)$ subtends at the origin an angle at most $cr_j/|z_j|$, with c a positive constant independent of j and j_0. So the measure of H is at most

$$c \sum_{j=j_0}^{\infty} r_j/|z_j| \to 0 \quad \text{as} \quad j_0 \to \infty.$$

3.7.3 Lemma

Let f be a transcendental meromorphic function of order $\rho < L < M < \infty$. Let z_j be the zeros and poles of f in $|z| > 2$, repeated according to multiplicity. Then the union U of the discs $D(z_j, |z_j|^{-M})$ is an R-set, and

$$|f'(z)/f(z)| = o(|z|^{L+M})$$

for all z with $|z|$ large and $z \notin U$. Also,

$$\sum_{|z_j| \geq r/2} |z_j|^{-M} = o(r^{L-M}) \tag{3.39}$$

as $r \to \infty$.

Note that if f is a rational function, not identically zero, then $f'(z)/f(z) = O(1/|z|)$ as $z \to \infty$.

Proof. Let $m(t)$ be the number of z_j in $|z| \leq t$. Then $m(t) \leq n(t, f) + n(t, 1/f)$. For large t we have

$$T(t, f) = o(t^L)$$

and so

$$N(2t, f) + N(2t, 1/f) = o(t^L)$$

for large t. Lemma 3.2.7 now gives

$$m(t) \leq n(t, f) + n(t, 1/f) \leq o(t^L) \tag{3.40}$$

for t large.

We prove (3.39) first, which will then show that U is an R-set. For large r and $R > r$ we set $s = r/4$ and we have

$$\sum_{r/2 \leq |z_j| \leq R} |z_j|^{-M} \leq \int_s^{2R} t^{-M} d(m(t) - m(s)) = (m(2R) - m(s))(2R)^{-M} + M \int_s^{2R} (m(t) - m(s)) t^{-M-1} dt,$$

using integration by parts. Using (3.40) this gives

$$\sum_{r/2 \leq |z_j| \leq R} |z_j|^{-M} \leq m(2R)(2R)^{-M} + M \int_s^{2R} m(t) t^{-M-1} dt \leq o(R^{L-M}) + M \int_s^{2R} o(t^{L-M-1}) dt.$$

Letting $R \to \infty$ we get

$$\sum_{r/2 \leq |z_j|} |z_j|^{-M} \leq M \int_s^{\infty} o(t^{L-M-1}) dt = o(s^{L-M}) = o(r^{L-M}),$$

which proves (3.39).

To estimate f'/f, take $z \notin U$ with $|z| = r$ large, and use (3.24), with $R = 2r$. Since

$$m(2r, f) + m(2r, 1/f) \leq 2T(2r, f) + O(1)$$

we get

$$|f'(z)/f(z)| \leq o(r^L) + 2 \sum \frac{1}{|z - A|}$$

with the sum over all zeros and poles A of f in $0 < |\zeta| < 2r$. Now if $|A| < r/2$ then $|z - A| > r/2$. On the other hand, if $r/2 \leq |A| < 2r$ then A is one of the z_j and so $|z - A| \geq |A|^{-M} \geq (2r)^{-M}$. Hence

$$|f'(z)/f(z)| \leq o(r^L) + (n(2r, f) + n(2r, 1/f))(2r)^M = o(r^{L+M}).$$

3.7.4 Lemma

Let f be a transcendental meromorphic function of finite order $\rho < L < M$, and let n be a positive integer. Then we can find an R-set U of discs $D(z_j, |z_j|^{-M})$, such that for $|z|$ large and $z \notin U$ we have

$$|f^{(m+1)}(z)/f^{(m)}(z)| = o(|z|^{L+M}) \tag{3.41}$$

for $0 \leq m \leq n-1$.

Proof. By Lemma 3.5.4, each derivative $f^{(m)}$ has order at most ρ. So for each m we form an R-set U_j of discs $D(z_{j,m}, |z_{j,m}|^{-M})$ such that for $|z|$ large and z outside U_m we have (3.41). Now just note that the union U of these finitely many R-sets is an R-set.

By writing $f''/f = (f''/f')(f'/f)$ etc., we also have

$$|f^{(m)}(z)/f(z)| = o(|z|^{n(L+M)})$$

for $|z|$ large, with z not in U.

3.8 Product representations

Taylor's theorem tells us that an entire function f has a power series representation $f(z) = \sum_{n=0}^{\infty} c_n z^n$: here we show that functions meromorphic in $\mathbb{C}$ can be represented as products.

3.8.1 The exponent of convergence

Let (a_n) be a sequence of non-zero complex numbers, tending to infinity. For $r > 0$ let $n(r)$ be the number of a_n in $|z| \leq r$, and set

$$N(r) = \int_0^r n(t)\frac{dt}{t}.$$

The exponent of convergence of the sequence (a_n) is then defined as

$$\lambda = \limsup_{r\to\infty} \frac{\log N(r)}{\log r} = \limsup_{r\to\infty} \frac{\log n(r)}{\log r}. \tag{3.42}$$

The equality in (3.42) follows easily from Lemma 1.2.4 and the inequalities, for large r,

$$N(r) \leq n(r)\log r + O(1), \quad N(2r) \geq \int_r^{2r} n(r)\frac{dt}{t} \geq n(r)\log 2.$$

If $q > 0$ then, assuming without loss of generality that all the a_n are non-zero,

$$\sum_{|a_n|\leq r} |a_n|^{-q} = \int_0^r t^{-q} dn(t) = n(r)r^{-q} + q\int_0^r n(t)t^{-q-1}dt. \tag{3.43}$$

3.8.2 Lemma

The exponent of convergence λ is the infimum of $q > 0$ such that $\sum |a_n|^{-q}$ converges.

Proof. Suppose first that $\lambda < p < q < \infty$. Then $n(t) < t^p$ for all large positive t, and so

$$n(r)r^{-q} + q\int_0^r n(t)t^{-q-1}dt$$

tends to a finite limit as $r \to \infty$, which implies using (3.43) that $\sum |a_n|^{-q}$ converges. Conversely, suppose that $\sum |a_n|^{-q} = S < \infty$. Then for $r > 0$ we have $n(r) \leq Sr^q$ by (3.43) and so $\lambda \leq q$.

It is now clear that $\sum |a_n|^{-\mu}$ converges for $\lambda < \mu < \infty$ and diverges for $0 \leq \mu < \lambda$.

3.8.3 Weierstrass products

Define

$$E(z,0) = (1-z), \quad E(z,p) = (1-z)\exp\left(\sum_{j=1}^{p} \frac{z^j}{j}\right), \quad p \in \mathbb{N},$$

(the Weierstrass primary factors). Then for $|z| \leq \frac{1}{2}$ we have

$$|\log E(z,p)| = |-z^{p+1}(p+1)^{-1} - ....| \leq |z|^{p+1} + |z|^{p+2} + \cdots \leq 2|z|^{p+1}. \tag{3.44}$$

Next, for $|z| \geq 1$ and $p \geq 1$ we have

$$\log |E(z,p)| \leq \log(1+|z|) + |z| + \ldots + \frac{|z|^p}{p}$$

and so, for any p,

$$\log |E(z,p)| \leq \log(1+|z|) + p|z|^p, \quad |z| \geq 1. \tag{3.45}$$

Applying the maximum principle gives

$$\log |E(z,p)| \leq A(p) = p + \log 2, \quad |z| \leq 1. \tag{3.46}$$

3.8.4 Lemma

Let (a_n) be a non-zero sequence tending to infinity, and let $q_n \geq 0$ be integers such that for every positive r we have

$$\sum \left(\frac{r}{|a_n|}\right)^{q_n+1} < \infty. \tag{3.47}$$

Then

$$F(z) = \prod E(z/a_n, q_n)$$

converges, and is an entire function with zero sequence (a_n).

Proof. Fix $K > 0$. Then for $|z| \leq K$ we have, by (3.44) and (3.47),

$$\sum_{|a_n| \geq 2K} |\log E(z/a_n, q_n)| \leq \sum_{|a_n| \geq 2K} 2|K/a_n|^{q_n+1} < \infty.$$

Hence

$$\sum_{|a_n| \geq 2K} \log E(z/a_n, q_n)$$

converges absolutely and uniformly on $|z| \leq K$, and

$$F(z) = \exp\left(\sum_{|a_n| \geq 2K} \log E(z/a_n, q_n)\right) \prod_{|a_n| < 2K} E(z/a_n, q_n)$$

is analytic on $D(0,K)$.

3.8.5 Theorem

Suppose that the non-zero sequence (a_n) has finite exponent of convergence λ, and let q be the least integer such that $\sum |a_n|^{-q-1}$ converges. Then the product

$$F(z) = \prod E(z/a_n, q)$$

converges in $\mathbb{C}$, and has order λ. Further, we have $\log M(r, F) = o(r^{q+1})$ as $r \to \infty$.

Proof. We obviously have $\lambda \leq q + 1$, by definition of λ, and the fact that $\sum |a_n|^{-\mu}$ converges for every $\mu > \lambda$ gives $q + 1 \leq \lambda + 1$. We note next that replacing q by $q + 1$ in (3.43) and letting $r \to \infty$ leads to

$$\int_0^\infty n(t)t^{-q-2}dt < \infty \tag{3.48}$$

and so

$$n(R)\int_R^{2R} t^{-q-2}dt \leq \int_R^{2R} n(t)t^{-q-2}dt = o(1),$$

which gives

$$n(R) = o(R^{q+1}), \quad R \to \infty. \tag{3.49}$$

The product $F(z)$ converges since (3.47) is satisfied for every $r > 0$, with $q_n = q$, and it is obvious that F has order at least λ, by Jensen's formula. Now suppose that

$$q < s \leq q + 1, \quad \lim_{r\to\infty} \frac{n(r)}{r^s} = 0. \tag{3.50}$$

In particular, (3.50) is satisfied by $s = q + 1$, by (3.49). Let $|z| = r$ be large. Then

$$\log |F(z)| \leq \sum \log |E(z/a_n, q)|.$$

Splitting the sum into those over (i) $|a_n| \leq r$, (ii) $r < |a_n| < 2r$ and (iii) $2r \leq |a_n|$ respectively, and using (3.44), (3.45) and (3.46), we obtain

$$\log |F(z)| \leq S_1 + S_2 + S_3,$$

in which

$$S_1 = \sum_{|a_n|\leq r} (\log(1 + r/|a_n|) + q(r/|a_n|)^q),$$

$$S_2 = \sum_{r<|a_n|<2r} A(q) \leq A(q)n(2r) = o(r^s)$$

by (3.50), and

$$S_3 = \sum_{|a_n|\geq 2r} 2|r/a_n|^{q+1}.$$

Now

$$S_1 = \int_0^r (\log(1 + r/t) + q(r/t)^q)dn(t) = n(r)(\log 2 + q) + \int_0^r (r/t(t + r) + q^2r^q/t^{q+1})n(t)dt$$

and so

$$S_1 < n(r)(\log 2 + q) + N(r) + q^2r^q \int_0^r n(t)/t^{q+1}dt.$$

Hence (3.50) gives

$$S_1 < o(r^s) + O(r^q) + q^2 r^q \int_0^r o(t^{s-q-1})dt = o(r^s),$$

using the fact that $s - q - 1 > -1$. Next, integration by parts and (3.49) give

$$S_3 = \int_{2r}^{\infty} 2(r/t)^{q+1} dn(t) \leq 2(q+1)r^{q+1} \int_{2r}^{\infty} n(t)t^{-q-2}dt. \tag{3.51}$$

Thus

$$S_3 = o(r^s) :$$

to see this, if $s = q + 1$ we use (3.48), which tells us that the integral from $2r$ to ∞ tends to 0, while if $q < s < q + 1$, then we use (3.50) and substitute $n(t) = o(t^s)$ into the integral. Hence $\log M(r, F) = o(r^s)$, and so F has order at most s. It follows that F has order at most λ: this is obvious if $\lambda = q + 1$, while if $q \leq \lambda < q + 1$ we take s with $\lambda < s < q + 1$.

For any function $f \not\equiv 0$ meromorphic in $\mathbb{C}$, we now define $\lambda(f)$ to be the exponent of convergence of the zero sequence of f. Obviously this is the same as the order of $N(r, 1/f)$, and by Jensen's formula is not greater than the order of f. Similarly $\lambda(1/f)$ is the exponent of convergence of the zeros of $1/f$ and so poles of f.

3.8.6 Hadamard representation theorem

Let $f \not\equiv 0$ be meromorphic in $\mathbb{C}$. Then there exist entire functions F_1, F_2, h and an integer m such that $\rho(F_1) = \lambda(f)$ and $\rho(F_2) = \lambda(1/f)$ and $f(z) \equiv z^m \frac{F_1(z)}{F_2(z)} e^{h(z)}$.

Proof. Let (a_n) be the sequence of zeros of f in $0 < |z| < \infty$, and let (b_n) be the sequence of poles of f in $0 < |z| < \infty$, in both cases repeated according to multiplicity. Then there exist entire functions F_1, F_2, of orders $\lambda(f), \lambda(1/f)$ respectively, such that the zero sequence of F_1 is (a_n), and that of F_2 is (b_n) (if either of these sequences is finite then F_j is a finite product, while if the sequence is empty we put $F_j = 1$). We then choose an integer m so that $f(z)z^{-m}F_2(z)F_1(z)^{-1} = g(z)$ is analytic and non-zero at 0, and it follows that g is analytic and non-zero in the plane, since all singularities of g and $1/g$ have been removed. Thus we may write $g = e^h$ with h entire.

3.9 Appendix: lemmas underlying the Cartan formula

Cartan's formula was derived in §3.4.1, and the following lemmas serve to show that certain quantities are in fact measurable functions.

3.9.1 Lemma

Let $0 < r < R$ and let f be a function non-constant and meromorphic on $D(0, R)$. Assume that the circle $|z| = r$ contains infinitely many points z with $|f(z)| = 1$. Then f is a rational function.

Proof. Let $S = \{s \in \mathbb{R} : |f(re^{is})| = 1\}$ and let T be the set of t in $\mathbb{R}$ such that t is a limit point of S.

Suppose that $t_0 \in T$. Then we can find $t_n \to t_0, n \to \infty$, with t_n real, $t_n \neq t_0$ and $|f(re^{it_n})| = 1$. Obviously $|f(re^{it_0})| = 1$, by continuity. For z near re^{it_0}, put

$$g(z) = \log f(z), \quad u = i \log z, \quad h(u) = g(z) = \log f(e^{-iu}).$$

Let $u_0 = -t_0 + i \log r$. Taylor's theorem allows us to write

$$h(u) = \sum_{n=0}^{\infty} a_n (u - u_0)^n,$$

with the power series absolutely convergent on an open disc D centred at u_0. Let

$$H(u) = \sum_{n=0}^{\infty} \mathrm{Re}(a_n)(u - u_0)^n,$$

so that H is analytic on D. Setting $u_n = -t_n + i \log r$ we see that $u_n - u_0$ is real and $H(u_n) = \mathrm{Re}(h(u_n)) = \log |f(re^{it_n})| = 0$, and so $H(u) \equiv 0$ on D, by the identity theorem. So if s is real and close to t_0 then $H(-s + i \log r) = \mathrm{Re}(h(-s + i \log r)) = 0$.

It follows that if $t \in T$ then there exists $\delta_t > 0$ such that $|f(re^{is})| = 1$ for $t - \delta_t < s < t + \delta_t$, and so T is open.

Now suppose that v is real, but not in T. Then v is not a limit point of S, and so there exists $\rho_v > 0$ such that $|f(re^{is})| \neq 1$ for $v - \rho_v < s < v$ and $v < s < v + \rho_v$. So no t in the interval $(v - \rho_v, v + \rho_v)$ is a limit point of S, and so $\mathbb{R} \setminus T$ is open.

But $\mathbb{R}$ is connected, and T is non-empty, since $S \cap [0, 2\pi]$ is infinite by hypothesis, so that S has a limit point in the compact set $[0, 2\pi]$. Thus we see that $\mathbb{R} = T$.

We have now proved that $|f(z)| \equiv 1$ on the circle $|z| = r$. Let a_μ be the zeros of f in $|z| < r$, and b_ν the poles of f in $|z| < r$, in both cases repeated according to multiplicity. For $|a| < r$ we have

$$|U_a(z)| = 1, \quad |z| = r, \quad U_a(z) = \frac{r(z - a)}{r^2 - \overline{a} z},$$

in which U_a is a Möbius transformation with a zero at a and a pole at $r^2/\overline{a}$ (except that $U_a(z) = z/r$ if $a = 0$). Let

$$F(z) = f(z) \prod_\mu U_{a_\mu}(z)^{-1} \prod_\nu U_{b_\nu}(z).$$

Then F is meromorphic in $D(0, R)$ and analytic and non-zero in $D(0, r)$, with $|F(z)| = 1$ on $|z| = r$. By the maximum principle applied to F and $1/F$, we see that $|F(z)| \equiv 1$ for $|z| \leq r$. Hence $\log F(z)$ has constant real part on $D(0, r)$ and is constant there, by the Cauchy-Riemann equations. Thus F is constant and f is a rational function, given by

$$f(z) = C \prod_\mu \left(\frac{r(z - a_\mu)}{r^2 - \overline{a_\mu} z} \right) \prod_\nu \left(\frac{r(z - b_\nu)}{r^2 - \overline{b_\nu} z} \right)^{-1}, \tag{3.52}$$

in which C is a constant of modulus 1, and the products are over all zeros a_μ and poles b_ν in $|z| < r$, in each case with repetition according to multiplicity. Notice that the zeros and poles of f in $0 < |z| < r$ determine the poles and zeros of f in $|z| > r$.

3.9.2 Lemma

Suppose that f is meromorphic in $D(0, R)$ and that $|f(z)| = 1$ on $|z| = r_1$ and $|z| = r_2$, where $0 < r_1 < r_2 < R$. Then f is constant.

Proof. Of all those zeros of f (if any) lying in $0 < |z| < r_2$, let a be the nearest to the origin. Applying formula (3.52) with $r = r_2$, we see that $f(c) = \infty, c = r_2^2/\overline{a}$. But, according to formula

(3.52) with $r = r_1$, the function f cannot have a pole at any ζ with $|\zeta| > |r_1^2/\overline{a}|$. This contradiction shows that there cannot be any such a, and so f has no zeros, and by the same argument no poles, in $0 < |z| < r_2$. Again by (3.52), f has no zeros or poles in $|z| > r_2$ either. So $f(z) = Dz^n$ for some constant D and integer n, and the fact that $|f(z)| = 1$ on $|z| = r_1$ and $|z| = r_2$ forces $n = 0$.

3.9.3 Lemma

Let $0 < r < R$ and let f be meromorphic and non-constant in $D(0,R)$. Then there exists $C > 0$ such that, for all real t,

$$n(r, e^{it}) < C, \quad n(r,a) = n(r, 1/(f-a)).$$

Proof. Take r, s, S with $r < s < S < R$. Choose z_0 with $|f(z_0)| \neq 0, 1, \infty$ and with z_0 so close to 0 that the circle $|z| = r$ lies in $D(z_0, s)$, and such that the circle $|z - z_0| = S$ lies in $D(0,R)$. Let $g(z) = f(z_0 + z)$. Then g is meromorphic on some disc $D(0,T)$, with $T > S$, and Lemma 3.2.7 gives, for real t,

$$n(r, 1/(f - e^{it})) \leq n(s, 1/(g - e^{it})) \leq DN(S, 1/(g - e^{it})), \quad D = (\log S/s)^{-1}.$$

Now we just note that (again with t a real constant)

$$N(S, 1/(g - e^{it})) \leq T(S, 1/(g - e^{it})) = T(S, g - e^{it}) - \log|g(0) - e^{it}|.$$

This equals

$$T(S, g - e^{it}) - \log|f(z_0) - e^{it}| \leq T(S,g) + T(S, e^{it}) + \log 2 + d = T(S,g) + \log 2 + d = C_1,$$

with d and C_1 constants, independent of t, using the fact that $|f(z_0) - e^{it}| \geq ||f(z_0)| - 1|$.

3.9.4 Lemma

Let $0 < r < R$ and let f be non-constant and meromorphic on $D(0,R)$. Then $h(t) = n(r, e^{it})$ and $H(t) = N(r, e^{it})$ are measurable functions on $\mathbb{R}$.

Proof. The following argument (communicated to the author by Christian Berg) shows that for fixed r the function $n(r, e^{it})$ is measurable in t. Rouché's theorem implies that $n_-(s,a)$ is lower semi-continuous in a, where $n_-(s,a)$ denotes the number of solutions of $f(z) = a$ in $|z| < s$. Hence $n(r,a) = \lim_{s \to r+} n_-(s,a)$ is measurable.

Now consider $N(r,a)$ for $a \in \mathbb{C}$. Take all zeros $z_1, \ldots, z_m$ for $f - a$ in $|z| \leq r$. Assume for now that all of these zeros are simple and that $f(0) \neq a$.

Now take a small positive δ and let $a_n \to a$ through a sequence. Then for large n there does not exist ζ_n with $|\zeta_n| \leq r$ and $|\zeta_n - z_j| \geq \delta$ for all j and such that $f(\zeta_n) = a_n$, since otherwise we may assume that $\zeta_{n_k} \to \zeta$ which gives $f(\zeta) = a$, a contradiction. So for large n there is a root $z_{j,n}$ of $f(z) = a_n$ near to z_j, and there are no other roots of $f(z) = a_n$ in $|z| \leq r$. Hence, as $n \to \infty$,

$$N(r, a_n) = \sum_{j=1}^{m} \log^+ \frac{r}{|z_{j,n}|} \to \sum_{j=1}^{m} \log^+ \frac{r}{|z_j|} = N(r,a).$$

This shows that, for fixed r, the function $N(r,a)$ is continuous off a finite set, and therefore measurable.

3.9.5 Lemma

Let f be non-constant and meromorphic on $D(0,R)$. For $0 \le s < R$ define

$$\psi(s) = \int_0^{2\pi} n(s, e^{it})dt.$$

Suppose that $0 < r < R$ and that there are only finitely many z with $|z| = r$ and $|f(z)| = 1$. Then ψ is continuous at r.

Proof. Take S with $r < S < R$ and take C as in Lemma 3.9.3, such that $n(S, e^{it}) < C$ for all real t. Let $z_j = re^{it_j}, 1 \le j \le n$, with t_j real, be the finitely many points on $|z| = r$ at which $|f(z)| = 1$. There is no loss of generality in assuming that $0 < t_1 < \ldots < t_n < 2\pi$, since replacing $f(z)$ by $f(ze^{iQ})$, for some real Q, does not change ψ. Let $\varepsilon > 0$ and let δ be small and positive, in particular so small that $2n\delta C < \varepsilon$.

Now suppose that $t \in [0, 2\pi]$, with t not one of the t_j. Then $f(z) \neq e^{it}$ on $|z| = r$ and so $f(z) \neq e^{it}$ for $|z|$ close to r. Thus we can find $\rho_t > 0$ and $\sigma_t > 0$ such that $|f(z) - e^{it}| \ge \sigma_t$ for $r - \rho_t \le |z| \le r + \rho_t$. This in turn gives us $\eta_t > 0$ such that if p is real with $|p - t| < \eta_t$, then $f(z) \neq e^{ip}$ for $r - \rho_t \le |z| \le r + \rho_t$.

This defines $\rho_t > 0, \eta_t > 0$ for $t \in [0, 2\pi] \setminus \{t_1, \ldots, t_n\}$. For $t = t_j$, we just set $\rho_t = \eta_t = \delta$.

Now the intervals $(t - \eta_t, t + \eta_t)$ cover the compact set $[0, 2\pi]$, and so we can find a finite set J such that $[0, 2\pi]$ is a subset of the union $\bigcup_{t \in J}(t - \eta_t, t + \eta_t)$. Let ρ be the minimum of all the $\rho_t, t \in J$. By reducing ρ if necessary, we can assume that $0 < r - \rho < r + \rho < S$.

Now if p is in $[0, 2\pi]$ but not in any of the intervals $(t_j - \delta, t_j + \delta)$, then p is in the interval $(t - \eta_t, t + \eta_t)$, for some $t \in J \setminus \{t_1, \ldots, t_n\}$ and so, by definition of η_t and ρ, we have $f(z) \neq e^{ip}$ for $r - \rho \le |z| \le r + \rho$. Hence $n(s, e^{ip}) = n(r, e^{ip})$ for $r - \rho < s < r + \rho$.

We now see that for $r - \rho < s < r + \rho$ we have

$$\psi(s) - \psi(r) = I = \int_E n(s, e^{it}) - n(r, e^{it})dt,$$

in which

$$E = [0, 2\pi] \cap \left(\bigcup_{j=1}^{n} (t_j - \delta, t_j + \delta) \right).$$

Since $|n(s, e^{it}) - n(r, e^{it})| \le n(S, e^{it}) < C$, we get

$$|\psi(s) - \psi(r)| \le |I| \le 2n\delta C < \varepsilon, \quad |s - r| < \rho.$$

Chapter 4

Applications to differential equations

4.1 Some basic facts about linear differential equations

4.1.1 Existence-uniqueness theorem

Let $k \geq 1$, let D be a simply connected domain in $\mathbb{C}$, and let $a_0(z), \ldots, a_{k-1}(z)$ be analytic in D. Let $z_0 \in D$ and let $c_0, \ldots, c_{k-1} \in \mathbb{C}$. Then there exists a unique solution f of the equation

$$w^{(k)} + \sum_{j=0}^{k-1} a_j w^{(j)} = 0, \tag{4.1}$$

such that f is analytic in D and $f^{(j)}(z_0) = c_j, 0 \leq j \leq k-1$.

Proof. Once we have an analytic solution f, the uniqueness is obvious. Given two such solutions f_1, f_2, we have $(f_1 - f_2)^{(j)}(z_0) = 0$ for all $j \geq 0$, and so $f_1 - f_2 \equiv 0$ on D, by the identity theorem.

The proof of existence can be deduced as follows from the counterpart Theorem 5.5.1 for matrix DEs in the next chapter. We first write the equation (4.1) in vector form using

$$\underline{c} = (c_0, \ldots, c_{k-1})^T, \quad \underline{w} = (w_0, \ldots, w_{k-1})^T, \quad w_j = w^{(j)} \tag{4.2}$$

and

$$w_0' = w_1, \ldots, w_{k-1}' = w^{(k)} = -\sum_{j=0}^{k-1} a_j w_j. \tag{4.3}$$

Here T denotes the transpose, so that $\underline{c}$ and $\underline{w}$ are column vectors. The equation (4.1) becomes a vector DE

$$\underline{w'} = a(z)\underline{w} \tag{4.4}$$

in which $a(z)$ is a k by k matrix with entries 1 immediately above the main diagonal, and with last row $-a_0(z), \ldots, -a_{k-1}(z)$, and all other entries 0. Now choose a non-singular constant matrix B whose first column is $\underline{c}$. Then Theorem 5.5.1 gives a holomorphic solution $x(z)$ on D of the equation $x' = a(z)x$ which satisfies $x(z_0) = B$, and the first column of $x(z)$ is the required solution $\underline{w}$ of (4.4).

In the case of a general domain D, we can sometimes cover D with finitely many simply connected domains. However, it may not be possible to obtain solutions analytic in all of D. For example, $1/z$ is analytic in $D = \mathbb{C} \setminus \{0\}$. On any simply connected subdomain of D we can define $w = \log z$, and w satisfies $w'' + (1/z)w' = 0$, but w is not analytic on D.

4.1.2 Oscillation theory on the real line

Suppose that u is a real-valued solution of

$$u'' + A(x)u = 0,$$

where A is a continuous real-valued function on an open interval I in $\mathbb{R}$. Then the zeros of u in I are isolated and do not coincide with zeros of u'. For if $t \in I$ and $u(t) = u'(t) = 0$ then $u \equiv 0$ by the (real) existence-uniqueness theorem, and this will be the case if u has distinct zeros $t_k \to t$, by continuity and Rolle's theorem.

Given such a solution u of a homogeneous linear differential equation on an unbounded interval, an obvious and important question is whether u tends to infinity (e.g. e^x on $(0, \infty)$) or decays to 0 (e.g. e^{-x} on $(0, \infty)$) or is oscillatory (e.g. $\sin x$ on $(0, \infty)$). There are a lot of criteria for oscillation, and one which is easy to prove and quite useful is:

4.1.3 Sturm's comparison theorem

Suppose that G_1, G_2 are continuous real-valued functions on an open interval I in $\mathbb{R}$, and that on I the functions u, v are real-valued, not identically zero, and satisfy

$$u'' + G_1 u = 0, \quad v'' + G_2 v = 0.$$

Suppose that $x_1, x_2 \in I$ with $x_1 < x_2$ and $u(x_1) = u(x_2) = 0$ and $u(x) \neq 0$ on (x_1, x_2), and that $G_2(x) \geq G_1(x)$ on $[x_1, x_2]$. Then either (i) v has a zero in (x_1, x_2) or (ii) on $[x_1, x_2]$ the function $G_2 - G_1$ vanishes identically and v is a constant multiple of u.

Proof. Suppose that v has no zero in (x_1, x_2): then it may be assumed that $u(x)$ and $v(x)$ are positive on (x_1, x_2), and that $u'(x_1) > 0, u'(x_2) < 0$. This delivers

$$(u'v - uv')(x_2) = (u'v)(x_2) \leq 0, \quad (u'v - uv')(x_1) = (u'v)(x_1) \geq 0, \tag{4.5}$$

and so

$$0 \quad \geq \quad (u'v - uv')(x_2) - (u'v - uv')(x_1) = \int_{x_1}^{x_2} (G_2(x) - G_1(x))u(x)v(x)\, dx \geq 0.$$

Thus it must be the case that $G_2(x) = G_1(x)$ on $[x_1, x_2]$, so that $u'v - uv'$ is constant there, and hence identically zero by (4.5).

In the complex domain, there are comparatively few such results. A good reference is [45, Ch. 8], but most result are *negative*, leading to zero-free regions, lower bounds for the distance between zeros etc. However, since the solutions of (4.1) are analytic when the coefficients are, we can use the value distribution theory for meromorphic functions developed by Nevanlinna.

4.2 Nevanlinna theory and differential equations

In this section we describe some applications of Nevanlinna theory to the equation

$$w'' + A(z)w = 0, \tag{4.6}$$

in which A is an entire function. By the existence-uniqueness theorem, all solutions are entire functions. The first result goes back to Wittich.

4.2.1 Theorem

Let f be a non-trivial (i.e. not identically zero) solution of (4.6), with $A \not\equiv 0$ entire. Then
(i) We have

$$T(r, A) = S(r, f). \tag{4.7}$$

(ii) If f has finite order then A is a polynomial.
(iii) If c is a finite, non-zero complex number then

$$m(r, 1/(f - c)) = S(r, f), \tag{4.8}$$

so that in particular $\delta(c, f) = 0$.

Proof. To prove (i), we just write $-A = f''/f = (f''/f')(f'/f)$ so that the lemma of the logarithmic derivative gives

$$T(r, A) = m(r, A) \leq S(r, f) + S(r, f') = S(r, f).$$

Also (ii) follows in the same way. Later we will see that the converse of (ii) is true.

Now that we have (i), we establish (iii) by writing

$$\frac{1}{f-c} = \frac{1}{Ac}\left(\frac{f''}{f} - \frac{f''}{f-c}\right).$$

However $\delta(0, f) = 1$ is possible. Indeed,

$$w'' - (g'' + (g')^2)w = 0$$

has the zero-free solution $f = e^g$. Consequently, in order to discuss zeros of solutions of (4.6), it is normally necessary to consider two linearly independent solutions.

Let f_1, f_2 be solutions of (4.6), and let W be the Wronskian determinant

$$W = W(f_1, f_2) = f_1 f_2' - f_1' f_2.$$

Then $W' = 0$ and $W = c$ is a constant. It is easy to see that $c = 0$ if and only if f_1 and f_2 are linearly dependent. We say that f_1 and f_2 are *normalized* LI solutions if $W(f_1, f_2) = 1$.

4.2.2 A result of Bank (Crelle's Journal, 1972)

The result of (iii) in §4.2.1 generalizes as follows. Suppose that f is a transcendental meromorphic function in the plane and satisfies a k'th order differential equation

$$0 = \sum_{j=1}^{p} a_j f^{m_{0,j}} (f')^{m_{1,j}} \dots (f^{(k)})^{m_{k,j}}, \tag{4.9}$$

with meromorphic coefficients a_j, which are not all identically zero and satisfy $T(r, a_j) = S(r, f)$. Let n be the degree of the equation (the largest of those sums $m_{0,j} + \ldots + m_{k,j}$ for which $a_j \not\equiv 0$), and set $F = f'/f$. Then for each positive integer k, we can write

$$f^{(k)} = Q_k(F) f,$$

in which $Q_k(F)$ is a polynomial in F and its derivatives, with constant coefficients. This is easily proved by induction, using

$$f' = Ff, \quad f'' = (F' + F^2)f,$$

and

$$f^{(k+1)} = (Q_k(F))' f + Q_k(F) F f.$$

Grouping together all terms of the same degree, we can write the equation (4.9) in the form

$$0 = \sum_{q=0}^{n} f^q L_q(z, F), \tag{4.10}$$

in which each L_q is a polynomial in F and its derivatives, with coefficients b satisfying $T(r, b) = S(r, f)$. There are now two cases.

Case 1: We have $L_q \equiv 0$ for every q.

In this case for each q the equation $0 = L_q(z, F)$ gives a *homogeneous* differential equation satisfied by f.

Case 2: Suppose $s+1$ is the greatest q for which $L_q \not\equiv 0$. Then we divide the equation (4.10) through by $f^s L_{s+1}$ to get an equation

$$f = \sum_{k=0}^{s} f^{k-s} M_k, \quad M_k = -L_k / L_{s+1},$$

where

$$T(r, M_k) \leq O(T(r, F)) + S(r, f).$$

Hence

$$m(r, f) \leq \sum_{k=0}^{s} m(r, M_k) + O(1) \leq O(T(r, F)) + S(r, f)$$

and

$$N(r, f) \leq \sum_{k=0}^{s} N(r, M_k) \leq O(T(r, F)) + S(r, f).$$

This gives

$$T(r, f) \leq O(T(r, F)) + S(r, f) \leq O(\overline{N}(r, f) + \overline{N}(r, 1/f)) + S(r, f). \tag{4.11}$$

We illustrate this with two examples. First, if $A \not\equiv 0$ is an entire function and c is a non-zero complex number then (4.6) may be written as

$$w'' + A(w - c) = -Ac.$$

Hence if f is a non-trivial solution of (4.6) we have a non-homogeneous differential equation in $g = f - c$ with coefficients which are small functions compared to g, and so we get

$$T(r, f) \leq T(r, g) + O(1) \leq O(\overline{N}(r, 1/g)) + S(r, g) = O(\overline{N}(r, 1/(f - c))) + S(r, f).$$

Next, suppose that f is a transcendental meromorphic solution in the plane of

$$aff'' + bf'^2 + cf^2 + Af'' + Bf' + Cf + D = 0,$$

with a, b, c, A, B, C, D rational functions. With $F = f'/f$ we get

$$f^2 L_2 + f L_1 + L_0 = 0,$$

in which

$$L_2 = a(F' + F^2) + bF^2 + c, \quad L_1 = A(F' + F^2) + BF + C, \quad L_0 = D.$$

If all the L_q are identically zero, we get three homogeneous equations, namely

$$aff'' + bf'^2 + cf^2 = 0, \quad Af'' + Bf' + Cf = 0, \quad D = 0,$$

which in principle may be easier to solve. If some L_q fails to vanish identically, we can estimate $T(r, f)$ in terms of $\overline{N}(r, f)$ and $\overline{N}(r, 1/f)$ using (4.11). In particular, if

$$\overline{N}(r, f) + \overline{N}(r, 1/f) = S(r, f)$$

then we must have $L_0 = L_1 = L_2 = 0$.

4.2.3 The Schwarzian

For meromorphic U define

$$S(U) = \{U, z\} = \frac{U'''}{U'} - \frac{3}{2}\left(\frac{U''}{U'}\right)^2. \tag{4.12}$$

Note that if U has a simple pole at a then there is a constant $c \neq 0$ such that

$$U'(z) = c(z-a)^{-2} + O(1), \quad U''(z) = -2c(z-a)^{-3} + O(1), \quad U'''(z) = 6c(z-a)^{-4} + O(1), \quad z \to a.$$

Hence the only poles of S are at zeros of U' and multiple poles of U, i.e. at multiple points of U.

If U is the quotient f_1/f_2 of LI solutions of (4.6), then we have $U' = cf_2^{-2}$ for some non-zero constant c, and an easy calculation gives

$$S(U) = 2A. \tag{4.13}$$

Also $U' \neq 0$ and, since f_2 has only simple zeros, U is locally one-one.

Conversely, suppose that F is meromorphic without multiple points on a simply connected domain D. Then (4.13) defines a function A analytic on D, and it is easy to check that $f_2 = (U')^{-1/2}$ is an analytic solution of (4.6) in D. If we choose a second solution f_1 of (4.6) such that $W(f_1, f_2) = -1$ then $U' = (f_1/f_2)'$ and U is the quotient of linearly independent solutions of (4.6).

The Schwarzian derivative plays an important role in conformal mapping. Suppose that U is meromorphic and locally one-one in the unit disc $D(0, 1)$. If U is one-one in D then

$$(1 - |z|^2)^2 |S(U)| \leq 6$$

there. In the other direction,

$$(1 - |z|^2)^2 |S(U)| \leq 2$$

is sufficient to imply that U is one-one. Both constants are sharp and the results are due to Nehari. The first uses coefficient inequalities and the second can be proved using differential equations or quasiconformal maps. Note that if $U = f_1/f_2$ is one-one on D then each of f_1 and f_2 has at most one zero in D.

4.2.4 The Bank-Laine product

This approach was introduced by Bank and Laine [8]. It is convenient first to note that if h and E are related by

$$\frac{h'}{h} = \frac{1}{2}\left(\frac{E'}{E} + \frac{c}{E}\right),$$

where $c = \pm 1$ is a constant, then a straightforward calculation shows that

$$\frac{h''}{h} = -\frac{1}{4}\left(\frac{(E')^2 - 2E''E - 1}{E^2}\right).$$

Now let f, g be LI solutions of (4.6), normalized so that $W(f, g) = fg' - gf' = 1$, and set

$$U = \frac{f}{g}, \quad E = fg, \quad \frac{U'}{U} = -\frac{1}{E}. \tag{4.14}$$

Then

$$\frac{E'}{E} = \frac{f'}{f} + \frac{g'}{g}, \quad \frac{1}{E} = \frac{g'}{g} - \frac{f'}{f}.$$

Solving thus gives

$$\frac{f'}{f} = \frac{E'-1}{2E}, \quad \frac{g'}{g} = \frac{E'+1}{2E}. \tag{4.15}$$

and so the identity above, with $h = f$, yields the *Bank-Laine equation*

$$4A = \frac{(E')^2 - 2E''E - 1}{E^2}. \tag{4.16}$$

Multiplying out by E^2 and differentiating, we also have

$$E''' + 4AE' + 2A'E = 0. \tag{4.17}$$

Note that (4.17) appears in [47], but (4.16) does not seem to have been used before Bank and Laine.

The product E is a *Bank-Laine function*: this means an entire function E such that $E = 0$ implies $E' = \pm 1$. Conversely, suppose that E is a Bank-Laine function. Then A as defined by (4.16) is entire, since the numerator has at least a double zero at any zero of E. Choose w with $E(w) \neq 0$ and define f and g near w by (4.15). Then f and g are solutions of (4.6) near w and so are entire functions. Since the Wronskian of f and g is then a constant, which has to be non-zero, and since (4.15) is unaffected if f and g are multiplied by a constant, it may be assumed that $W(f, g) = 1$. But then (4.15) gives

$$\frac{1}{E} = \frac{g'}{g} - \frac{f'}{f} = \frac{1}{fg}.$$

Thus we obtain:

4.2.5 Theorem (Bank-Laine 1982-3)

An entire function E is a Bank-Laine function if and only if E is the product of linearly independent normalized solutions of an equation (4.6) with A entire.

4.2.6 The advantages of the product

Let f_1, f_2 be normalized LI solutions of (4.6), with product $E = f_1 f_2$. Let c denote a positive constant (not necessarily the same at every occurrence).

(i) *We have*

$$T(r, A) \leq cT(r, E) + S(r, E). \tag{4.18}$$

This follows at once from (4.16).

(ii) *We have*

$$T(r, E) \leq \frac{1}{2} T(r, A) + N(r, 1/E) + S(r, E). \tag{4.19}$$

To see this, write $T(r, E) = m(r, 1/E) + N(r, 1/E) + O(1)$ and note from (4.16) that

$$2m(r, 1/E) = m(r, 1/E^2) \leq T(r, A) + S(r, E).$$

(iii) *If A has finite order and the zeros of E have finite exponent of convergence, then E has finite order.*

(iv) *If E has finite order then A is a polynomial if and only if $m(r, 1/E) = O(\log r)$.*

4.2.7 Examples of Bank-Laine functions

(i) Let $E = e^Q$ with Q a polynomial. Then E is a Bank-Laine function and A has the form

$$4A = -2Q'' - (Q')^2 - e^{-2Q}.$$

(ii) Let P be a polynomial with only simple zeros, and let Q be a non-constant polynomial, chosen using Lagrange interpolation, so that $E = Pe^Q$ is a Bank-Laine function. Here both E and A have order equal to the degree of Q.

(iii) Let $K = (2n + 1)^2/16$ with n a non-negative integer, and define

$$Q(\zeta) = \sum_{m=0}^{n} a_m \zeta^m$$

by $a_0 = 1$ and, with $c = \pm i$,

$$(4m^2 + 4m + 1 - 16K)a_m = 16c(m + 1)a_{m+1}.$$

Then $W(z) = Q(e^{-z/2})$ satisfies

$$W'' + W'(2ce^{z/2} - 1/2) + W(-K + 1/16) = 0$$

and $w(z) = W(z) \exp(2ce^{z/2} - z/4)$ solves

$$w'' + (e^z - K)w = 0. \tag{4.20}$$

We thus have linearly independent solutions whose zeros have exponent of convergence at most 1. In fact, the change of variables $\zeta = 2e^{z/2}, u(\zeta) = w(z)$, turns (4.20) into Bessel's equation (this is in [45]). There are quite a lot of similar examples of equations (4.6), with A a polynomial in $e^{\alpha z}$ and $e^{-\alpha z}$, having LI solutions with $\lambda(f_1 f_2) \leq 1$.

4.2.8 The Bank-Laine conjecture

It is conjectured that if A is a transcendental entire function and the equation (4.6) has linearly independent solutions f_1, f_2 with $\lambda(f_1 f_2) < \infty$, then the order of A is either ∞ or a positive integer.

It has been proved (Rossi, Shen 1986) that if A is transcendental and $\rho(A) \leq 1/2$ then $\lambda(f_1 f_2) = \infty$.

4.2.9 Theorem (Bank-Laine)

Suppose that A is a transcendental entire function of order $\rho < \alpha < 1/2$, and that $E = f_1 f_2$ is the product of normalized LI solutions of (4.6). Then $\lambda(E) = \infty$.

Proof. Suppose that $\lambda(E) < \infty$. Then E has finite order. By Lemmas 3.7.2 and 3.7.4 there exists a constant $M > 0$ such that provided $|z|$ lies outside a set of finite measure we have

$$|E''(z)/E(z)| + |E'(z)/E(z)| \leq |z|^M. \tag{4.21}$$

The next ingredient is a classical result known as the $\cos \pi\rho$ theorem: since A has order $\rho < \alpha < 1/2$ we have

$$\frac{\log |A(z)|}{\log M(r, A)} > \cos \pi\alpha > 0, \quad |z| = r, \tag{4.22}$$

for all r in a set H of lower logarithmic density at least $1 - \rho/\alpha$, so that

$$\int_{H \cap [1,s]} \frac{dt}{t} > (1 - \rho/\alpha - o(1)) \log s, \quad s \to \infty.$$

This gives us arbitrarily large r satisfying (4.22), such that (4.21) also holds on $|z| = r$. Since

$$\log r = o(T(r, A)) = o(\log M(r, A)),$$

we deduce from (4.16) that E must be small on the whole circle $|z| = r$, which is obviously impossible, by the maximum principle.

4.2.10 Theorem (Bank-Laine)

Suppose that A is a transcendental entire function of finite order ρ, and that (4.6) has normalized LI solutions f_1, f_2 such that $\lambda(f_1 f_2) < \rho$. Then ρ is a positive integer.

Proof. With $E = f_1 f_2$ we have

$$\lambda(E) < \rho(A) \leq \rho(E) < \infty.$$

Hence we may write $E = \Pi e^g$ with Π entire of order $\lambda(E)$ and g a polynomial, of degree $\rho(E)$. We now have

$$m(r, 1/E) = (1 + o(1))T(r, E)$$

and so $\rho(A) \geq \rho(E)$.

4.3 Polynomial coefficients

There is an extensive literature on the asymptotic behaviour of solutions of (4.1), when the a_j are polynomials or rational functions. We will describe here the solutions of

$$w'' + b(z)w = 0, \tag{4.23}$$

when $b(z)$ is a rational function with $b(z) = cz^n(1 + o(1)), z \to \infty, n \geq -1, c \neq 0$.

4.3.1 Hille's method

Let $c > 0$ and $0 < \varepsilon < \pi$. Then there exists a constant $d > 0$, depending only on c and ε, with the following properties.

Suppose that the function F is analytic, with $|F(z)| \leq c|z|^{-2}$, in

$$\Omega = \{z : 1 \leq R \leq |z| \leq S < \infty, |\arg z| \leq \pi - \varepsilon\}. \tag{4.24}$$

Then the equation

$$w'' + (1 - F(z))w = 0 \tag{4.25}$$

has linearly independent solutions $U(z), V(z)$ satisfying

$$\begin{aligned} U(z) = e^{-iz}(1 + \delta_1(z)), \quad U'(z) = -ie^{-iz}(1 + \delta_2(z)), \\ V(z) = e^{iz}(1 + \delta_3(z)), \quad V'(z) = ie^{iz}(1 + \delta_4(z)), \end{aligned} \tag{4.26}$$

in which

$$|\delta_j(z)| \leq \frac{d}{|z|} \quad \textit{for} \quad z \in \Omega_1 = \Omega \setminus \{z : \mathrm{Re}(z) < 0, |\mathrm{Im}(z)| < R\}. \tag{4.27}$$

Here Ω_1 can be thought of as Ω with the "shadow" of $D(0, R)$ removed.

To prove this, let $X = Se^{i\sigma}$, where $\sigma = \min\{\pi/2, \pi - \varepsilon\}$. Choose a solution v of the equation

$$v'' + 2iv' - Fv = 0, \tag{4.28}$$

analytic in Ω, such that $v(X) = 1, v'(X) = 0$. Set, for $z \in \Omega$,

$$L(z) = v(z) - 1 + \frac{1}{2i}\int_X^z (e^{2i(t-z)} - 1)F(t)v(t)dt, \tag{4.29}$$

the integration being independent of path in Ω, by Cauchy's theorem. Now

$$L'(z) = v'(z) - \int_X^z e^{2i(t-z)}F(t)v(t)dt, \tag{4.30}$$

and

$$\begin{aligned} L''(z) &= v''(z) + 2i\int_X^z e^{2i(t-z)}F(t)v(t)dt - F(z)v(z) \\ &= v''(z) + 2i(v'(z) - L'(z)) - F(z)v(z) = -2iL'(z). \end{aligned}$$

Since $L(X) = L'(X) = 0$, the existence-uniqueness theorem gives $L(z) \equiv 0$ on Ω.

Now let $z \in \Omega_1$. Choose the path of integration γ_z to be the arc of the circle $|t| = S$ from X clockwise to the first point x of intersection of the circle $|t| = S$ with the line $\mathrm{Im}(t) = \mathrm{Im}(z)$, followed by the straight line segment from x to z. Then $\mathrm{Im}(t - z) \geq 0$ and hence $|e^{2i(t-z)}| \leq 1$ on γ_z, and this is the reason for the choice of Ω_1 and X.

Since $L(z) = 0$, (4.29) gives

$$|v(z) - 1| \leq \int_X^z |F(t)v(t)|\,|dt|, \quad |v(z)| \leq 1 + \int_X^z |F(t)v(t)|\,|dt|. \tag{4.31}$$

We apply the method generally known as Gronwall's lemma. Let s denote arc length on γ_z, and parametrize γ_z with respect to s. Set

$$H(s) = 1 + \int_X^{\zeta(s)} |F(t)v(t)|\,|dt| = 1 + \int_0^s |F(\zeta(s))v(\zeta(s))|\,ds, \quad \zeta \in \gamma_z.$$

Then the second estimate of (4.31) gives

$$\frac{dH}{ds} = |F(\zeta(s))v(\zeta(s))| \leq |F(\zeta(s))|H(s)$$

so that

$$H(s) = \frac{H(s)}{H(0)} \leq \exp\left(\int_X^{\zeta(s)} |F(t)|\,|dt|\right).$$

Thus the first estimate of (4.31) becomes

$$|v(z) - 1| \leq H(s) - 1 \leq \exp\left(\int_X^{z} |F(t)|\,|dt|\right) - 1. \tag{4.32}$$

Let $d_1, d_2, \ldots$ denote positive constants depending only on c and ε. The circle $|t| = S$ evidently contributes at most $d_1 S^{-1} \leq d_1|z|^{-1}$ to the integral in (4.32). Similarly, if $|\arg z| \leq \pi/4$ then $\mathrm{Re}(z) > 0$ and the horizontal part of γ_z contributes at most

$$\int_{Re(z)}^{\infty} \frac{c}{t^2}\,dt \leq \frac{d_2}{\mathrm{Re}(z)} \leq \frac{d_3}{|z|}.$$

Finally, if $\pi/4 \leq |\arg z| \leq \pi - \varepsilon$ we write $z = a + ib$ with a, b real and $|b| > d_4|z|$, and the contribution from the horizontal part of γ_z to the integral in (4.32) is at most

$$\int_{\mathbb{R}} \frac{c}{x^2 + b^2}\,dx \leq \frac{d_5}{|b|} \leq \frac{d_6}{|z|}.$$

Thus (4.32) gives

$$|v(z) - 1| \leq \exp\left(\frac{d_7}{|z|}\right) - 1 \leq \frac{d_8}{|z|} \leq d_8,$$

using the fact that $R \geq 1$, and (4.30) gives

$$|v'(z)| \leq \int_X^z |F(t)| d_9\,|dt| \leq \frac{d_{10}}{|z|}.$$

Now we need only set $V(z) = v(z)e^{iz}$ so that V solves (4.25), by (4.28), and (4.26) for V follows at once.

To obtain U, we set $Y = \overline{X}$ and choose a solution u of

$$u'' - 2iu' - Fu = 0,$$

with $u(Y) = 1, u'(Y) = 0$, and the integral equation for u is

$$u = 1 + \frac{1}{2i}\int_Y^z (e^{-2i(t-z)} - 1)F(t)u(t)dt.$$

The path of integration has $\mathrm{Im}(t - z) \leq 0$. Finally we set $U(z) = u(z)e^{-iz}$.

4.3.2 Other regions

An almost identical argument works if Ω is replaced by

$$\{z : 1 \leq R \leq |z| \leq S < \infty, |\arg z - \pi| \leq \pi - \varepsilon\},$$

with this time

$$\Omega_1 = \Omega \setminus \{z : \mathrm{Re}(z) > 0, |\mathrm{Im}(z)| < R\}.$$

We may also replace Ω with an unbounded region. Suppose that F is analytic, with $|F(z)| \leq c|z|^{-2}$, in

$$\Omega' = \{z : 1 \leq R \leq |z| < \infty, |\arg z| \leq \pi - \varepsilon\}.$$

We take a sequence $S_n \to \infty$, and obtain corresponding solutions U_n, V_n in

$$\{z : R \leq |z| \leq S_n, |\arg z| \leq \pi - \varepsilon\} \setminus \{z : \mathrm{Re}(z) < 0, |\mathrm{Im}(z)| < R\}.$$

The corresponding error terms $\delta_{j,n}(z), j = 1, 2, 3, 4$, are uniformly bounded, since the constant d is independent of S in §4.3.1. Thus by normal families we may assume, passing to a subsequence if necessary, that the $U_n, V_n, \delta_{j,n}$ converge locally uniformly on

$$\Omega'' = \{z : 1 \leq R < |z| < \infty, |\arg z| < \pi - \varepsilon\} \setminus \{z : \mathrm{Re}(z) \leq 0, |\mathrm{Im}(z)| \leq R\}.$$

The limit functions U, V solve (4.26), and the corresponding $\delta_j(z)$ satisfy (4.27) on Ω''.

4.3.3 Equations with a polynomial coefficient

The standard application of Hille's method is to the equation (4.23), when b is a polynomial, not identically zero. Slightly more generally, suppose that $b(z)$ is analytic in $R_0 < |z|$, with

$$b(z) = cz^n(1 + o(1)), \quad z \to \infty,$$

in which c is a non-zero constant and n is an integer not less than -1.

4.3.4 The case $n = -1$

If $n = -1$ it is convenient to set

$$z = u^2, \quad g(u) = f(z) = f(u^2),$$

in which f is a solution of (4.23). Then g solves

$$g''(u) = 2f'(u^2) + 4u^2 f''(u^2) = g'(u)/u - 4u^2 b(u^2) g(u)$$

and so

$$g''(u) - g'(u)/u + c(u)g(u) = 0, \quad c(u) = 4u^2 b(u^2) = 4c(1 + o(1)), \quad u \to \infty.$$

Now set $h(u) = u^{-1/2} g(u) = u^{-1/2} f(u^2)$ so that h satisfies

$$h''(u) + (c(u) - 3/4u^2)h(u) = 0.$$

In the equation for h we have $n = 0$, and from the asymptotic behaviour of h we can deduce that of f. We assume henceforth that $n \geq 0$.

4.3.5 Critical rays

The *critical rays* are those rays $\arg z = \theta \in \mathbb{R}$ for which

$$\arg c + (n+2)\theta = 0 \pmod{2\pi}. \tag{4.33}$$

Assume that $\arg z = \theta_0$ is a critical ray, let R_0 be large and positive, and with ε small and positive define

$$Z = \int_{2R_0 e^{i\theta_0}}^{z} b(t)^{1/2} dt = \frac{2c^{1/2}}{n+2} z^{(n+2)/2}(1+o(1)), \quad z \to \infty, \quad |\arg z - \theta_0| \le \frac{2\pi}{n+2} - \varepsilon.$$

Here we are free to choose either branch of $b(t)^{1/2}$ (each of which is of course -1 times the other). The condition (4.33) implies that cz^{n+2} is real and positive on the critical ray, and so we may choose the branch of $b(t)^{1/2}$ in order to ensure that $c^{1/2}z^{(n+2)/2}$ is also real and positive on $\arg z = \theta_0$. We assume henceforth that this has been done.

4.3.6 Lemma

Let R_1 be large and let σ be small and positive. Let $V = V(z)$ satisfy

$$V(z) = \frac{2c^{1/2}}{n+2} z^{(n+2)/2}(1+o(1)) \quad \textit{as} \quad z \to \infty, \quad |\arg z - \theta_0| \le \frac{2\pi}{n+2} - \tau, \tag{4.34}$$

where $0 < \tau < \sigma$. Then V is univalent on the region T_1 given by

$$|z| > R_1, \quad |\arg z - \theta_0| < \frac{2\pi}{n+2} - \sigma,$$

and V maps T_1 onto a region containing

$$T_1^* = \{w : \, |w| > R_1^*, |\arg w| < \pi - \sigma^*\}.$$

Here we may take any large R_1^* and any σ^* with $\sigma^* > (n+2)\sigma/2$.

To prove the lemma, note first that

$$\zeta = \frac{2c^{1/2}}{n+2} z^{(n+2)/2}$$

(with the same choice of square roots as before) is univalent on the region T_2 given by

$$|z| > 0, \quad |\arg z - \theta_0| < \frac{2\pi}{n+2} - \frac{\sigma}{2},$$

and ζ maps T_2 onto the sector

$$T_3 = \left\{ \zeta : |\zeta| > 0, \quad |\arg \zeta| < \pi - \frac{\sigma(n+2)}{4} \right\}.$$

But (4.34) and Cauchy's estimate for derivatives give

$$\frac{dV}{dz} \sim c^{1/2} z^{n/2} = \frac{d\zeta}{dz}, \quad \frac{dV}{d\zeta} = \frac{dV}{dz}\frac{dz}{d\zeta} = 1 + o(1) \quad \text{as} \quad z \to \infty, \quad |\arg z - \theta_0| < \frac{2\pi}{n+2} - \sigma.$$

Thus if R_1 is large and z_1, z_2 are distinct and in T_1, we set $\zeta_j = \zeta(z_j)$, and we may integrate from ζ_1 to ζ_2 along a a straight line, to obtain

$$V(z_1) - V(z_2) = \int_{\zeta_2}^{\zeta_1} \frac{dV}{d\zeta}\, d\zeta = \int_{\zeta_2}^{\zeta_1} 1 + o(1)\, d\zeta = \zeta_1 - \zeta_2 + o(|\zeta_1 - \zeta_2|) \neq 0.$$

This shows that V is univalent on T_1. To see that $V(T_1)$ contains T_1^*, just take R large and positive, and σ' with $(n+2)\sigma/2 < (n+2)\sigma'/2 < \sigma^*$, and look at the image under ζ of

$$U_R = \{z : R < |z| < 2R, \quad |\arg z - \theta_0| < \frac{2\pi}{n+2} - \sigma'\}.$$

This is, for some large S,

$$V_R = \left\{ w : S < |w| < 2^{(n+2)/2} S, \quad |\arg w| < \pi - \frac{(n+2)\sigma'}{2} \right\}.$$

As z goes once around the boundary ∂U_R we see that ζ goes once around ∂V_R, and $V(z)$ describes a simple closed curve Γ_R which is close to ∂V_R, since $V(z) \sim \zeta$. But $V(T_1)$ is simply connected, and so the interior of Γ_R lies in $V(T_1)$, which gives

$$\left\{ w : S(1+\sigma) < |w| < 2^{(n+2)/2} S(1-\sigma), \quad |\arg w| < \pi - \sigma^* \right\} \subseteq V(T_1).$$

This proves the last conclusion.

4.3.7 The Liouville transformation

Let δ be small and positive, let R_1 be large and write

$$W(Z) = b(z)^{1/4} w(z), \tag{4.35}$$

in which w is a solution of (4.23), and z lies in

$$Q_1 = \left\{ z : |z| > \frac{R_1}{4}, \qquad |\arg z - \theta_0| < \frac{2\pi}{n+2} - \frac{\delta}{4} \right\}.$$

By Lemma 4.3.6, we have, for some large R_2,

$$Q_2 = \left\{ w : |w| > R_2, \quad |\arg w| < \pi - \frac{(n+2)\delta}{4} \right\} \subseteq Z(Q_1),$$

and the same asymptotics for Z show that

$$Z(S_1) \subseteq Q_2, \quad \text{where} \quad S_1 = \left\{ z : |z| > R_1, \qquad |\arg z - \theta_0| < \frac{2\pi}{n+2} - \delta \right\}.$$

The equation (4.23) transforms to

$$\frac{d^2 W}{dZ^2} + (1 - F_0(Z))W = 0, \quad F_0(Z) = \frac{b''(z)}{4b(z)^2} - \frac{5b'(z)^2}{16 b(z)^3}, \tag{4.36}$$

and we have $|F_0(Z)| = O(|Z|^{-2})$ in Q_2. By §4.3.1 there exist solutions $U_1(Z), U_2(Z)$ of (4.36) satisfying (4.26) in Q_2 and these give *principal* solutions

$$u_j(z) = b(z)^{-1/4} \exp((-1)^j iZ + o(1)) \tag{4.37}$$

of (4.23) in S_1.

The u_j are zero-free in S_1, but if A, B are non-zero constants we show that

$$w = Au_1 - Bu_2$$

has zeros near the critical ray, as follows. Set

$$V(z) = \frac{1}{2i} \log \frac{u_2(z)}{u_1(z)}.$$

Now, $w(z) = 0$ if and only if $u_2/u_1 = A/B$, which is the same as

$$2iV(z) = \log \frac{u_2(z)}{u_1(z)} = 2iZ + o(1) = \log(A/B) + k2\pi i, \tag{4.38}$$

with k an integer and any (fixed) determination of $\log(A/B)$. First of all, if $z \in S_1$ is large and $w(z) = 0$ then (4.38) gives

$$\frac{2c^{1/2}}{n+2} z^{(n+2)/2}(1 + o(1)) \sim V(z) \sim k\pi,$$

and in particular this leads to $\arg V(z) = o(1)$ and hence $\arg z \sim \theta_0$. Thus zeros z of w in S_1 with $|z|$ large must lie near the critical ray.

Now let k be a large positive integer. Then

$$V_k = \frac{1}{2i} \log \frac{A}{B} + k\pi$$

lies near the positive real axis, and so by Lemma 4.3.6 there is a solution z_k of $V(z_k) = V_k$ in S_1. Moreover, this z_k is unique by the univalence of V and z_k lies near the critical ray. Now the number of these V_k inside a disc of centre 0 and large radius R is $(1 + o(1))R/\pi$. Hence by (4.34) the number of these zeros z_k of w in $|z| \leq S$ is $(1 + o(1))c_1 S^{(n+2)/2}$ as $S \to \infty$, for some positive constant c_1, which gives the following result [8].

Theorem. *Let $b \not\equiv 0$ be a polynomial of degree n and let w be a solution of (4.23) with infinitely many zeros. Then*

$$\liminf_{r \to \infty} \frac{N(r, 1/w)}{r^{(n+2)/2}} > 0.$$

4.4 Asymptotics for equations with transcendental coefficients

For a linear differential equation with transcendental entire coefficients it is in general much harder to obtain asymptotic representations for the solutions. However, when one coefficient is sufficiently dominant it is possible to obtain local representations for solutions with few zeros. For the case $k = 2$ it is interesting to compare the results of the next theorem with the solutions (4.37) obtained for polynomial coefficients.

4.4.1 Theorem

Let $k \geq 2$ and let $A_0, \ldots, A_{k-2}$ be entire functions of finite order, with $A = A_0$ transcendental. Let E_1 be a subset of $[1, \infty)$, of infinite logarithmic measure, and with the following property. For each $r \in E_1$ there exists an arc

$$a_r = \{re^{it} : 0 \leq \alpha_r \leq t \leq \beta_r \leq 2\pi\} \tag{4.39}$$

of the circle $S(0,r)$, such that

$$\lim_{r\to\infty, r\in E_1} \frac{\min\{\log|A(z)| : z\in a_r\}}{\log r} = \infty, \tag{4.40}$$

and, if $k\geq 3$,

$$\lim_{r\to\infty, r\in E_1} \max\left\{\frac{\log^+|A_j(z)|}{\log|A(z)|} : z\in a_r\right\} = 0, \tag{4.41}$$

for $j=1,\ldots,k-2$.

Let f be a solution of

$$y^{(k)} + \sum_{j=0}^{k-2} A_j y^{(j)} = 0, \tag{4.42}$$

with $\lambda(f) < \infty$. Then there exists a subset $E_2\subseteq [1,\infty)$ of finite measure, such that for large $r\in E_0 = E_1\setminus E_2$ the following is true. We have

$$\frac{f'(z)}{f(z)} = c_r A(z)^{1/k} - \frac{k-1}{2k}\frac{A'(z)}{A(z)} + O(r^{-2}), \quad z\in a_r. \tag{4.43}$$

Here c_r is a constant which may depend on r, but satisfies $c_r^k = -1$. The branch of $A^{1/k}$ in (4.43) is analytic on a_r (including in the case where a_r is the whole circle $S(0,r)$).

We may summarize (4.40) and (4.41) as saying that, as $r\to\infty$ in E_1,

$$|z| + \sum_{1\leq j\leq k-2} |A_j(z)| \leq |A(z)|^{o(1)} \tag{4.44}$$

for $z\in a_r$. To prove the theorem, we start by writing

$$f = Ve^h, \quad \rho(V) < \infty, \tag{4.45}$$

where V and h are entire functions. We may assume that $h'\not\equiv 0$ (if $h'\equiv 0$ then h is constant and we can replace $h(z)$ by $h(z)+z$ and $V(z)$ by $V(z)e^{-z}$, which has finite order).

Now

$$\frac{f'}{f} = \frac{V'}{V} + h'$$

and it is easy to prove by induction that, for $m=1,2,\ldots$,

$$\frac{f^{(m)}}{f} = (h')^m + m(h')^{m-1}\frac{V'}{V} + \frac{m(m-1)}{2}(h')^{m-2}h'' + T_{m-2}(h'), \tag{4.46}$$

where $T_{m-2}(h')$ is a polynomial in h' of degree at most $m-2$, with coefficients which are polynomials in the logarithmic derivatives $V^{(j)}/V, h^{(j)}/h', j=1,\ldots,m$ (for $m=1$ we set $T_{m-2}=0$).

Denote positive constants by M_j. Substituting (4.46) into (4.42) gives

$$(h')^k + k(h')^{k-1}\frac{V'}{V} + \frac{k(k-1)}{2}(h')^{k-2}h'' + T_{k-2}(h') +$$

$$+ \sum_{1\leq j\leq k-2} A_j\left((h')^j + j(h')^{j-1}\frac{V'}{V} + \frac{j(j-1)}{2}(h')^{j-2}h'' + T_{j-2}(h')\right) + A_0 = 0. \tag{4.47}$$

Claim 1: h' has finite order (and therefore so has h).

To prove this suppose $|z| = r$ is large and $|h'(z)| \geq 1$, and divide (4.47) through by $h'(z)^{k-1}$. Since

$$m(r, V^{(j)}/V) = O(\log r), \quad r \to \infty,$$

for each $j \in \mathbb{N}$, and since $A_0, \ldots, A_{k-2}$ have finite order, we obtain

$$m(r, h') \leq S(r, h') + O(\log r) + O(r^{M_0})$$

outside a set E_2 of finite measure, giving

$$m(r, h') = O(r^{M_0}), \quad r \notin E_2.$$

For large $r \in E_2$, choose $s \in [r, 2r] \setminus E_2$ to obtain

$$m(r, h') \leq m(s, h') = O(s^{M_0}) = O(r^{M_0}).$$

This proves Claim 1.

Since V, h' and the coefficients A_μ have finite order we can use §3.7 to find points u_m with $|u_m| \geq 4$ and $u_m \to \infty$ as $m \to \infty$ such that

$$\left|\frac{V^{(j)}(z)}{V(z)}\right| + \left|\frac{h^{(j)}(z)}{h'(z)}\right| + \left|\frac{A'_\mu(z)}{A_\mu(z)}\right| \leq |z|^{M_1} \tag{4.48}$$

for $1 \leq j \leq k$ and $0 \leq \mu \leq k-2$ and for all large z satisfying

$$z \notin U_0 = \bigcup_{m=1}^{\infty} D(u_m, |u_m|^{-M_2}), \tag{4.49}$$

and this can be done so that

$$\sum_{m=1}^{\infty} |u_m|^{-M_2} < \infty. \tag{4.50}$$

Let U be the set obtained by doubling the radii of all the discs of U_0. Since the set of $r \geq 1$ such that the circle $S(0, r)$ meets the disc $D(u_m, 2|u_m|^{-M_2})$ has linear measure at most $2|u_m|^{-M_2} \leq 2$, it follows using (4.50) that there exists a set E_2 of finite linear measure such that for $r \notin E_2$ the circle $S(0, r)$ meets none of the discs of U.

Let $E_0 = E_1 \setminus E_2$ be as in the statement of the theorem. Then E_0 is unbounded. Let $M_3 > 0$ be large compared to M_1 and M_2.

Claim 2: for large $r \in E_0$ and $z_0 \in a_r$ we have (4.44) and (4.48) for $z \in D(z_0, |z_0|^{-M_3})$.

To prove Claim 2, note first that if $r \in E_0$ is large then the circle $S(0, r)$ does not meet U, and so provided M_3 was chosen large enough the disc $D(z_0, |z_0|^{-M_3})$ does not meet any of the discs $D(u_m, |u_m|^{-M_2})$, so that (4.48) holds for $z \in D(z_0, |z_0|^{-M_3})$. In particular, integrating A'_μ/A_μ shows that

$$|\log |A_\mu(z)/A_\mu(z_0)|| = \left|\int_{z_0}^{z} A'_\mu(t)/A_\mu(t)dt\right| < \ln 2$$

for $z \in D(z_0, |z_0|^{-M_3})$, again provided M_3 was chosen large enough, which gives

$$|A(z)| \geq \frac{1}{2}|A(z_0)|, \quad |A_\mu(z)| \leq 2|A_\mu(z_0)|,$$

and so (4.44) for such z. This proves Claim 2.

Claim 3: for large $r \in E_0$ and $z_0 \in a_r$ we have

$$\frac{1}{2}|A(z)|^{1/k} \leq |h'(z)| \leq 2|A(z)|^{1/k}. \tag{4.51}$$

Suppose first that $|h'(z)| < \frac{1}{2}|A(z)|^{1/k}$. Then (4.44), (4.47) and (4.48) give

$$|A(z)| < 2^{-k}|A(z)|+|A(z)|^{(k-1)/k}(O(|z|^{M_4})+O(|A(z)|^{o(1)})) < 2^{-k}|A(z)|+|A(z)|^{(k-1)/k}O(|A(z)|^{o(1)}),$$

which is clearly impossible. Now suppose that $|h'(z)| > 2|A(z)|^{1/k}$. Then $h'(z)$ is large and (4.44), (4.47) and (4.48) yield

$$|h'(z)|^k < 2^{-k}|h'(z)|^k + |h'(z)|^{k-1}(O(|z|^{M_4}) + O(|A(z)|^{o(1)})) < 2^{-k}|h'(z)|^k + |h'(z)|^{k-1}O(|h'(z)|^{o(1)})$$

which is again impossible. Claim 3 is proved.

For large $r \in E_0$ and $z_0 \in a_r$ we may now define a branch of $A(z)^{1/k}$, analytic on $D(z_0, |z_0|^{-M_3})$, since A is large there and so in particular non-zero.

Claim 4: we have

$$h'(z) = c_{z_0}A(z)^{1/k} + O(r^{M_5}), \quad z \in D(z_0, |z_0|^{-M_3}). \tag{4.52}$$

Here the constant $c = c_{z_0}$ may depend on z_0 but satisfies $c^k = -1$.

To prove Claim 4 set $u(z) = h'(z)A(z)^{-1/k}$. Dividing (4.47) through by $A(z)$ and using (4.44), (4.48) and (4.51) we get

$$0 = u^k + O(r^{M_5}|A(z)|^{-1/k}) + 1 = u^k + 1 + o(1).$$

Since u is continuous on $D(z_0, |z_0|^{-M_3})$ there is a fixed c with $c^k = -1$ such that $u = c + o(1)$ on $D(z_0, |z_0|^{-M_3})$, and the binomial theorem gives

$$u = (-1 + O(r^{M_5}|A(z)|^{-1/k}))^{1/k} = c(1 + O(r^{M_5}|A(z)|^{-1/k}))$$

from which (4.52) follows on multiplying out by $A(z)^{1/k}$. This proves Claim 4.

For large $r \in E_0$ and $z_0 \in a_r$ we now set

$$f(z) = W(z)\exp\left(\int_{z_0}^{z} c_{z_0}A(t)^{1/k}dt\right), \quad \frac{f'(z)}{f(z)} = c_{z_0}A(z)^{1/k} + \frac{W'(z)}{W(z)}, \quad z \in D(z_0, |z_0|^{-M_3}). \tag{4.53}$$

By (4.45), (4.48) and (4.52) we have

$$w(z) = \frac{W'(z)}{W(z)} = O(r^{M_6}), \quad z \in D(z_0, |z_0|^{-M_3}). \tag{4.54}$$

Now (4.54) and Cauchy's estimate for derivatives give

$$w^{(j)}(z_0) = \frac{j!}{2\pi i} \int_{|z-z_0|=\frac{1}{2}|z_0|^{-M_3}} \frac{w(z)}{(z-z_0)^{j+1}} dz = O(r^{M_7}), \quad j = 1, \ldots, k,$$

and so we get

$$\frac{W^{(j)}(z_0)}{W(z_0)} = O(r^{M_8}), \quad j = 1, \ldots, k. \tag{4.55}$$

Also, writing

$$H(z) = \int_{z_0}^{z} c_{z_0} A(t)^{1/k} dt, \quad \frac{H''(z)}{H'(z)} = \frac{A'(z)}{kA(z)} \tag{4.56}$$

gives, using (4.48),

$$\frac{H^{(j)}(z_0)}{H'(z_0)} = O(r^{M_9}), \quad j = 1, \ldots, k. \tag{4.57}$$

Substituting $f = We^H$ into (4.42) gives at z_0 (compare (4.47))

$$(H')^k + k(H')^{k-1}\frac{W'}{W} + \frac{k(k-1)}{2}(H')^{k-2}H'' + T_{k-2}(H') +$$

$$+ \sum_{1 \le j \le k-2} A_j \left((H')^j + j(H')^{j-1}\frac{W'}{W} + \frac{j(j-1)}{2}(H')^{j-2}H'' + T_{j-2}(H') \right) + A_0 = 0,$$

which by (4.56) we may write in the form

$$k(H')^{k-1}\frac{W'}{W} + \frac{k(k-1)}{2}(H')^{k-2}H'' + O(r^{M_{10}})(H')^{k-2} = 0,$$

so that

$$\frac{W'}{W} = -\frac{k(k-1)}{2k}\frac{H''}{H'} + O(r^{-2}) = -\frac{k(k-1)}{2k^2}\frac{A'}{A} + O(r^{-2}),$$

using (4.56) again. Substituting this estimate into (4.53) we obtain (4.43) at z_0.

We show now that we may take the same branch of $A^{1/k}$ and the same k'th root c_r of -1 for all $z_0 \in a_r$. Suppose first that $\beta_r - \alpha_r < 2\pi$ in (4.39). Then we may define an analytic branch of $A(z)^{1/k}$ on a simply connected domain containing a_r, since $A(z)$ is large near a_r. Then we have, for each $z_0 \in a_r$, using (4.43), (4.44) and (4.48),

$$\frac{f'(z_0)}{f(z_0)A(z_0)^{1/k}} = c_{z_0} + o(1)$$

in which $c_{z_0}^k = -1$. Since the left hand side is continuous, we see that the root c_{z_0} is the same for all $z_0 \in a_r$.

Suppose finally that $0 = \alpha_r, \beta_r = 2\pi$. Then we take a small $\delta > 0$ and obtain (4.43) on $a'_r = \{z : |z| = r, 0 \le \arg z \le 2\pi - \delta\}$. Here $c = c_r$ does not depend on δ. As we then let $\delta \to 0+$ both sides of (4.43) are continued analytically around the circle $S(0,r)$ and since the left hand side is continuous and $A(z)$ is large on $S(0,r)$ it follows that $A(z)^{1/k}$ must return to the same branch of $A^{1/k}$ as we continue once around $S(0,r)$, since otherwise it would return to the original branch of $A^{1/k}$ multiplied by a constant $d \ne 1$ with $d^k = 1$.

Chapter 5

Asymptotics for matrix linear differential equations

In this chapter we discuss asymptotics for solutions of linear differential equations with rational coefficients, combining a slightly non-standard approach to the regular singular point case with methods from Wasow's and Balser's texts [4, 72].

5.1 Some facts from linear algebra

Lemma 5.1.1 *Let $A = (a_{jk})$ be a matrix and suppose that rows $j_1, \ldots, j_s$ of A are linearly independent. Then there exists pairwise distinct $k_1, \ldots, k_s$ with $a_{j_\mu k_\mu} \neq 0$ for each μ.*

Proof. It may be assumed that A has s rows and rank s and, by taking s linearly independent columns, that A is a square matrix, with $\det A \neq 0$. Now determine k_1 by choosing a non-zero entry in row 1 with non-zero minor, then delete row 1 and column k_1, and repeat. □

5.1.1 Nilpotent matrices

A $\nu \times \nu$ matrix A is called *nilpotent* if there exists $t \in \mathbb{N} = \{1, 2, \ldots\}$ with $A^t = (0)$, in which case 0 is the only eigenvalue of A, because $Ax = \lambda x$ gives $0 = A^t x = \lambda^t x$. Conversely, if 0 is the only eigenvalue of a $\nu \times \nu$ matrix B then the characteristic equation of B is just $\lambda^\nu = 0$, and so $B^\nu = (0)$ by the Cayley-Hamilton theorem. Thus if $A^t = (0)$ for some $t \in \mathbb{N}$ then $A^s = (0)$ for some $s \leq \nu$.

5.1.2 Upper triangular shifting matrices

The m-dimensional *(upper) triangular shifting matrix* N_m is the $m \times m$ square matrix with all entries 0, excepts for 1s immediately to the right of the main diagonal (i.e. $n_{jk} = 0$, except that $n_{jk} = 1$ if $k - j = 1$). For example,

$$N_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Left multiplication (of an $m \times n$ matrix) by N_m shifts every row up one place, and replaces the last row by 0s. Right multiplication (of an $n \times m$ matrix) by N_m shifts every column right one place, and replaces the first column by 0s.

Lemma 5.1.2 *Suppose that an $m \times m$ matrix*

$$B = \begin{pmatrix} a_1 & 1 & 0 & \dots & 0 \\ a_2 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ a_{m-1} & 0 & 0 & \dots & 1 \\ a_m & 0 & 0 & \dots & 0 \end{pmatrix} = A + N_m$$

is nilpotent, where columns 2 to m of A are all zero. Then $A = (0)$.

Proof. Since B is nilpotent, 0 is the only eigenvalue of B, and the characteristic equation of B can be written (with $\lambda = -x$)

$$\begin{aligned} 0 &= \det(B - \lambda I_m) = \begin{vmatrix} a_1 + x & 1 & 0 & \dots & 0 \\ a_2 & x & 1 & \dots & 0 \\ \vdots & & & & \\ a_{m-1} & 0 & \dots & x & 1 \\ a_m & 0 & 0 & \dots & x \end{vmatrix} \\ &= (x + a_1)x^{m-1} - a_2 x^{m-2} + a_3 x^{m-3} + \ldots \pm a_m = x^m. \end{aligned}$$

To see this, observe that each entry in column 1 of $B - \lambda I_m$ has minor of form $\begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$, where C is lower triangular with $1s$ on the main diagonal, and D is upper triangular with all diagonal entries x.

□

5.1.3 Direct sums

A block matrix

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & A_s \end{pmatrix}$$

is written $A = A_1 \oplus \ldots \oplus A_s$. Note that if $A = A_1 \oplus \ldots \oplus A_s$ and $B = B_1 \oplus \ldots \oplus B_s$ have blocks of matching sizes then $AB = A_1B_1 \oplus \ldots \oplus A_sB_s$.

Lemma 5.1.3 *Given a block matrix $A = A_1 \oplus \ldots \oplus A_s$ and any permutation $B_1, \ldots, B_s$ of $A_1, \ldots, A_s$, there is a similarity transformation $B = T^{-1}AT$ which produces $B = B_1 \oplus \ldots \oplus B_s$.*

Proof. The proof is by induction on s, and the blocks are interchanged by conjugation of matrices. First, if $s = 2$ and I_1 and I_2 are appropriately sized identity matrices then

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} 0 & I_1 \\ I_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & A_1 \\ A_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I_1 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} A_2 & 0 \\ 0 & A_1 \end{pmatrix}. \tag{5.1}$$

Thus, if $s \geq 3$ and $B_1 = A_p$, where $1 < p \leq s$, then the above method for $s = 2$ turns $A = A_1 \oplus \ldots \oplus A_s$ into $C = A_p \oplus \ldots \oplus A_s \oplus A_1 \ldots \oplus A_{p-1} = B_1 \oplus \ldots \oplus A_s \oplus A_1 \ldots \oplus A_{p-1}$. It remains only to note that if conjugation by T turns $D_1 \oplus \ldots \oplus D_{s-1}$ into $E_1 \oplus \ldots \oplus E_{s-1}$ then conjugation by a matrix of form

$$\begin{pmatrix} I & 0 \\ 0 & T \end{pmatrix}$$

turns $F \oplus D_1 \oplus \ldots \oplus D_{s-1}$ into $F \oplus E_1 \oplus \ldots \oplus E_{s-1}$.

□

5.1.4 Jordan form

A square matrix of form $\lambda I + N$, where $\lambda \in \mathbb{C}$ and N is an upper triangular shifting matrix, is called an upper Jordan block (or just Jordan block). A Jordan matrix is a block matrix of form

$$J = \begin{pmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & J_s \end{pmatrix}, \quad J_k = \lambda_k I_{m_k} + N_{m_k}.$$

This is expressed as a direct sum

$$J = J_1 \oplus J_2 \oplus \ldots \oplus J_s, \quad J^p = J_1^p \oplus J_2^p \oplus \ldots \oplus J_s^p \quad (p \in \mathbb{N}). \tag{5.2}$$

Every square matrix A is similar (via a conjugation $A = S^{-1}JS$) to a Jordan matrix J.

Lemma 5.1.4 *Let A be an $n \times n$ matrix. Then A has n linearly independent "generalised eigenvectors" w_j each with the property that $(A - \lambda_j I_n)^{p_j} w_j = 0$ for some $p_j \in \mathbb{N}$ and eigenvalue λ_j of A.*

Proof. Suppose first that $A = \lambda I_n + N$, where $N = N_n$ is the $n \times n$ upper triangular shifting matrix in §5.1.2. Then $N^n = (0)$, and so $(A - \lambda I_n)^n x = 0$ for every n-dimensional column vector x.

Now suppose that $A = A_1 \oplus \ldots \oplus A_s$, with each A_j of form $A = \lambda_j I_{\mu_j} + N_{\mu_j}$. Take any vector w such that its first $\mu_1 + \ldots + \mu_{j-1}$ and last $\mu_{j+1} + \ldots + \mu_s$ entries are all 0. Since

$$\begin{aligned} (A - \lambda_j I_n)^{\mu_j} &= (A_1 - \lambda_j I_{\mu_1})^{\mu_j} \oplus \ldots \oplus (A_j - \lambda_j I_{\mu_j})^{\mu_j} \oplus \ldots \oplus (A_s - \lambda_j I_{\mu_s})^{\mu_j} \\ &= (A_1 - \lambda_j I_{\mu_1})^{\mu_j} \oplus \ldots \oplus (0) \oplus \ldots \oplus (A_s - \lambda_j I_{\mu_s})^{\mu_j} \end{aligned}$$

we have $(A - \lambda_j I_n)^{\mu_j} w = 0$. Thus each A_j gives rise to μ_j vectors w with $(A - \lambda_j I_n)^{\mu_j} w = 0$, and the collection of all of these is linearly independent.

In the general case, choose an invertible matrix P such that $B = P^{-1}AP$ is in Jordan form. Then $Bx = \lambda x$ if and only if $A(Px) = PBx = P(\lambda x) = \lambda Px$. Thus B has the same eigenvalues as A. By the previous paragraph there exist n linearly independent vectors v_j each with the property that $(B - \lambda_j I_n)^{p_j} v_j = 0$ for some $p_j \in \mathbb{N}$ and eigenvalue λ_j of B (and hence of A). Now

$$(A - \lambda_j I_n)^{p_j} P v_j = PP^{-1}(A - \lambda_j I_n)^{p_j} P v_j = P(P^{-1}AP - \lambda_j I_n)^{p_j} v_j = 0.$$

□

5.2 Some basic facts from matrix analysis

For vectors $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$ in $\mathbb{C}^n$ write

$$\langle a, b \rangle = \sum_{j=1}^n a_j \bar{b}_j = \overline{\langle b, a \rangle}, \quad \|a\| = \sqrt{\langle a, a \rangle} = \sqrt{\sum_{j=1}^n |a_j|^2}.$$

The Cauchy-Schwarz inequality then reads $|\langle a, b \rangle| \le \|a\| \cdot \|b\|$: to prove this assume without loss of generality that $\langle a, b \rangle$ is real and positive and write, for $t \in \mathbb{R}$,

$$0 \le \langle a + tb, a + tb \rangle = \|a\|^2 + t(\langle a, b \rangle + \langle b, a \rangle) + t^2\|b\|^2 = \|a\|^2 + 2t\langle a, b \rangle + t^2\|b\|^2 = At^2 + 2Bt + C$$

so that $B^2 \leq AC$. The triangle inequality $\|a+b\| \leq \|a\| + \|b\|$ then follows via

$$\|a+b\|^2 = \|a\|^2 + \langle a, b\rangle + \langle b, a\rangle + \|b\|^2 \leq \|a\|^2 + 2\|a\| \cdot \|b\| + \|b\|^2 = (\|a\| + \|b\|)^2,$$

and this extends by induction to finite sums. For a positive measure μ on a space Y and a simple function $f = \sum_j a_j \chi_{Y_j} : Y \to \mathbb{C}^n$, the triangle inequality leads to

$$\left\| \int_Y f \, d\mu \right\| = \left\| \sum_j a_j \mu(Y_j) \right\| \leq \sum_j \|a_j\| \, \mu(Y_j) = \int_Y \|f\| \, d\mu,$$

so that

$$\left\| \int_Y f \, d\mu \right\| \leq \int_Y \|f\| \, d\mu \tag{5.3}$$

for integrable $f : Y \to \mathbb{C}^n$.

If A is an $n \times n$ matrix (a_{jk}), then the Frobenius norm of A is defined by

$$\|A\| = \|A\|_{\mathcal{F}} = \sqrt{\sum_{jk} |a_{jk}|^2}.$$

This is the same as the $\mathbb{C}^{n^2}$ norm of the n^2-dimensional vector obtained by writing out the entries of A, and $\|A\|_{\mathcal{F}}^2$ is the sum of the squares of the $\mathbb{C}^n$ norms of the rows (or columns) of A. Hence (5.3) holds for matrix-valued f with the Frobenius norm. For a matrix product $C = AB$, the Cauchy-Schwarz inequality gives (with all sums from 1 to n)

$$|c_{jk}|^2 = \left| \sum_r a_{jr} b_{rk} \right|^2 \leq \sum_r |a_{jr}|^2 \cdot \sum_r |b_{rk}|^2$$

and so

$$\sum_k |c_{jk}|^2 \leq \sum_r |a_{jr}|^2 \cdot \sum_{r,k} |b_{rk}|^2 = \sum_r |a_{jr}|^2 \cdot \|B\|_{\mathcal{F}}^2$$

and

$$\|C\|_{\mathcal{F}}^2 = \sum_{j,k} |c_{jk}|^2 \leq \sum_{j,r} |a_{jr}|^2 \cdot \|B\|_{\mathcal{F}}^2 = \|A\|_{\mathcal{F}}^2 \cdot \|B\|_{\mathcal{F}}^2.$$

Thus the Frobenius norm is submultiplicative.

5.2.1 The exponential and logarithm of a matrix

If A is a square matrix then

$$\exp(A) = \sum_{m=0}^{\infty} \frac{A^m}{n!},$$

this being convergent, with norm at most $\exp(\|A\|)$. If A and B commute, i.e. $AB = BA$, then $\exp(A+B) = \exp(A)\exp(B) = \exp(B)\exp(A)$, and so $\exp(-A)$ is the inverse of $\exp(A)$.

If $A(z)$ is a holomorphic matrix and $A(z)$ commutes with $A'(z)$, which is always the case if $A(z)$ is a holomorphic diagonal matrix, then

$$\frac{d}{dz}\left(\exp(A(z))\right) = A'(z)\exp(A(z)) = \exp(A(z))A'(z).$$

If F is a constant square matrix, then $z^F = \exp(F \log z)$, and continuing this matrix function once counter-clockwise around the origin multiplies it by $\exp(2\pi i F)$. If F is nilpotent, then the entries of z^F are polynomials in $\log z$. For example, the notation of §5.1.2 gives

$$\begin{aligned} z^{N_4} &= I_4 + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \log z + \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{(\log z)^2}{2} + \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{(\log z)^3}{6} \\ &= \begin{pmatrix} 1 & \log z & (1/2)(\log z)^2 & (1/6)(\log z)^3 \\ 0 & 1 & \log z & (1/2)(\log z)^2 \\ 0 & 0 & 1 & \log z \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

and

$$z^{\lambda I_4 + N_4} = z^{\lambda I_4} z^{N_4} = \begin{pmatrix} z^\lambda & z^\lambda \log z & (1/2) z^\lambda (\log z)^2 & (1/6) z^\lambda (\log z)^3 \\ 0 & z^\lambda & z^\lambda \log z & (1/2) z^\lambda (\log z)^2 \\ 0 & 0 & z^\lambda & z^\lambda \log z \\ 0 & 0 & 0 & z^\lambda \end{pmatrix}.$$

Lemma 5.2.1 *Let A be a $\mu \times \mu$ nilpotent matrix. Then there exists a $\mu \times \mu$ matrix D with $\exp(D) = I - A$.*

Proof. Since A is nilpotent we have $A^\mu = (0)$. For $t \in \mathbb{C}$ write $I = I_\mu$ and

$$B(t) = \sum_{m=1}^{\mu-1} \frac{1}{m} (tA)^m, \quad B'(t) = A \sum_{m=1}^{\mu-1} (tA)^{m-1},$$

as well as

$$(I - tA) B'(t) = A(I - tA)(I + tA + \ldots + (tA)^{\mu-2}) = A(I - (tA)^{\mu-1}) = A.$$

This gives, since the matrices $B'(t)$ and $B(t)$ commute,

$$\begin{aligned} (I - tA) \frac{d}{dt} (\exp(B(t)) &= (I - tA) \sum_{m=0}^{\infty} \frac{m B'(t) B(t)^{m-1}}{m!} \\ &= A \sum_{m=1}^{\infty} \frac{B(t)^{m-1}}{(m-1)!} = A \exp(B(t)). \end{aligned}$$

Now write

$$C(t) = (I - tA) \exp(B(t)), \quad C'(t) = -A \exp(B(t)) + A \exp(B(t)) = (0),$$

so that $C(t)$ is constant, with $C(0) = \exp(B(0)) = \exp((0)) = I$. Hence $\exp(-B(t)) = I - tA$ and the result follows with $D = -B(1)$.

$\square$

Lemma 5.2.2 *Let $H = \lambda I_\mu + N_\mu$ be a $\mu \times \mu$ Jordan block, with $\lambda \in \mathbb{C} \setminus \{0\}$ and N_μ a shifting matrix. Then there exists a $\mu \times \mu$ matrix B with $\exp(B) = H$.*

Proof. Choose $b \in \mathbb{C}$ with $e^b = \lambda$. Then $\exp(bI_\mu) = \lambda I_\mu$. Now let $K = -\lambda^{-1}N_\mu$; then $K^\mu = (0)$, and Lemma 5.2.1 gives a matrix M with $\exp(M) = I_\mu - K = I_\mu + \lambda^{-1}N_\mu$. This gives, because the matrices bI_μ and M commute,

$$\exp(bI_\mu + M) = \lambda I_\mu(I_\mu + \lambda^{-1}N_\mu) = \lambda I_\mu + N_\mu = H.$$

□

Lemma 5.2.3 *Let B be a non-singular matrix in Jordan form. Then there exists a matrix C with $\exp(C) = B$.*

Proof. Write $B = H_1 \oplus \ldots \oplus H_s$, where each H_j is as in Lemma 5.2.2, and use the fact that $\exp(C_1 \oplus \ldots \oplus C_s) = \exp(C_1) \oplus \ldots \oplus \exp(C_s)$.

□

Lemma 5.2.4 *Let B be a non-singular matrix. Then there exists a matrix C with $\exp(C) = B$.*

Proof. Write $B = P^{-1}DP$, where D is a non-singular matrix in Jordan form, and use Lemma 5.2.3 to choose E with $\exp(E) = D$. Then $\exp(P^{-1}EP) = P^{-1}DP = B$.

□

Lemma 5.2.5 *Let $B = (b_{jk})$ be a square matrix and $c \in \mathbb{C}$. Then $\exp(cB)$ has determinant $\exp(c\,\mathrm{tr}\,B)$, where $\mathrm{tr}\,B = \sum_j b_{jj}$.*

Proof. If $B = (0)$ this is obvious, and if B is in (upper triangular) Jordan form then $\exp(cB)$ is an upper triangular matrix whose diagonal entries are the exponentials of the diagonal entries of cB. In the general case write $B = P^{-1}DP$, where D is in Jordan form, and

$$\det\left(\exp(cB)\right) = \det\left(\exp(P^{-1}cDP)\right) = \det\left(P^{-1}\exp(cD)P\right) = \exp(c\,\mathrm{tr}\,D).$$

But if $E = (e_{jk})$ and $F = (f_{jk})$ are square matrices of the same size then

$$\mathrm{tr}\,(EF) = \sum_j\left(\sum_k e_{jk}f_{kj}\right) = \sum_k\left(\sum_j f_{kj}e_{jk}\right) = \mathrm{tr}\,(FE),$$

which gives

$$\mathrm{tr}\,(B) = \mathrm{tr}\,(P^{-1}DP) = \mathrm{tr}\,(PP^{-1}D) = \mathrm{tr}\,D.$$

□

5.2.2 A hierarchy of nilpotent matrices

Let A and B be $\nu \times \nu$ nilpotent matrices. Following [4], the matrix B is called *superior* to A if $\mathrm{rank}\,A^l \leq \mathrm{rank}\,B^l$ for every $l \in \mathbb{N}$ and there exists $m \in \mathbb{N}$ with $\mathrm{rank}\,A^m < \mathrm{rank}\,B^m$.

Since $A^\nu = B^\nu = (0)$ it must be the case that $m \leq \nu - 1$, and because there are only ν possible values for the rank (namely 0 to $\nu - 1$) it is not possible to have arbitrarily long sequences A_j of $\nu \times \nu$ nilpotent matrices such that A_{j+1} is superior to A_j. To see this, write the ranks of the powers A_j^m, for $m = 1, \ldots, \nu - 1$, as $(r_{j,1}, \ldots, r_{j,\nu-1})$. Then $r_{j,l} \leq r_{j+1,l}$, with strict inequality for at least one l, so the number of matrices in such a chain is at most $1+(\nu-1)^2$, because the $\nu-1$ entries $r_{j,k}$ can each increase at most $\nu-1$ times.

For example, if

$$C=\begin{pmatrix} 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}=N_3\oplus N_2, \quad D=\begin{pmatrix} 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}=N_4\oplus N_1,$$

then C and D have rank 3, while

$$C^2=\begin{pmatrix} 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad D^2=\begin{pmatrix} 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

so C^2 has rank 1, while D^2 has rank 2, and D is superior to C.

Note that if E is similar to A, and F is similar to B, while B is superior to A, then F is superior to E, because E^l and A^l have the same rank for every $l\in\mathbb{N}$, as have F^l and B^l.

Lemma 5.2.6 *Let A_0 and B_0 be $\nu\times\nu$ matrices given by*

$$A_0=\begin{pmatrix} M_1 & 0 & \dots & 0 & 0\\ 0 & M_2 & \dots & 0 & 0\\ \dots & \dots & \dots & \dots & \dots\\ 0 & \dots & 0 & M_\tau & 0\\ 0 & \dots & 0 & 0 & M \end{pmatrix}, \quad B_0=\begin{pmatrix} M_1 & 0 & \dots & 0 & 0\\ 0 & M_2 & \dots & 0 & 0\\ \dots & \dots & \dots & \dots & \dots\\ C_1 & C_2 & \dots & M_\tau & 0\\ 0 & \dots & 0 & 0 & M \end{pmatrix}, \tag{5.4}$$

in which the following conditions all hold:
the M_j are upper triangular shifting matrices of dimension s_j, where $s_1\geq\ldots\geq s_\tau$;
the last block M satisfies $M^\nu=(0)$;
all columns, bar possibly the first, of each block C_j vanish;
at least one C_j is not the zero matrix.

Then B_0 is superior to A_0, but is nilpotent.

Here we allow for the case that M is 0×0, so that the blocks above and to the immediate left of M do not appear.

Proof. We first show by induction that (5.4) yields representations

$$A_0^l=\begin{pmatrix} M_1^l & 0 & \dots & 0 & 0\\ 0 & M_2^l & \dots & 0 & 0\\ \dots & \dots & \dots & \dots & \dots\\ 0 & \dots & 0 & M_\tau^l & 0\\ 0 & \dots & 0 & 0 & M^l \end{pmatrix}, \quad B_0^l=\begin{pmatrix} M_1^l & 0 & \dots & 0 & 0\\ 0 & M_2^l & \dots & 0 & 0\\ \dots & \dots & \dots & \dots & \dots\\ C_1^{(l)} & C_2^{(l)} & \dots & M_\tau^l & 0\\ 0 & \dots & 0 & 0 & M^l \end{pmatrix} \tag{5.5}$$

for $l\in\mathbb{N}$. Only the formula for B_0^l needs proof, and it is clearly true for $l=1$, with $C_k^{(l)}=C_k$.

Assuming the result for some $l \in \mathbb{N}$ gives

$$B_0^{l+1} = \begin{pmatrix} M_1 & 0 & \dots & 0 & 0 \\ 0 & M_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ C_1 & C_2 & \dots & M_\tau & 0 \\ 0 & \dots & 0 & 0 & M \end{pmatrix} \begin{pmatrix} M_1^l & 0 & \dots & 0 & 0 \\ 0 & M_2^l & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ C_1^{(l)} & C_2^{(l)} & \dots & M_\tau^l & 0 \\ 0 & \dots & 0 & 0 & M^l \end{pmatrix}$$

and thus (5.5) is proved for $l+1$, with

$$C_k^{(l+1)} = C_k M_k^l + M_\tau C_k^{(l)}.$$

Since the M_j are all nilpotent, the matrix $D_0 = B_0^\nu$ is zero on and above the diagonal, so that the only eigenvalue of D_0 is 0. Thus D_0 is nilpotent and so is B_0.

Note next that each C_k is an $s_\tau \times s_k$ matrix. We now claim that for $1 \leq l \leq s_k$ the lth column of $C_k^{(l)}$ is the first column of C_k, and that all columns of $C_k^{(l)}$ from the $(l+1)$th onwards are zero. Again this is clear for $l = 1$, and assuming it true for some $l \in \{1, \dots, s_k - 1\}$ gives the following. First, postmultiplying by M_k^l shifts columns right l places, so the $(l+1)$th column of $C_k M_k^l$ is the first column of C_k and all other columns of $C_k M_k^l$ vanish. Second, all columns of $M_\tau C_k^{(l)}$ from the $(l+1)$th onwards are zero, because this is true of $C_k^{(l)}$. This proves the claim.

Consider now the pth column of A_0^l, where $p \leq s_1 + \dots + s_\tau$, and assume that this column of A_0^l is not the zero vector. This column then has exactly one non-zero entry, a 1 lying in M_k^l for some $k \leq \tau$; moreover, this 1 must lie in at least the $(l+1)$th column of M_k^l, and it must be the case that $l+1 \leq s_k$. We claim that this column of A_0^l is the same as the corresponding column of B_0^l, this being obvious from (5.5) if $k = \tau$, while if $k < \tau$ then the corresponding column of $C_k^{(l)}$ is zero. Thus rank $A_0^l \leq$ rank B_0^l for every $l \in \mathbb{N}$.

Now observe that, by (5.5),

$$A_0^{s_\tau} = \begin{pmatrix} M_1^{s_\tau} & 0 & \dots & 0 & 0 \\ 0 & M_2^{s_\tau} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & M^{s_\tau} \end{pmatrix}, \quad B_0^{s_\tau} = \begin{pmatrix} M_1^{s_\tau} & 0 & \dots & 0 & 0 \\ 0 & M_2^{s_\tau} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ C_1^{(s_\tau)} & C_2^{(s_\tau)} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & M^{s_\tau} \end{pmatrix}$$

There is at least one k with $1 \leq k \leq \tau - 1$ for which the first column of C_k does not vanish: since $s_\tau \leq s_k$ this column is then the s_τth column of $C_k^{(s_\tau)}$, and so at least one column of $B_0^{s_\tau}$ is not a linear combination of columns of $A_0^{s_\tau}$. Therefore rank $A_0^{s_\tau} <$ rank $B_0^{s_\tau}$ and the lemma is proved. □

There is a companion version for rows, in which we again permit the case where M is 0×0.

Lemma 5.2.7 *Let A_0 and B_0 be $\nu \times \nu$ matrices given by*

$$A_0 = \begin{pmatrix} M_1 & 0 & \dots & 0 & 0 \\ 0 & M_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & M_\tau & 0 \\ 0 & \dots & 0 & 0 & M \end{pmatrix}, \quad B_0 = \begin{pmatrix} M_1 & 0 & \dots & D_1 & 0 \\ 0 & M_2 & \dots & D_2 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & M_\tau & 0 \\ 0 & \dots & 0 & 0 & M \end{pmatrix}, \tag{5.6}$$

in which the following conditions all hold:
the M_j are upper triangular shifting matrices of dimension s_j, where $s_1 \geq \dots \geq s_\tau$;

the last block M has $M^\nu = (0)$;
all rows, bar possibly the last, of each block D_j vanish;
at least one D_j is not the zero matrix.

Then B_0 is nilpotent but superior to A_0.

Proof. This time (5.6) yields representations

$$A_0^l = \begin{pmatrix} M_1^l & 0 & \dots & 0 & 0 \\ 0 & M_2^l & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & M_\tau^l & 0 \\ 0 & \dots & 0 & 0 & M^l \end{pmatrix}, \quad B_0^l = \begin{pmatrix} M_1^l & 0 & \dots & D_1^{(l)} & 0 \\ 0 & M_2^l & \dots & D_2^{(l)} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & M_\tau^l & 0 \\ 0 & \dots & 0 & 0 & M^l \end{pmatrix} \tag{5.7}$$

for $l \in \mathbb{N}$. To check this write $D_k^{(1)} = D_k$ and

$$B_0^{l+1} = \begin{pmatrix} M_1^l & 0 & \dots & D_1^{(l)} & 0 \\ 0 & M_2^l & \dots & D_2^{(l)} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & M_\tau^l & 0 \\ 0 & \dots & 0 & 0 & M^l \end{pmatrix} \begin{pmatrix} M_1 & 0 & \dots & D_1 & 0 \\ 0 & M_2 & \dots & D_2 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & M_\tau & 0 \\ 0 & \dots & 0 & 0 & M \end{pmatrix}$$

so that the recurrence relation is

$$D_k^{(l+1)} = M_k^l D_k + + D_k^{(l)} M_\tau .$$

Here each D_k is an $s_k \times s_\tau$ matrix.

We now claim that for $1 \le l \le s_k$ the following holds for $D_k^{(l)}$: the lth row from the bottom is the last row of D_k, and all rows above it vanish. This is clear for $l = 1$, and assuming it true for some $l \in \{1, \dots, s_k - 1\}$ gives the following. First, premultiplying by M_k^l shifts rows up l places, so the $(l+1)$th row from the bottom of $M_k^l D_k$ is the last row of D_k, and all other rows of $M_k^l D_k$ vanish. Second, if we count from the bottom then all rows of $D_k^{(l)} M_\tau$ from the $(l+1)$th onwards are zero, because this is true of $D_k^{(l)}$. This proves the claim.

Consider now the pth row of A_0^l, where $p \le s_1 + \dots + s_\tau$, and assume that this row of A_0^l is not the zero vector. This row then has exactly one non-zero entry, a 1 lying in M_k^l for some $k \le \tau$. This 1 must lie in at least the $(l+1)$th row from the bottom of M_k^l, and we must have $l+1 \le s_k$. Again we assert that this row of A_0^l is the same as the corresponding row of B_0^l, this being obvious if $k = \tau$, while if $k < \tau$ then the corresponding row of $D_k^{(l)}$ is zero. Thus we see that $\operatorname{rank} A_0^l \le \operatorname{rank} B_0^l$ for every $l \in \mathbb{N}$.

Now observe that

$$A_0^{s_\tau} = \begin{pmatrix} M_1^{s_\tau} & 0 & \dots & 0 & 0 \\ 0 & M_2^{s_\tau} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & M^{s_\tau} \end{pmatrix}, \quad B_0^{s_\tau} = \begin{pmatrix} M_1^{s_\tau} & 0 & \dots & D_1^{(s_\tau)} & 0 \\ 0 & M_2^{s_\tau} & \dots & D_2^{(s_\tau)} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & M^{s_\tau} \end{pmatrix}$$

There is at least one k with $1 \le k \le \tau - 1$ for which the last row of D_k does not vanish: since $s_\tau \le s_k$, this row is then the s_τth row from the bottom of $D_k^{(s_\tau)}$, and so at least one row of $B_0^{s_\tau}$ is not a linear combination of rows of $A_0^{s_\tau}$. Therefore $\operatorname{rank} A_0^{s_\tau} < \operatorname{rank} B_0^{s_\tau}$ and the lemma is proved. □

5.2.3 The solution of certain equations

Lemma 5.2.8 *Let P and Q be upper triangular shifting matrices, of dimensions p and q respectively, let C be a given $p \times q$ matrix, and consider the equation*

$$PX - XQ = C - B. \tag{5.8}$$

Then there exists a $p \times q$ matrix B such that (5.8) has a $p \times q$ solution X. This may be done so that one of the following holds:
(i) all columns of B are zero, bar possibly the first, and the last column of X is zero;
(ii) all rows of B are zero, bar possibly the last, and the first row of X is zero.

Note that in the subsequent application of Lemma 5.2.8 we do not use the conclusions regarding the columns/rows of X, only those involving B.

Proof. Premultiplying by P moves rows of X up one place, and replaces the last row by 0s. Similarly, postmultiplying by Q moves columns one place right, replacing the first by 0s. Thus (5.8) may be written in case (i) in the form

$$\begin{aligned}
Y &= PX - XQ \\
&= \begin{pmatrix} x_{2,1} & x_{2,2} & \dots & x_{2,q-1} & x_{2,q} \\ \dots & \dots & \dots & \dots & \dots \\ x_{p-1,1} & x_{p-1,2} & \dots & x_{p-1,q-1} & x_{p-1,q} \\ x_{p,1} & x_{p,2} & \dots & x_{p,q-1} & x_{p,q} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & x_{1,1} & x_{1,2} & \dots & x_{1,q-1} \\ 0 & x_{2,1} & x_{2,2} & \dots & x_{2,q-1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & x_{p-1,1} & x_{p-1,2} & \dots & x_{p-1,q-1} \\ 0 & x_{p,1} & x_{p,2} & \dots & x_{p,q-1} \end{pmatrix} \\
&= \begin{pmatrix} c_{1,1} - b_{1,1} & c_{1,2} & \dots & c_{1,q-1} & c_{1,q} \\ c_{2,1} - b_{2,1} & c_{2,2} & \dots & c_{2,q-1} & c_{2,q} \\ \dots & \dots & \dots & \dots & \dots \\ c_{p-1,1} - b_{p-1,1} & c_{p-1,2} & \dots & c_{p-1,q-1} & c_{p-1,q} \\ c_{p,1} - b_{p,1} & c_{p,2} & \dots & c_{p,q-1} & c_{p,q} \end{pmatrix}.
\end{aligned} \tag{5.9}$$

Consider the last rows in (5.9); we see that we need $b_{p,1} = c_{p,1}$; thus $x_{p,1}$ up to $x_{p,q-1}$ are now determined, and we set $x_{p,q} = 0$. Thus the last row of X has been determined. Now looking at the penultimate row in both sides shows that we need $c_{p-1,1} - b_{p-1,1}$ to equal $x_{p,1}$, which has already been determined. This gives us $b_{p-1,1}$ and the penultimate row of X, with the stipulation that its last entry be 0. The rows of X are thus determined moving upwards: once the kth row of X is known, we need $c_{k-1,1} - b_{k-1,1} = x_{k,1}$, and we can determine $x_{k-1,1}, \dots, x_{k-1,q-1}$ and set $x_{k-1,q} = 0$.

Now consider case (ii); here (5.8) may be written as

$$\begin{aligned}
Y &= PX - XQ \\
&= \begin{pmatrix} x_{2,1} & x_{2,2} & \dots & x_{2,q-1} & x_{2,q} \\ \dots & \dots & \dots & \dots & \dots \\ x_{p-1,1} & x_{p-1,2} & \dots & x_{p-1,q-1} & x_{p-1,q} \\ x_{p,1} & x_{p,2} & \dots & x_{p,q-1} & x_{p,q} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & x_{1,1} & x_{1,2} & \dots & x_{1,q-1} \\ 0 & x_{2,1} & x_{2,2} & \dots & x_{2,q-1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & x_{p-1,1} & x_{p-1,2} & \dots & x_{p-1,q-1} \\ 0 & x_{p,1} & x_{p,2} & \dots & x_{p,q-1} \end{pmatrix} \\
&= \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,q-1} & c_{1,q} \\ c_{2,1} & c_{2,2} & \dots & c_{2,q-1} & c_{2,q} \\ \dots & \dots & \dots & \dots & \dots \\ c_{p-1,1} & c_{p-1,2} & \dots & c_{p-1,q-1} & c_{p-1,q} \\ c_{p,1} - b_{p,1} & c_{p,2} - b_{p,2} & \dots & c_{p,q-1} - b_{p,q-1} & c_{p,q} - b_{p,q} \end{pmatrix}.
\end{aligned} \tag{5.10}$$

Comparing the first columns in (5.10) we see that we need $b_{p,1} = c_{p.1}$. Now $x_{2,1}$ up to $x_{p,1}$ are determined, and we set $x_{1,1} = 0$. Thus the first column of X has been determined. Now looking at the second columns shows that we need $c_{p,2} - b_{p,2}$ to equal $-x_{p,1}$, which has already been determined. This then gives us the second column of X (with the stipulation that its first entry be 0). The columns of X are thus determined moving rightwards: once the $(k-1)$th column of X is known, we need $c_{p,k} - b_{p,k} = -x_{p,k-1}$, and we can determine $x_{2,k}, \ldots, x_{p,k}$ and set $x_{1,k} = 0$.

□

Comment. Balser [4] imposes conditions on the dimensions of P and Q and states in passing that these are required to ensure uniqueness. For the *existence* of a solution as in (i) or (ii) the dimensions p and q can be arbitrary.

In particular, if $p = 1$ then cases (i) and (ii) require, respectively,

$$-\begin{pmatrix} 0 & x_{1,1} & x_{1,2} & \ldots & x_{1,q-1} \end{pmatrix} = \begin{pmatrix} c_{1,1} - b_{1,1} & c_{1,2} & \ldots & c_{1,q-1} & c_{1,q} \end{pmatrix}, \quad x_{1,q} = 0,$$

and

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \end{pmatrix} = \begin{pmatrix} c_{1,1} - b_{1,1} & c_{1,2} - b_{1,2} & \ldots & c_{1,q-1} - b_{1,q-1} & c_{1,q} - b_{1,q} \end{pmatrix},$$

both of which are plainly solvable.

Similarly, when $q = 1$ the required equations for cases (i) and (ii) are, respectively,

$$(i) \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c_{1,1} - b_{1,1} \\ c_{2,1} - b_{2,1} \\ \ldots \\ c_{p-1,1} - b_{p-1,1} \\ c_{p,1} - b_{p,1} \end{pmatrix}, \quad (ii) \begin{pmatrix} x_{2,1} \\ \ldots \\ x_{p-1,1} \\ x_{p,1} \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c_{1,1} \\ c_{2,1} \\ \ldots \\ c_{p-1,1} \\ c_{p,1} - b_{p,1} \end{pmatrix},$$

and these are obviously solvable. □

Lemma 5.2.9 *Let A and B be square matrices, where A is $m \times m$ and B is $n \times n$. Then the equation*

$$AX - XB = (0) \tag{5.11}$$

has a unique $m \times n$ solution X if and only if A and B have no common eigenvalue.

Now let C be an $m \times n$ matrix. If A and B have no common eigenvalue then the equation

$$AX - XB = C \tag{5.12}$$

has an $m \times n$ solution X, and this solution is unique.

Proof. Obviously one solution to (5.11) is to make X be the $m \times n$ zero matrix. Suppose A and B share the eigenvalue λ. Then so do A and the transpose B^T (because $B^T - \lambda I = (B - \lambda I)^T$ has determinant 0), and there exist non-zero column vectors v, w with $Av = \lambda v$ and $B^T w = \lambda w$, so $w^T B = \lambda w^T$. The matrix $X = v \cdot w^T$ is $m \times 1 \times 1 \times n$ and so $m \times n$, and

$$X \neq (0), \quad AX - XB = Av \cdot w^T - v \cdot w^T B = \lambda v \cdot w^T - v \cdot \lambda w^T = (0).$$

Now suppose that A and B share no eigenvalues, and that (5.11) has a solution X. Then $A^m X = XB^m$ for every integer $m \geq 0$. Thus

$$(A - \lambda I_m)^p X = X(B - \lambda I_n)^p$$

for every $\lambda \in \mathbb{C}$ and $p \in \mathbb{N}$. The matrix B has n linearly independent vectors w_j each with the property that $(B - \lambda_j I_n)^{p_j} w_j = 0$ for some $p_j \in \mathbb{N}$ and eigenvalue λ_j of B, and each of these satisfies

$$(A - \lambda_j I_m)^{p_j} X w_j = X(B - \lambda_j I_n)^{p_j} w_j = 0.$$

Since $\det(A - \lambda_j I_m) \neq 0$, this forces $X w_j = 0$ for each j, and so X annihilates every n-dimensional column vector and is the zero matrix.

Next, form an mn-dimensional column vector E by writing the columns of C one after another, and let Y be formed from X in matching fashion. Each entry of C is a linear combination of entries from X, with coefficients which are entries of A and B. Thus the equation (5.12) can be written in the form $DY = E$, where D is a square matrix. If A, B have no common eigenvalue, then the equation $DY = 0$ has no non-trivial solution, by the first part. Hence D is non-singular and $DY = E$ has a solution, which is then unique.

□

5.3 A class of formal expressions

Let $p \in \mathbb{N}$; then a *formal series in descending powers* of $z^{1/p}$ will mean a series $v(z) = \sum_{n\in\mathbb{Z}} a_n z^{n/p}$, with the $a_n \in \mathbb{C}$ and $a_n = 0$ for all but finitely many positive n. Let $\mathcal{V} = \mathcal{V}_p$ be the collection of these formal series.

Two elements $a = \sum_{n\in\mathbb{Z}} a_n z^{n/p}$ and $b = \sum_{n\in\mathbb{Z}} b_n z^{n/p}$ of $\mathcal{V}_p$ are equal if and only if $a_n = b_n$ for every n. The product ab is determined by multiplying term by term and gathering up like powers. Thus the set $\mathcal{V}$ forms a field, since if $v(z)$ is not the zero series then $1/v(z)$ can be computed formally by writing

$$v(z) = a_n z^{n/p} + a_{n-1} z^{(n-1)/p} + \ldots, \quad a_n \neq 0, \quad \frac{1}{v(z)} = a_n^{-1} z^{-n/p}(1 + a_{n-1}/a_n z^{1/p} + \ldots)^{-1}.$$

It follows that a square matrix with entries in $\mathcal{V}$ has an inverse matrix with entries in $\mathcal{V}$ if and only if its determinant is not the zero series.

Lemma 5.3.1 *Let n be a positive integer. Then the powers $(\log z)^m$, $m = 0, \ldots, n$, of the formal logarithm are linearly independent over $\mathcal{V}$.*

Proof. Suppose that we have a formal identity

$$\sum_{m=0}^{n} a_m(z)(\log z)^m = 0,$$

in which the coefficients $a_m(z)$ belong to $\mathcal{V}$ and do not all vanish. It may be assumed that $a_n(z)$ is not the zero series and that n is the least positive integer for which such an identity holds. Formally differentiating then gives

$$\sum_{m=0}^{n-1} b_m(z)(\log z)^m = 0, \quad b_m(z) \in \mathcal{V}, \quad b_{n-1}(z) = \frac{n}{z} + a'_{n-1}(z).$$

Since $b_{n-1}(z)$ cannot be the zero series, this contradicts the minimality of n. □

Lemma 5.3.1 motivates the following definition. Let $\mathcal{W}$ be the collection of polynomials in the formal logarithm $\log z$ with coefficients in $\mathcal{V}$, that is, sums $\sum_{n=0}^{\infty} a_n(z)(\log z)^n$, where $a_n(z) \in \mathcal{V}$ and all but finitely many a_n vanish. Two elements of $\mathcal{W}$ are the same if and only if they have the same coefficients.

Lemma 5.3.2 *Suppose that we have a formal identity*

$$\sum_{j=1}^{Q} P_j(z) \sum_{m=0}^{n_j} V_{j,m}(z)(\log z)^m = 0, \quad P_j(z) = e^{q_j(z)} z^{d_j},$$

in which: each $q_j(z)$ is a polynomial in $z^{1/p}$ and each d_j is a complex number; each $V_{j,m}(z)$ belongs to $\mathcal{V}$; if $j \neq k$ then either $q_j - q_k$ is non-constant or $p(d_j - d_k) \notin \mathbb{Z}$. Then $V_{j,m}(z)$ is the zero series for each j and m.

Proof. Suppose that we have such an identity, in which not all the $V_{j,m}$ vanish. It may be assumed that each V_{j,n_j} is not the zero series, while $P_Q = V_{Q,n_Q} = 1$ and $R = \sum_{j=1}^{Q}(1+n_j)$ is minimal. Formal differentiation yields

$$\begin{aligned} 0 &= (n_Q/z)(\log z)^{n_Q-1} + \sum_{m=0}^{n_Q-1} (V'_{Q,m}(z)(\log z)^m + (m/z)V_{Q,m}(z)(\log z)^{m-1}) + \\ &+ \sum_{j=1}^{Q-1} P_j(z) \sum_{m=0}^{n_j} ((V'_{j,m}(z) + (q'_j(z) + d_j/z)V_{j,m}(z))(\log z)^m + (m/z)V_{j,m}(z)(\log z)^{m-1}). \end{aligned}$$

If $n_Q > 0$ then the minimality of R forces $n_Q/z + V'_{Q,n_Q-1}(z) = 0$, which is impossible. Hence we have $n_Q = 0$ and so $Q > 1$. Moreover, again since R is minimal, we get

$$0 = V'_{j,n_j}(z)/V_{j,n_j}(z) + q'_j(z) + d_j/z$$

for $1 \leq j < Q$. Expanding out $V'_{j,n_j}(z)/V_{j,n_j}(z)$ in a formal series in descending powers of $z^{1/p}$ then shows that q_j is constant and pd_j is an integer for each $j < Q$, which is again impossible. □

Lemma 5.3.3 *Let $U(z)$ be a formal expression*

$$U(z) = e^{q(z)} z^d \sum_{m=0}^{n} V_m(z)(\log z)^m,$$

in which $d \in \mathbb{C}$, while q is a polynomial in $z^{1/p}$ and each $V_m(z)$ is a formal series in descending integer powers of $z^{1/p}$. If the formal derivative U' vanishes then q is constant and $pd \in \mathbb{Z}$, while $V_m(z)$ vanishes for all $m > 0$ and U reduces to a constant.

Proof. We have, with the notation $V_{n+1} = 0$,

$$\begin{aligned} U'(z) &= e^{q(z)} z^d \sum_{m=0}^{n} (V'_m(z) + (q'(z) + d/z)V_m(z))(\log z)^m + \\ &\quad + e^{q(z)} z^d \sum_{m=0}^{n} V_m(z)(m/z)(\log z)^{m-1} \\ &= e^{q(z)} z^d \sum_{m=0}^{n} (V'_m(z) + (q'(z) + d/z)V_m(z) + (m+1)V_{m+1}(z)/z)(\log z)^m. \end{aligned}$$

The fact that this expression for U' vanishes then requires that

$$V'_m(z) + (q'(z) + d/z)V_m(z) + (m+1)V_{m+1}(z)/z = 0$$

for $0 \leq m \leq n$. Taking $m = n$ gives

$$V_n'(z) + (q'(z) + d/z)V_n(z) = 0, \quad V_n'(z)/V_n(z) + q'(z) + d/z = 0.$$

This forces q to be constant and pd to be an integer. We may assume that $d = 0$, and we then have $V_n'(z) = 0$ so that V_n is a non-zero constant. Moreover, n must be 0, since otherwise

$$V_{n-1}'(z) + (n/z)V_n = 0,$$

which is impossible. □

5.4 Formal solutions and uniqueness

Lemma 5.4.1 *Let H be a $\nu \times \nu$ Jordan matrix with diagonal entries $\eta_1, \ldots, \eta_\nu \in \mathbb{C}$. Then all non-zero entries in column k of z^H have the form $c_{jk} z^{\eta_k} (\log z)^{m_{jk}}$, and all non-zero entries in row j of z^H have the form $d_{jk} z^{\eta_j} (\log z)^{n_{jk}}$, where $c_{jk}, d_{jk} \in \mathbb{C}$ and m_{jk}, n_{jk} are non-negative integers.*

Proof. If H is a single Jordan block $H = \eta I + N$, where η is the eigenvalue, I is the identity matrix and N is a shifting matrix, then $z^{\eta I} = z^\eta I$ and z^N is a matrix whose non-zero entries are constant multiples of non-negative integer powers of $\log z$. The result then follows by writing $z^H = z^{\eta I} z^N = z^N z^{\eta I}$, using the fact that I and N commute. In the general case we have $H = H_1 \oplus \ldots \oplus H_s$, where the H_j are Jordan blocks, and $z^H = z^{H_1} \oplus \ldots \oplus z^{H_s}$. □

Lemma 5.4.2 *Let $p \in \mathbb{N}$ and let H be a $\nu \times \nu$ Jordan matrix with diagonal entries $\eta_1, \ldots, \eta_\nu \in \mathbb{C}$, and let $R(z)$ be a $\nu \times \nu$ diagonal matrix with diagonal entries $r_1(z), \ldots, r_\nu(z)$, each of these being a polynomial in $z^{1/p}$. Let $V(z)$ be a $\nu \times \nu$ square matrix with entries which are formal series in descending powers of $z^{1/p}$. Then the following statements hold:*
(i) the entry in row j, column k of $Y(z) = V(z) z^H e^{R(z)}$ is $e^{r_k(z)} z^{\eta_k} T_{jk}(z)$, where $T_{jk}(z) \in \mathcal{W}$, that is, $T_{jk}(z)$ is a polynomial in $\log z$ with coefficients which are formal series in descending powers of $z^{1/p}$;
(ii) the entry in row j, column k of $e^{R(z)} z^H V(z)$ is $e^{r_j(z)} z^{\eta_j} U_{jk}(z)$, where $U_{jk}(z) \in \mathcal{W}$.

Proof. The entries of column k of $V(z) z^H$ are formed by taking the dot product of each row of $V(z)$ with column k of z^H, and Lemma 5.4.1 shows that each non-zero entry in column k of z^H has form $c z^{\eta_k} (\log z)^m$ for some $c \in \mathbb{C}$ and integer $m \geq 0$. Now right-multiplying by $e^{R(z)}$ multiplies column k by $e^{r_k(z)}$.

Similarly, the entries in row j of $z^H V(z)$ are formed by taking the dot product of row j of z^H with each column of $V(z)$, and each non-zero entry in row j of z^H has form $c z^{\eta_j} (\log z)^m$ for some $c \in \mathbb{C}$ and integer $m \geq 0$. Now left-multiplying by $e^{R(z)}$ multiplies row j by $e^{r_j(z)}$. □

Now consider the differential equation

$$y' = B(z) y, \tag{5.13}$$

where $B(z)$ is a $\nu \times \nu$ matrix whose entries are formal series in descending powers of z.

Definition 5.4.1 *A basic formal matrix will mean a $\nu \times \nu$ matrix $Y(z)$ with the following property. There exist $q \in \mathbb{N}$ and $q_1(z), \ldots, q_\nu(z)$, each a polynomial in $z^{1/q}$ with zero constant term, as well as complex numbers $\sigma_1, \ldots, \sigma_\nu$, such that the entry $Y_{jk}(z)$ in row j, column k of $Y(z)$ is $e^{q_k(z)} z^{\sigma_k} S_{jk}(z)$, where $S_{jk}(z)$ is a polynomial in $\log z$ with coefficients which are formal series in descending powers of $z^{1/q}$.*

Equivalently, $Y(z)$ has the form $Y(z) = E(z)D(z)$, where $E(z)$ is a matrix whose entries are polynomials in $\log z$ with coefficients which are formal series in descending powers of $z^{1/q}$, while $D(z)$ is a diagonal matrix with entries $e^{q_k(z)}z^{\sigma_k}$.

It is clear that if $Y(z)$ is a basic formal matrix, then so are its formal derivative $Y'(z)$ and the matrix $B(z)Y(z)$, and their columns have the same exponential parts q_k and powers σ_k as $Y(z)$. Thus we will define formal solutions of (5.13) as follows.

Definition 5.4.2 *A basic formal matrix solution of (5.13) will mean a basic formal matrix $Y(z)$ such that $Y'(z)$ and $B(z)Y(z)$ agree: that is, the powers of $\log z$ and their series coefficients in each entry of $Y'(z)$ match those of $B(z)Y(z)$.*

Definition 5.4.3 *A principal formal matrix solution of (5.13) will mean a $\nu \times \nu$ matrix solution $X(z) = U(z)z^F e^{P(z)}$ satisfying the following, for some $p \in \mathbb{N}$.*
(i) F is a constant matrix in Jordan form given by

$$F = J_1 \oplus \ldots \oplus J_s,$$

where J_j is $\mu_j \times \mu_j$ and a Jordan block.
(ii) $P(z)$ is a diagonal matrix of form

$$P(z) = P_1(z)I_{\mu_1} \oplus \ldots \oplus P_s(z)I_{\mu_s},$$

where $P_j(z)$ is a polynomial in $z^{1/p}$ with constant term 0; this implies that $P'(z)$, $P(z)$ and $e^{P(z)}$ all commute with any matrix $M = M_1 \oplus \ldots \oplus M_s$ such that M_j is $\mu_j \times \mu_j$, and in particular with F and z^F.
(iii) $U(z)$ is a matrix over $\mathcal{V}$ (that is, its entries are formal series in descending powers of $z^{1/p}$), and $\det U(z)$ is not the zero series.

Lemma 5.4.2 implies that any principal formal matrix solution is a basic formal matrix solution. Moreover, $X(z)$ in Definition 5.4.3 has determinant $\det U(z) \cdot z^{\mathrm{tr}\, F} \cdot \exp(\mathrm{tr}\, P(z))$, by Lemma 5.2.5.

Lemma 5.4.3 *If $X(z) = U(z)z^F e^{P(z)}$ is a principal formal matrix solution of (5.13) as in Definition 5.4.3, then F may be chosen so that all its eigenvalues have real part lying in $[0, 1/p)$.*

Proof. Choose a diagonal matrix $D_0 = \sigma_1 I_{\mu_1} \oplus \ldots \oplus \sigma_s I_{\mu_s}$ so that $F = D_0 + F_0$, where all eigenvalues of F_0 have real part lying in $[0, 1/p)$. Then D_0 and F_0 commute and it is possible to write

$$U(z)z^F = U(z)z^{D_0}z^{F_0} = U_0 z^{F_0},$$

in which $\det U_0(z)$ is not the zero series. □

If $X(z) = U(z)z^F e^{P(z)}$ is a principal formal matrix solution as in Definition 5.4.3 then we have, since F, z^F, $P'(z)$ and $P(z)$ all commute with each other,

$$\frac{d}{dz}\left(z^F\right) = \frac{F}{z} \cdot z^F = z^F \cdot \frac{F}{z}, \quad \frac{d}{dz}\left(e^{P(z)}\right) = P'(z)e^{P(z)}$$

and

$$(0) = X'(z) - B(z)X(z) = \left(U'(z) + U(z)\frac{F}{z} + U(z)P'(z) - B(z)U(z)\right) z^F e^{P(z)},$$

so that

$$R(z) = U'(z) + U(z)\frac{F}{z} + U(z)P'(z) - B(z)U(z)$$

must be the zero series in powers of $z^{1/p}$. The question of existence will be treated later, but some initial results concerning uniqueness will be developed following an example.

5.4.1 Uniqueness of formal solutions

Example 5.4.1 *Suppose that an equation $x' = B(z)x$ has a solution*

$$X = U(z)\begin{pmatrix} z^a & 0 \\ 0 & z^b \end{pmatrix}\begin{pmatrix} e^P & 0 \\ 0 & e^Q \end{pmatrix}$$

with $a, b \in \mathbb{C}$, P and Q polynomials in z and $U(z)$ a matrix whose entries are analytic functions of z, or formal series in descending integer powers of z. Then another solution is

$$\begin{aligned} Y &= U(z)\begin{pmatrix} z^a & 0 \\ 0 & z^b \end{pmatrix}\begin{pmatrix} e^P & 0 \\ 0 & e^Q \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= U(z)\begin{pmatrix} z^a & 0 \\ 0 & z^b \end{pmatrix}\begin{pmatrix} 0 & e^P \\ e^Q & 0 \end{pmatrix} \\ &= U(z)\begin{pmatrix} z^a & 0 \\ 0 & z^b \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} e^Q & 0 \\ 0 & e^P \end{pmatrix} \\ &= U(z)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} z^b & 0 \\ 0 & z^a \end{pmatrix}\begin{pmatrix} e^Q & 0 \\ 0 & e^P \end{pmatrix}. \end{aligned}$$

Here the powers a, b and exponential parts P and Q have been interchanged.

Lemma 5.4.4 *Let X and Y be formal solutions of (5.13), such that Y is a basic formal matrix solution as in Definition 5.4.2, and X is a principal formal matrix solution as in Definition 5.4.3. Then there exists a constant matrix C with $Y = XC$. Furthermore, if C is invertible, then the polynomials appearing in the exponential terms in the columns of Y form a permutation of the diagonal entries of P: in particular, this holds if Y is also a principal formal matrix solution $Y(z) = V(z)z^G e^{Q(z)}$ as in Definition 5.4.3.*

Proof. By taking the least common multiple, it may be assumed that the integers q and p occurring in Definitions 5.4.2 and 5.4.3 are the same. Write

$$C(z) = X(z)^{-1}Y(z) = e^{-P(z)}z^{-F}U(z)^{-1}Y(z) = (c_{jk}(z)). \tag{5.14}$$

Here $U(z)^{-1}$ exists because $\det U(z)$ is not the zero series. Let $p_1(z), \ldots, p_\nu(z)$ be the diagonal entries of $P(z)$, and $\lambda_1, \ldots, \lambda_\nu$ those of F. Then Lemma 5.4.2 and the notation of Definition 5.4.2 show that

$$c_{jk}(z) = e^{q_k(z)-p_j(z)}z^{\sigma_k-\lambda_j}v_{jk}(z), \tag{5.15}$$

where $v_{jk}(z) \in \mathcal{W}$, that is, $v_{jk}(z)$ is a polynomial in $\log z$ with coefficients which are formal series in descending powers of $z^{1/p}$. Thus $C(z)$ has a formal derivative, and

$$Y = XC, \quad (0) = Y' - BY = X'C + XC' - BXC = XC', \quad C' = (0),$$

so that $c'_{jk}(z) = 0$ for each j, k. Lemma 5.3.3 shows that $c_{jk}(z) = c_{jk}$ is a constant, and if $c_{jk} \neq 0$ then (5.15) implies that $q_k - p_j$ is constant, and so 0.

Suppose that C is invertible, and that p^* occurs s times in the list $p_1, \ldots, p_\nu$, say

$$p_{j_1} = \ldots = p_{j_s} = p^*,$$

with the j_μ pairwise distinct. Since C is invertible, Lemma 5.1.1 shows that there exist pairwise distinct k_μ with $c_{j_\mu k_\mu} \neq 0$, forcing $p_{j_\mu} - q_{k_\mu}$ to be constant. Hence p^* occurs at least s times in the list

$q_1, \ldots, q_\nu$. This implies that if the distinct polynomials which occur in the list $p_1, \ldots, p_\nu$ are $r_1, \ldots, r_\tau$, with frequencies $s_1, \ldots, s_\tau$, then these occur with frequencies $t_1 \geq s_1, \ldots, t_\tau \geq s_\tau$ in the list $q_1, \ldots, q_\nu$, and

$$\nu = \sum_{k=1}^{\tau} s_k \leq \sum_{k=1}^{\tau} t_k \leq \nu,$$

which forces $s_k = t_k$ and $\sum_{k=1}^{\tau} t_k = \nu$. Thus each list is a permutation of the other.

Now suppose that $Y(z)$ is also a principal formal matrix solution $Y(z) = V(z)z^G e^{Q(z)}$. Then the $q_k(z)$ in Definition 5.4.2 are precisely the diagonal entries of $Q(z)$, and C is invertible, because

$$\det V(z) \cdot z^{\operatorname{tr} G} \cdot \exp(\operatorname{tr} Q(z)) = \det Y(z) = \det U(z) \cdot z^{\operatorname{tr} F} \cdot \exp(\operatorname{tr} P(z)) \cdot \det C$$

and $\det V(z)$ does not vanish identically. □

In the case where $X = Uz^F e^P$ and $Y = Vz^G e^Q$ are both principal formal matrix solutions, with the same integer p, and the eigenvalues of F and G are normalised as in Lemma 5.4.3, it is possible to say more. We can write

$$F = J + D, \quad G = K + E,$$

where D and E are diagonal constant matrices, whose entries all have real part in $[0, 1/p)$, and J, K are Jordan matrices, all of whose eigenvalues are 0; moreover, this can be done so that J, D and P commute, as do K, E and Q. As before, C is a constant matrix, and if $c_{jk} \neq 0$ then Lemma 5.3.2 and (5.15) imply that $p_j = q_k$ and $p(\lambda_j - \sigma_k) \in \mathbb{Z}$, which forces $\lambda_j = \lambda_k$ by virtue of the normalisation of the eigenvalues. Hence we always have

$$c_{jk}\sigma_k = \lambda_j c_{jk} \quad \text{and} \quad c_{jk} q_k(z) = p_j(z) c_{jk},$$

whether or not $c_{jk} = 0$. It follows that

$$CE = DC, \quad CQ = PC, \quad E = C^{-1}DC, \quad Q = C^{-1}PC,$$

which leads in turn to

$$Cz^E = z^D C, \quad Ce^Q = e^P C.$$

Furthermore, Y satisfies

$$Y(z) = U(z)z^F e^{P(z)} C = U(z)z^J z^D e^{P(z)} C = U(z)z^J C z^E e^{Q(z)} = V(z) z^K z^E e^{Q(z)},$$

which forces

$$U(z)z^J C = V(z)z^K, \quad z^J C z^{-K} = H(z) = U(z)^{-1}V(z). \tag{5.16}$$

Here $H(z)$ is given by a formal series in descending powers of $z^{1/p}$, because $U(z)^{-1} \in \mathcal{V}$, which follows from the fact that $\det U(z)$ is not the zero series. But, since the eigenvalues of the Jordan matrices J and K are all 0, the entries of $z^J C z^{-K}$ are all polynomials in $\log z$. Thus H is a constant matrix and so

$$Jz^J C z^{-K} - z^J C z^{-K} K = (0), \quad JH - HK = (0). \tag{5.17}$$

Since H is invertible, (5.17) implies that $J = HKH^{-1}$, so that J and K are similar matrices, and $z^J = Hz^K H^{-1}$. This now gives, by (5.16),

$$H = Hz^K H^{-1} C z^{-K}, \quad I = z^K H^{-1} C z^{-K}, \quad z^K = z^K H^{-1} C, \quad H = C.$$

Finally, this delivers

$$CG = C(K + E) = HK + CE = JH + DC = (J + D)C = FC,$$

and so

$$G = C^{-1}FC, \quad Q = C^{-1}PC, \quad V = UC.$$

5.5 Holomorphic matrix differential equations

Lemma 5.5.1 *Let $a(z)$ be a holomorphic $\nu \times \nu$ matrix function on a domain $D \subseteq \mathbb{C}$, let $x(z)$ be a holomorphic $\nu \times \nu$ matrix solution of*

$$x' = a(z)x \tag{5.18}$$

on D, and let B be a constant $\nu \times \nu$ matrix. Then $x(z)B$ also solves (5.18) on D. Furthermore, $W(z) = \det x(z)$ satisfies $W'(z) = b(z)W(z)$ on D, where $b(z)$ is the trace of $a(z)$. In particular, if $z_0 \in D$ and $\det x(z_0) = 0$, then $\det x(z) = 0$ for all $z \in D$.

Proof. The first assertion is obvious. Next, by the product rule, we have

$$W'(z) = \sum_{j=1}^{\nu} \det x^{[j]}(z),$$

where $x^{[j]}(z)$ means the matrix $x(z)$, but with row j replaced by its derivative, which is

$$\begin{aligned} (x'_{j1}(z), \ldots, x'_{j\nu}(z)) &= \left(\sum_{t=1}^{\nu} a_{jt}(z) x_{t1}(z), \ldots, \sum_{t=1}^{\nu} a_{jt}(z) x_{t\nu}(z) \right) \\ &= \sum_{t=1}^{\nu} a_{jt}(z) \left(x_{t1}(z), \ldots, x_{t\nu}(z) \right), \end{aligned}$$

this being a linear combination of the rows of $x(z)$. Since a determinant is left unchanged by adding to one row multiples of the other rows, we get $\det x^{[j]}(z) = a_{jj}(z) \det x(z)$. □

If $\det x(z) \neq 0$ for all $z \in D$ then x will be called a non-singular solution.

Theorem 5.5.1 (The existence-uniqueness theorem) *Let $a(z)$ be a holomorphic $\nu \times \nu$ matrix function on a simply connected domain $D \subseteq \mathbb{C}$, let B be a constant $\nu \times \nu$ matrix, and let $z_0 \in \mathbb{C}$. Then the equation (5.18) has a unique holomorphic $\nu \times \nu$ matrix solution $x(z)$ on D with $x(z_0) = B$.*

Proof. This uses the (standard) Newton-Picard successive approximations method coupled with the Riemann mapping theorem. The first step is to prove existence and uniqueness on a neighbourhood of z_0. The equation can be written in integral form as

$$x(z) = x(z_0) + \int_{z_0}^{z} a(t)x(t)\, dt. \tag{5.19}$$

Define

$$x_0(z) = (0), \quad x_1(z) = B, \quad \ldots, \quad x_{q+1}(z) = B + \int_{z_0}^{z} a(t)x_q(t)\, dt \quad (q \geq 0). \tag{5.20}$$

Using the Frobenius norm for matrices, suppose that $\|a(z)\| \leq M < \infty$ on $D(z_0, \delta) \subseteq D$, and take ρ with $0 < \rho \leq \delta$ and $\rho M \leq \frac{1}{2}$. It will be shown that there exists a unique solution x of (5.18), analytic on $D(z_0, \rho)$, with $x(z_0) = B$. To this end write, for $q \geq 0$,

$$M_q = \sup\{\|x_{q+1}(z) - x_q(z)\| : z \in D(z_0, \rho)\}. \tag{5.21}$$

Then $M_0 = \|B\|$. But (5.20) gives

$$x_{q+2}(z) - x_{q+1}(z) = \int_{a}^{z} a(t)(x_{q+1}(t) - x_q(t))\, dt,$$

and $\|a(t)(x_{q+1}(t) - x_q(t))\| \leq MM_q$ on $D(z_0, \rho)$, which implies that

$$M_{q+1} \leq \rho M M_q \leq \frac{1}{2} M_q.$$

Thus $M_q \leq (1/2)^q \|B\|$ and the series

$$x(z) = \sum_{j=0}^{\infty} (x_{j+1}(z) - x_j(z)) = \lim_{q \to \infty} \sum_{j=0}^{q-1} (x_{j+1}(z) - x_j(z)) = \lim_{q \to \infty} x_q(z)$$

converges absolutely and uniformly on $D(z_0, \rho)$; moreover, the limit function $x(z)$ is analytic there, by Weierstrass' theorem. Since $x_{q+1}(z)$ and $x_q(z)$ both converge to $x(z)$, we get

$$x(z) = B + \int_a^z a(t)x(t)\, dt, \quad x' = ax, \quad x(z_0) = B.$$

The uniqueness is established as follows. If $B = (0)$ then, with ρ as above and

$$T = \sup\{\|x(z)\| : z \in D(z_0, \rho)\},$$

we get $T \leq \frac{1}{2}T$ and so $T = 0$. Moreover, with this same value of ρ, fix a solution X of (5.18) which is holomorphic on $D(z_0, \rho)$ and satisfies $X(z_0) = I$. Then the uniqueness property implies that any solution x of (5.18) which is holomorphic on $D(z_0, \rho)$ must satisfy $x(z) = X(z)x(z_0)$.

We now extend the solutions to all of the simply connected domain D. We have seen how to define solutions on $D(z_0, \rho)$, for $z_0 \in D$, where ρ depends on the coefficient A but not on x or $B = x(z_0)$. If D is a disc $D(0, R)$, where $0 < R \leq \infty$, and B is given, let S be the supremum of $r > 0$ such that there exists an analytic solution x on $D(0, r)$ with $x(0) = B$: then $S > 0$. By the identity theorem and the fact that we can choose r arbitrarily close to S, there exists such a solution on $D(0, S)$, and so if $S = R$ we have finished. If $S < R$, choose $S_1 > S$ and $M_1 > 0$ such that $\|a(z)\| \leq M_1$ for $|z| \leq S_1$. Then there exists a small positive σ such that if $|b| \leq S$ we may take $z_0 = b$ and $\rho = \sigma$ in the above construction. Choose S_2 with $S_2 < S < S_2 + \sigma$ and finitely many b_j with $|b_j| = S_2$ such that the discs $D(b_j, \sigma)$ together cover the circle $|z| = S$. For each j, we can then choose a solution y_j defined on $D(b_j, \sigma)$ with $y_j = x$ at b_j, from which it follows that $y_j = x$ on all of the domain $D(0, S) \cap D(b_j, \sigma)$. If j and m are such that $D(b_j, \sigma) \cap D(b_m, \sigma)$ is non-empty, then $D(b_j, \sigma) \cap D(b_m, \sigma)$ is connected, and $D(b_j, \sigma) \cap D(b_m, \sigma) \cap D(0, S)$ is non-empty. Hence $y_j \equiv y_m$ on $D(b_j, \sigma) \cap D(b_m, \sigma)$. But this allows us to extend x to the union of the $D(b_j, \sigma)$ and so to $D(0, S')$, where $S' > S$. This contradiction shows that $S = R$.

Thus we have proved the existence-uniqueness theorem for the whole plane and for any disc. Now if D is any simply connected domain, not the whole plane, and $z_0 \in D$, choose an analytic one-one function ϕ such that $z = \phi(w)$ maps the unit disc $D(0, 1)$ onto D, with $\phi(0) = z_0$, and let $b(w) = a(\phi(w))\phi'(w)$. Then there exists $y(w)$ on $D(0, 1)$ with $y(0) = B$ satisfying $y'(w) = b(w)y(w)$, and x may be defined by $x(z) = x(\phi(w)) = y(w)$, which gives

$$x'(z) = y'(w)\frac{dw}{dz} = \frac{y'(w)}{\phi'(w)} = \frac{b(w)y(w)}{\phi'(w)} = a(z)x(z).$$

□

5.6 The regular singular point case

This section will mainly be concerned with the equation

$$zx' = A(z)x, \tag{5.22}$$

where $A(z)$ is bounded and holomorphic in a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$. Here it is convenient to allow the possibility that $\beta - \alpha > 2\pi$, so that S is understood to lie on the Riemann surface of $\log z$, on which we no longer identify points whose arguments differ by 2π, and both $A(z)$ and $x(z)$ are continued analytically. In any case, any ambiguity may be eliminated here by considering $y(w) = x(e^w)$ and $B(w) = A(e^w)$ on the half-strip T given by $\mathrm{Re}\, w > \log R$, $\alpha < \mathrm{Im}\, w < \beta$; here $y'(w) = B(w)y(w)$ and any local solution extends to the whole of T by the existence-uniqueness theorem.

In the case where $A(z)$ is bounded and holomorphic in the annulus $R < |z| < +\infty$, the equation (5.22) will be said to have a *regular singular point* at infinity.

Lemma 5.6.1 *Let $A(z)$ be a holomorphic $\nu \times \nu$ matrix function on on an annulus Ω given by $0 < R < |z| < \infty$. Let x be a holomorphic solution of (5.22) on a domain $D \subseteq \Omega$. If $\det x(z_0) \neq 0$ for some $z_0 \in D$ then there exists a non-singular constant $\nu \times \nu$ matrix C with $\widetilde{x} = xC$ on D, where $\widetilde{x}$ denotes the solution of (5.18) obtained by analytically continuing $x(z)$ once around a circle $|z| = r > R$.*

Proof. Note that $\det(\widetilde{x}(z_0)) \neq 0$, by Lemma 5.5.1 and analytic continuation. To prove the lemma just choose C such that $\widetilde{x}(z_0) = x(z_0)C$, so that $\widetilde{x}(z) = x(z)C$ on a neighbourhood of z_0, by the existence-uniqueness theorem, and hence for all $z \in D$ by the identity theorem. □

Lemma 5.6.2 *Suppose that $x(z)$ and $A(z)$ are holomorphic $\nu \times \nu$ matrix functions on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and that $\|A(z)\| \leq M < \infty$ on S. Suppose further that x satisfies $zx' = A(z)x$ or $zx' = xA(z)$ on S and let $z_0 \in S$: then*

$$\|x(z)\| \leq \|x(z_0)\| \left|\frac{z}{z_0}\right|^M e^{(\beta-\alpha)M}$$

for $z \in S, |z| \geq |z_0|$.

Proof. This is a straightforward application of a method going back to T.H. Gronwall. As already noted, the change of variables $w = \log z$ maps S onto the horizontal half-strip T given by $\mathrm{Re}\, w > \log R$, $\alpha < \mathrm{Im}\, w < \beta$. Setting $X(w) = x(z)$ then gives

$$\|X'\| \leq M\|X\|$$

on T. Fix $w_0 \in T$, and parametrize with respect to arc length s a straight line L starting from w_0. This gives, for $w = w(s) \in T \cap L$,

$$\|X(w(s))\| \leq \|X(w_0)\| + \int_{w_0}^{w(s)} M\|X(w)\|\, |dw| \leq H(s),$$

where

$$H(s) = \|X(w_0)\| + \int_0^s M\|X(w(t))\|\, dt.$$

Then

$$H'(s) = M\|X(w(s))\| \leq MH(s), \quad H(s) \leq H(0)e^{Ms},$$

which yields, if $z \in S$ with $|z| \geq |z_0|$, and $w_0 = \log z_0$ and $w = \log z$,

$$\begin{aligned} \|X(w)\| &\leq \|X(w_0)\| e^{M|w-w_0|} \\ &= \|X(w_0)\| \exp(M|\log z/z_0|) \\ &\leq \|X(w_0)\| \exp(M \log |z/z_0| + (\beta - \alpha)M). \end{aligned}$$

□

Lemma 5.6.3 *Suppose that $x(z)$ and $A(z)$ are holomorphic $\nu \times \nu$ matrix functions on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and that $\|A(z)\| \leq M < \infty$ on S. Suppose further that $x(z)$ is non-singular for all $z \in S$ and satisfies (5.22) on S. Then $u = x^{-1}$ satisfies*

$$\|u(z)\| \leq \|u(z_0)\| \left| \frac{z}{z_0} \right|^M e^{(\beta-\alpha)M}$$

for $z, z_0 \in S$ with $|z| \geq |z_0|$.

Proof. This follows from Lemma 5.6.2 since

$$I_\nu = ux, \quad (0) = zu'x + zux' = zu'x + uAx, \quad zu' = -uA.$$

□

Lemma 5.6.4 *Suppose that $x(z)$ and $A(z)$ are holomorphic $\nu \times \nu$ matrix functions on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and that $\|A(z)\| \leq M < \infty$ on S. Suppose further that $x(z)$ satisfies (5.22) on S. If there exists $N > M$ such that $\|x(z)\| = o(|z|^{-N})$ as $z \to \infty$ in S then $x(z) \equiv 0$.*

Proof. It may be assumed that $\alpha = -\beta < 0$, and it suffices to show that $x(z)$ vanishes for large z on the positive real axis. For large positive t write

$$y(t) = -\int_t^\infty \frac{A(s)x(s)}{s}\, ds, \quad y'(t) = x'(t).$$

Since $x(t)$ and $y(t)$ both tend to 0 as $t \to \infty$, we have $x(t) = y(t)$. Because $x(t)t^N \to 0$, there must exist large positive t with $\|x(s)s^N\| \leq \|x(t)t^N\|$ for $t \leq s < +\infty$. This implies that

$$\|x(t)\| = \|y(t)\| \leq \int_t^\infty M\|x(t)\| \frac{t^N}{s^{N+1}}\, ds = \frac{M}{N}\|x(t)\|,$$

which forces $x(s) = x(t) = 0$ for $t \leq s < \infty$. □

Lemma 5.6.5 *Suppose that $x(z)$, $A(z)$ and $B(z)$ are holomorphic $\nu \times \nu$ matrix functions on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and that $\|A(z)\| \leq M < \infty$ on S. Suppose further that $x(z)$ is non-singular for all $z \in S$ and satisfies (5.22) on S, and that $C(z) = B(z) - A(z) = O(|z|^{-N})$ as $z \to \infty$ in S, where $N > 2M$. Then the equation $zy' = B(z)y$ has a solution y on S which satisfies*

$$y(z) = x(z)(I_\nu + O(|z|^{2M-N})) = (I_\nu + O(|z|^{4M-N}))x(z)$$

as $z \to \infty$ in S.

Proof. We first determine a solution $u(z) = I_\nu + O(|z|^{2M-N})$ on S of

$$u' = Du, \quad D = z^{-1}x^{-1}Cx. \tag{5.23}$$

Here D is a holomorphic $\nu \times \nu$ matrix function, and there exists $c > 0$ with $\|D(z)\| \leq c|z|^{2M-N-1}$ as $z \to \infty$ in S, by Lemmas 5.6.2 and 5.6.3. A suitable solution u will be generated in the standard way via

$$u_{-1}(z) = (0), \quad u_0(z) = I_\nu, \quad u_{n+1}(z) = I_\nu - \int_z^\infty D(t)u_n(t)\, dt, \tag{5.24}$$

in which the integration is eventually along $\arg z = (\alpha + \beta)/2$. Let T be large and positive. We assert that $\|u_n(z) - u_{n-1}(z)\| \leq 2^{-n}$ and $u_n(z)$ is bounded for $n \geq 0$, $z \in S$, $|z| \geq T$. This is evidently true for $n = 0$, and assuming it true for $0 \leq k \leq n$ implies that $u_{n+1}(z)$ is well defined by (5.24), since $N > 2M$, and that, for $z \in S$, $|z| \geq T$,

$$\|u_{n+1}(z) - u_n(z)\| = \left\| \int_z^\infty D(t)(u_n(t) - u_{n-1}(t))\, dt \right\| \leq 2^{-n} \left| \int_z^\infty c|t|^{2M-N-1}\, |dt| \right| \leq 2^{-n-1}.$$

Hence the series $\sum_{n=1}^\infty (u_n(z) - u_{n-1}(z))$ converges uniformly for $z \in S$, $|z| \geq T$, which makes it possible to write

$$u(z) = u_0(z) + \sum_{n=1}^\infty (u_n(z) - u_{n-1}(z)) = \lim_{n\to\infty} u_n(z) = I_\nu - \int_z^\infty D(t)u(t)\, dt.$$

Here u is holomorphic and bounded for $z \in S$, $|z| \geq T$, and satisfies $u' = Du$ and $\|u(z) - I_\nu\| = O(|z|^{2M-N})$ as required. Now write, using (5.23),

$$y = xu, \quad By = Bxu = (A + C)xu = zx'u + Cxu = zx'u + zxu' = zy'.$$

Then y satisfies

$$y(z) = x(z)u(z) = x(z)(I_\nu + O(|z|^{2M-N})) = x(z)(I_\nu + \delta(z)) = (I_\nu + \varepsilon(z))x(z),$$

where, in view of Lemmas 5.6.2 and 5.6.3,

$$\varepsilon(z) = x(z)\delta(z)x(z)^{-1} = O(|z|^{4M-N}).$$

□

Theorem 5.6.1 *Let $A(z)$ be a bounded holomorphic $\nu \times \nu$ matrix function on an annulus Ω given by $0 < R < |z| < \infty$, and let $D \subseteq \Omega$ be a simply connected domain. Take a non-singular solution $x(z)$ of (5.22) on D, and let $\widetilde{x}$ be the solution of (5.22) on D obtained by continuing x once counter-clockwise around the origin. Then there exists a constant matrix B such that*

$$\widetilde{x}(z) = x(z)B, \quad x(z) = W(z)z^G$$

on D, where G is any constant matrix with $\exp(2\pi i G) = B$, while $W(z)$ is a non-singular holomorphic matrix function on Ω and each entry of $W(z)$ has at most a pole at infinity. Moreover, the solution $x(z) = W(z)z^G$ continues analytically to any sector given by $|z| > R$, $-\infty < \alpha < \arg z < \beta < +\infty$.

Equations (5.22) with $A(z)$ holomorphic and bounded on an annulus $R < |z| < \infty$ will be said to have a regular singular point at infinity.

Proof. By Lemmas 5.2.4 and 5.6.1 there exist constant matrices B and C with B non-singular such that

$$\widetilde{x}(z) = x(z)B, \quad \exp(2\pi iC) = B^{-1}.$$

Since $x(z)$ may be continued analytically throughout Ω we may write

$$W(z) = x(z)z^C, \quad \widetilde{W}(z) = \widetilde{x}(z)\exp(2\pi iC)z^C = \widetilde{x}(z)B^{-1}z^C = x(z)z^C = W(z).$$

Thus W is a holomorphic non-singular matrix function on Ω, and applying Lemma 5.6.2 to $x(z)$ in $|\arg z| < \pi$ and $0 < \arg z < 2\pi$ shows that there exist positive M_1, M_2 such that

$$\|W(z)\| \le \|x(z)\| \cdot |z|^{M_1} \le |z|^{M_2} \quad \text{on } \Omega.$$

Hence each entry of $W(z)$ has at most a pole at infinity. Now set $G = -C$. □

5.7 Asymptotic series

Let $p \in \mathbb{N}$ and consider a formal series $a(z)$ in descending powers of $z^{1/p}$ given by

$$a(z) = \sum_{m \in \mathbb{Z}} a_m z^{m/p},$$

with $a_m \in \mathbb{C}$ and $a_m = 0$ for all sufficiently large $m > 0$. If a branch of $z^{1/p}$ is chosen on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and if $b(z)$ is holomorphic on S, then $a(z)$ is called an asymptotic series for $b(z)$ on S if the following is true: for each $n \in \mathbb{N}$ we have

$$b(z) - \sum_{m \in \mathbb{Z}, m \ge -n} a_m z^{m/p} = o(|z|^{-n/p})$$

as $z \to \infty$ in S. This will be written $b(z) \sim a(z)$ on S, and an equivalent condition is, for each $n \in \mathbb{N}$,

$$b(z) - \sum_{m \in \mathbb{Z}, m \ge -n} a_m z^{m/p} = O(|z|^{-(n+1)/p}).$$

As before, it is convenient to allow the possibility that $\beta - \alpha > 2\pi$, which is facilitated by mapping to a half-strip via $w = \log z$.

Lemma 5.7.1 *Suppose that $b(z)$ and $d(z)$ are holomorphic on the sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, each having an asymptotic series in descending powers of $z^{1/p}$. Then so have $b(z) + d(z)$ and $b(z)d(z)$. If the asymptotic series for $b(z)$ is not the zero series, then $1/b(z)$ also has an asymptotic series on S. Finally, if $\varepsilon > 0$ then $b'(z)$ has an asymptotic series on the sector $\alpha + \varepsilon < \arg z < \beta - \varepsilon$, obtained by differentiating that of b term by term.*

Proof. To obtain an asymptotic series for $1/b$ assume without loss of generality that $p = 1$ and $b(z) = 1 - f(z)$, where $f(z) \sim \sum_{m<0} a_m z^m$. This gives, for $N \in \mathbb{N}$,

$$\begin{aligned}
\frac{1}{b(z)} &= \sum_{n=0}^{N} f(z)^n + O(|z|^{-1-N}) \\
&= \sum_{n=0}^{N} \left(\sum_{-N \le m < 0} a_m z^m + O(|z|^{-1-N}) \right)^n + O(|z|^{-1-N}) \\
&= \sum_{n=0}^{N} \left(\sum_{-N \le m < 0} a_m z^m \right)^n + O(|z|^{-1-N}) \\
&= \sum_{n=0}^{N} d_n z^{-n} + O(|z|^{-1-N}),
\end{aligned}$$

in which $d_0, \ldots, d_N$ are the coefficients in the formal reciprocal of $1 - \sum_{m<0} a_m z^m$ and are independent of those a_m with $m > N$. The proof of the other assertions is routine. □

Lemma 5.7.2 *Suppose that $a(z)$ is holomorphic on the sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and has an asymptotic series $a(z) \sim b(z) = \sum_{n=1}^{\infty} b_n z^{-n}$ there. Then $c(z) = \exp(a(z))$ has asymptotic series $c(z) \sim d(z) = \sum_{n=0}^{\infty} d_n z^{-n}$, where $d_0 = 1$ and $d(z)$ is the formal exponential of $b(z)$. Furthermore, $c'(z)$ has asymptotic series $b'(z)d(z)$.*

Proof. Let $N \in \mathbb{N}$. As $z \to \infty$ in S, we have

$$\begin{aligned}
c(z) &= \sum_{n=0}^{N} \frac{1}{n!} \left(\sum_{m=1}^{N} b_m z^{-m} + O(|z|^{-1-N}) \right)^n + O(|z|^{-1-N}) \\
&= \sum_{n=0}^{N} \frac{1}{n!} \left(\sum_{m=1}^{N} b_m z^{-m} \right)^n + O(|z|^{-1-N}) \\
&= \sum_{n=0}^{N} d_n z^{-n} + O(|z|^{-1-N}),
\end{aligned}$$

in which $d_0, \ldots, d_N$ are the coefficients in the formal exponential of $b(z)$ and are independent of those b_m with $m > N$.

□

Theorem 5.7.1 *Given a formal series $a(z) = \sum_{m \in \mathbb{Z}} a_m z^{m/p}$ in descending powers of $z^{1/p}$, and any choice of the branch of $z^{1/p}$ on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, there exists a holomorphic function $f(z)$ on S with $f(z) \sim a(z)$ on S.*

Proof. It may be assumed that that $p = 1$, since if $p > 1$ then $w = z^{1/p}$ maps S onto a sector. It may also be assumed that $\alpha = -\beta < 0$ and $R \ge 2$, and that $a_m = 0$ for all $m \ge 0$, as this involves only subtracting a polynomial from $a(z)$.

Since the function $(1 - e^z)/z$ is entire, and bounded in the left halfplane, there exists $C > 0$ such that if $\operatorname{Re} z < 0$ then $|1 - e^z| \le C|z|$. Choose a small positive d, in particular with $d\beta < \pi/4$, and for $m < 0$ set

$$b_m(z) = 1 - \exp\left(-c_m(z)\right), \quad c_m(z) = \frac{z^d}{1 + |a_m|},$$

so that $|\arg z| < \beta$ gives $|\arg c_m(z)| < \pi/4$ and $|a_m b_m(z)| \leq C|z|^d$. Therefore

$$\sum_{m<0} |a_m b_m(z) z^m| \leq C \sum_{m<0} |z|^{d+m} \leq C \sum_{m<0} R^{d+m} < \infty$$

on S, and so the series

$$f(z) = \sum_{m<0} a_m b_m(z) z^m$$

converges absolutely and uniformly, and is holomorphic, there. Let $n \in \mathbb{N}$ and write

$$f(z) - \sum_{-n \leq m < 0} a_m z^m = \sum_{-n \leq m < 0} a_m b_m(z) z^m - \sum_{-n \leq m < 0} a_m z^m + \sum_{m < -n} a_m b_m(z) z^m,$$

in which, as $z \to \infty$ in S,

$$\sum_{m<-n} |a_m b_m(z) z^m| \leq \sum_{m<-n} C|z|^{d+m} = C|z|^{d-n-1} \sum_{m \leq 0} |z|^m \leq C|z|^{d-n-1} \sum_{m \leq 0} R^m = o(|z|^{-n}),$$

while

$$\sum_{-n \leq m < 0} a_m z^m - \sum_{-n \leq m < 0} a_m b_m(z) z^m = \sum_{-n \leq m < 0} a_m z^m \exp\left(-c_m(z)\right)$$

tends to 0 faster than any power of $|z|$.

□

5.7.1 Asymptotic series and the inverse matrix

In general, a non-singular holomorphic function $A(z)$ can have an asymptotic series in descending powers of z without its algebraic inverse necessarily having one: for example $e^{-z} \sim 0$ on the sector $|\arg z| < \pi/4$, but e^z has there no asymptotic series in descending powers of z.

However, suppose that we have a formal $\nu \times \nu$ matrix series $\widetilde{A}(z) = \sum_{n \in \mathbb{Z}} A_n z^n$, such that $A_n = 0$ for all sufficiently large $n > 0$ and $\widetilde{d}(z) = \det \widetilde{A}(z)$ is not the zero series. Then a formal inverse $\widetilde{B}(z) = \sum_{n \in \mathbb{Z}} B_n z^n$ is given by the standard formula for the inverse matrix as the adjugate matrix divided by the determinant $\widetilde{d}(z)$.

Suppose next that $A(z)$ is a holomorphic matrix function on a sector S, with

$$A(z) \sim \widetilde{A}(z) = \sum_{n \in \mathbb{Z}} A_n z^n$$

as $z \to \infty$ in S, the series again having $A_n = 0$ for all sufficiently large $n > 0$, and suppose that $\widetilde{d}(z) = \det \widetilde{A}(z)$ is not the zero series. Then $\widetilde{A}(z)$ has a formal inverse $\widetilde{B}(z) = \sum_{n \in \mathbb{Z}} B_n z^n$. Moreover, $d(z) = \det A(z) \sim \widetilde{d}(z)$, and so $A(z)$ is a non-singular matrix for each large $z \in S$. Hence an inverse matrix function $B(z)$ of $A(z)$ is defined by the adjugate-determinant quotient formula, and taking asymptotic series in this formula shows that $B(z) \sim \widetilde{B}(z)$.

5.7.2 Asymptotic series and the equation (5.22)

Lemma 5.7.3 *Given a formal series $\sum_{m=0}^{\infty} A_m z^{-m}$, where each A_m is a constant $\nu \times \nu$ matrix, there exist $M > 0$, a non-negative integer Q, an increasing real sequence (R_n) and a constant matrix G with the following properties. First,*

$$D_n(z) = \sum_{m=0}^{n} A_m z^{-m} \quad \textit{satisfies} \quad \|D_n(z)\| \leq M \quad \textit{for } |z| \geq R_n. \tag{5.25}$$

Next, let $-\infty < \alpha < \beta < +\infty$. If N is sufficiently large then for all $n \geq N$ the equation

$$zx' = D_n(z)x \tag{5.26}$$

has a holomorphic solution $x_n(z) = W_n(z)z^G$ on the sector S^* on the Riemann surface of $\log z$ given by $|z| > R_N$, $\alpha < \arg z < \beta$, such that W_n is a non-singular holomorphic matrix function on $R_N < |z| < \infty$, each entry of W_n having at most a pole of order Q at infinity. Moreover, there exist $P > 0$ and a formal series $\sum_{m=0}^{\infty} C_m z^{-m}$, independent of n, such that the W_n satisfy

$$\left\| W_n(z) - z^Q \sum_{m=0}^{n} C_m z^{-m} \right\| \leq |z|^{P-n} \quad \text{as } z \to \infty. \tag{5.27}$$

Proof. Let $R_0 = 1$; once R_{n-1} has been chosen, choose $R_n > R_{n-1}$ such that $\|A_n z^{-n}\| \leq 2^{-n}$ for $|z| \geq R_n$. Thus (5.25) holds with $M = \|A_0\| + 1$.

Now let N be a large positive integer and assume without loss of generality that $\beta - \alpha > 4\pi$. It will be shown that there exist, for each $n \geq N$, a constant matrix G_n and a non-singular solution $x_n(z) = W_n(z)z^{G_n}$ of (5.26) on S^*, where W_n is a non-singular holomorphic matrix function on $R_N < |z| < \infty$ and each entry of W_n has at most a pole at infinity. Moreover, provided N is large enough, this will be accomplished so that each matrix G_n satisfies $G_n = G_N = G$.

For $n = N$ the existence of such a solution $W_N(z)z^{G_N}$, with G_N a constant matrix and W_N holomorphic on $R_N < |z| < +\infty$, follows from Theorem 5.6.1. The solutions x_n for $n > N$ are now determined inductively as follows. If $n \geq N$ and $x_n(z) = W_n(z)z^{G_n}$ has been determined, combining (5.25) with Lemmas 5.6.2 and 5.6.5 shows that there exists a solution x_{n+1} of

$$zx' = D_{n+1}(z)x, \tag{5.28}$$

holomorphic on the sector $|z| > R_N$, $\alpha < \arg z < \beta$, such that

$$x_{n+1}(z) = x_n(z)(I_\nu + O(|z|^{2M-n})) = x_n(z) + O(|z|^{3M-n}) = O(|z|^M) \tag{5.29}$$

as $z \to \infty$ there. Starting near the ray $\arg z = \alpha + \pi/4$ and continuing (5.28) once counter-clockwise around the origin then gives a continued solution

$$\widetilde{x}_{n+1}(z) = \widetilde{x}_n(z)(I_\nu + O(|z|^{2M-n})) = x_n(z)(B_n + O(|z|^{2M-n})), \quad B_n = \exp(2\pi i G_n).$$

Hence (5.29) yields, as $z \to \infty$ near $\arg z = \alpha + \pi/4$,

$$\widetilde{x}_{n+1}(z) = x_{n+1}(z)(I_\nu + O(|z|^{2M-n}))(B_n + O(|z|^{2M-n})) = x_{n+1}(z)B_n + \phi_n(z),$$

in which $\phi_n(z) = O(|z|^{3M-n})$ satisfies (5.28) and so vanishes identically by Lemma 5.6.4, since N is large. Applying Theorem 5.6.1 then makes it possible to write $x_{n+1}(z) = W_{n+1}(z)z^{G_n}$, where W_{n+1} is a non-singular holomorphic matrix function on $R_N < |z| < \infty$, and each entry of W_n has at most a pole at infinity, this holding initially near $\arg z = \alpha + \pi/4$, but extending to $\alpha < \arg z < \beta$ by continuation of z^{G_n}. This completes the induction, and shows that $G_n = G_N = G$ for all $n \geq N$.

Now (5.25) and Lemma 5.6.2 yield $Q \in \mathbb{N}$ such that

$$W_n(z) = z^Q \sum_{m=0}^{\infty} C_{m,n} z^{-m}$$

as $z \to \infty$, in which each $C_{m,n}$ is a constant matrix (here Q depends only on M and G_N). Moreover, (5.29) delivers $P_1 > 0$, independent of n, such that

$$\sum_{m=0}^{\infty} (C_{m,n+1} - C_{m,n})z^{-m} = z^{-Q}(W_{n+1}(z) - W_n(z)) = z^{-Q}(x_{n+1}(z) - x_n(z))z^{-G} = O(|z|^{P_1-n})$$

as $z \to \infty$. This implies that $C_{m,n+1} = C_{m,n} = C_m$ for $m < n - P_1$, which proves (5.27). $\square$

Theorem 5.7.2 *For each integer $m \geq 0$ let A_m be a constant matrix. Then the formal differential equation*

$$zx' = \left(\sum_{m=0}^{\infty} A_m z^{-m} \right) x \tag{5.30}$$

has a formal solution $S(z) = T(z)z^G = \sum_{m=0}^{\infty} C_m z^{-m} z^G$, where G and C_m, $m \geq 0$, are constant matrices, and the determinant of $T(z) = \sum_{m=0}^{\infty} C_m z^{-m}$ is not the zero series.

Moreover, if $A(z)$ is a holomorphic $\nu \times \nu$ matrix function on a sector $S(R, \alpha, \beta)$ given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, such that $A(z)$ has on $S(R, \alpha, \beta)$ the asymptotic series

$$A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m},$$

then (5.22) has a holomorphic solution $x(z) = Y(z)z^G$ on $S(R, \alpha, \beta)$, where $Y(z)$ has the asymptotic series $Y(z) \sim \sum_{m=0}^{\infty} C_m z^{-m}$ there.

This theorem is the key result of this section. It may be applied, in particular, when $A(z)$ is holomorphic and bounded on an annulus $R < |z| < +\infty$, in which case its asymptotic series is a convergent Laurent series, and the theorem gives the existence of holomorphic solutions, on any sector $S(R, \alpha, \beta)$, of the equation (5.22), which has a regular singular point at infinity.

Proof. Let D_n, G, x_n, W_n and the sector S^* be as in Lemma 5.7.3. By incorporating a term $z^{\lambda I_\nu}$ into z^G, where $\lambda \in \mathbb{Z}$, it may be assumed further that $C_0 \neq (0)$ and $Q = 0$ in (5.27), so that $W_n(\infty) = C_0$ is a finite matrix. The fact that $x_n(z) = W_n(z)z^G$ solves (5.26) gives $P \in \mathbb{N}$ such that, for all large n,

$$\begin{aligned}
(0) &= zW_n'(z) + W_n(z)G - D_n(z)W_n(z) \\
&= z \sum_{m=0}^{n} mC_m z^{-m-1} + \sum_{m=0}^{n} C_m G z^{-m} - \left(\sum_{m=0}^{n} A_m z^{-m} \right) \left(\sum_{m=0}^{n} C_m z^{-m} \right) + O(z^{P-n}) \\
&= z \sum_{m=0}^{\infty} mC_m z^{-m-1} + \sum_{m=0}^{\infty} C_m G z^{-m} - \left(\sum_{m=0}^{\infty} A_m z^{-m} \right) \left(\sum_{m=0}^{\infty} C_m z^{-m} \right) + O(z^{P-n}),
\end{aligned}$$

where $O(z^{P-n})$ means a formal series involving no powers of z higher than $P - n$. Since n is arbitrary, this gives the formal solution $S(z) = T(z)z^G$ of (5.30).

To establish the non-vanishing of of $\det T(z)$, observe first that, by Lemmas 5.2.5 and 5.5.1, (5.26) and the fact that $W_n(z)$ is non-singular, there exists $c_n \neq 0$ such that, as $z \to \infty$ in S^*,

$$\begin{aligned}
\det W_n(z) &= z^{-\operatorname{tr} G} \det x_n(z) = z^{-\operatorname{tr} G} \exp \left(\int^z u^{-1} (\operatorname{tr} D_n(u)) \, du \right) \\
&= c_n z^{\operatorname{tr}(A_0 - G)} \exp \left(\int^z \sum_{m=1}^{n} \operatorname{tr} A_m u^{-m-1} \right) \sim c_n z^{\operatorname{tr}(A_0 - G)}.
\end{aligned}$$

Provided n is so large that $n - P > |\operatorname{tr}(A_0 - G)|$, formula (5.27) now yields, for large n, as $z \to \infty$.

$$\det \left(\sum_{m=0}^{n} C_m z^{-m} \right) = \det W_n(z) + O(|z|^{P-n}) \sim c_n z^{\operatorname{tr}(A_0 - G)}.$$

The left-hand side of this equation is a rational function and the leading term of its Laurent series, valid near infinity, is independent of n for large n, from which it follows that so is c_n. This implies that for large n, in the sense of formal series,

$$\begin{aligned}\det T(z) &= \det\left(\sum_{m=0}^{n} C_m z^{-m}\right) + O(z^{-n-1}) \\ &= c_n z^{\mathrm{tr}(A_0-G)} + O\left(z^{-1+\mathrm{tr}(A_0-G)}\right) + O(z^{-n-1}) \neq 0.\end{aligned}$$

Now suppose that $A(z)$ and the sector $S(R,\alpha,\beta)$ are as in the hypotheses. By Lemma 5.6.5 there exist $M_1 > 1$ and, for each large n, a solution

$$y_n(z) = (I_\nu + O(|z|^{M_1-n}))x_n(z) = (I_\nu + O(|z|^{M_1-n}))W_n(z)z^G = Y_n(z)z^G \tag{5.31}$$

of (5.22) on $S(R,\alpha,\beta)$. Then (5.27) shows that there exists $M_2 > 1$ with, for each large n,

$$\begin{aligned}Y_{n+1}(z) - Y_n(z) &= (I_\nu + O(|z|^{M_1-n-1}))W_{n+1}(z) - (I_\nu + O(|z|^{M_1-n}))W_n(z) \\ &= (I_\nu + O(|z|^{M_1-n-1}))(W_n(z) + O(|z|^{P-n})) - (I_\nu + O(|z|^{M_1-n}))W_n(z) \\ &= O(|z|^{M_2-n})\end{aligned}$$

as $z \to \infty$ on $S(R,\alpha,\beta)$. It follows from Lemma 5.6.4 that $y_{n+1} = y_n = y$ and $Y_{n+1} = Y_n = Y$ on $S(R,\alpha,\beta)$, for all large n. Now (5.27), (5.31) and the formula $W_n(\infty) = C_0$ together show that there exists $M_3, M_4 > 1$ with, for each large n,

$$Y(z) = W_n(z) + O(|z|^{M_3-n}) = \sum_{m=0}^{n} C_m z^{-m} + O(|z|^{M_4-n}).$$

It follows that $\sum_{m=0}^{\infty} C_m z^{-m}$ is an asymptotic series for Y on $S(R,\alpha,\beta)$. $\square$

In Theorem 5.7.2 it may be assumed further that G is in Jordan form, so that $S(z)$ becomes a principal formal matrix solution as in Definition 5.4.3. This may be seen by choosing an invertible constant matrix H such that $J = H^{-1}GH$ is in Jordan form, and right-multiplying $S(z)$ and $x(z)$ by H, using the fact that

$$S(z)H = T(z)z^G H = T(z)HH^{-1}z^G H = T(z)Hz^J, \quad x(z)H = Y(z)Hz^J.$$

Example

In Theorem 5.7.2 it cannot in general be asserted that $\det C_0 \neq 0$. Write

$$H = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \quad x(z) = \begin{pmatrix}1 & -1\\ 0 & z\end{pmatrix} z^H = \begin{pmatrix}1 & -1\\ 0 & z\end{pmatrix}\begin{pmatrix}z & z\log z\\ 0 & z\end{pmatrix} = \begin{pmatrix}z & z\log z - z\\ 0 & z^2\end{pmatrix},$$

so that

$$zx'(z) = \begin{pmatrix}z & z\log z\\ 0 & 2z^2\end{pmatrix} = \begin{pmatrix}1 & 1/z\\ 0 & 2\end{pmatrix}\begin{pmatrix}z & z\log z - z\\ 0 & z^2\end{pmatrix} = \begin{pmatrix}1 & 1/z\\ 0 & 2\end{pmatrix} x(z).$$

Here $x(z) = T(z)z^H$ with

$$T(z) = \begin{pmatrix}1 & -1\\ 0 & z\end{pmatrix} = z\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} + \begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}, \quad \det T(z) = z, \quad T(z)^{-1} = \begin{pmatrix}1 & 1/z\\ 0 & 1/z\end{pmatrix}.$$

Suppose that $x(z) = U(z)z^F$ with $U(z) = U_0 + U_1 z^{-1} + \ldots$ and $\det U_0 \neq 0$, where F and the U_m are all constant matrices. Then there exist constant matrices M and G, with M non-singular and G in Jordan form, such that

$$FM = MG, \quad x(z)M = U(z)z^F M = U(z)Mz^G = V(z)z^G,$$

where $V(z) = U(z)M = V_0 + V_1 z^{-1} + \ldots$ and $\det V_0 \neq 0$. This gives

$$z^3 \det M = \det(x(z)M) = z^g \det V(z) = z^g(\det V_0 + o(1)),$$

where g is the trace of G, which must therefore be 3. Hence the sum of the eigenvalues of G must be 3. Now $3/2$ cannot be the unique eigenvalue of G, since otherwise $x(z)M = V(z)z^G$ would involve fractional powers of z, and so the eigenvalues of G are distinct. But then $V(z)z^G$ cannot involve logarithms, and nor can $x(z)M$, so $M_{21} = M_{22} = 0$, contradicting the fact that M is non-singular.

The same $x(z)$ can be written, in accordance with Lemma 5.4.3, in the form

$$\begin{pmatrix} z & z\log z - z \\ 0 & z^2 \end{pmatrix} = \begin{pmatrix} z & -z \\ 0 & z^2 \end{pmatrix} \begin{pmatrix} 1 & \log z \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} z & -z \\ 0 & z^2 \end{pmatrix} z^K, \quad K = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

in which K has 0 as its only eigenvalue, and so is not similar to H.

5.8 Scalar equations and asymptotic series

Theorem 5.8.1 *Given an integer p and a formal series $A(z) = \sum_{m=-\infty}^{p} A_m z^m$, with each $A_m \in \mathbb{C}$, there exist a polynomial P and a complex number Q such that the equation*

$$x' = A(z)x = \left(\sum_{m=-\infty}^{p} A_m z^m \right) x \tag{5.32}$$

has a formal solution $X(z) = z^Q e^{P(z)} U(z)$, where $U(z) = \sum_{m=0}^{\infty} u_m z^{-m}$ with $u_m \in \mathbb{C}$ and $u_0 = 1$.

Moreover, if $B(z)$ is a holomorphic function on a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$, and $B(z)$ has on S the asymptotic series $B(z) \sim A(z) = \sum_{m=-\infty}^{p} A_m z^m$, then the equation

$$x' = B(z)x \tag{5.33}$$

has a holomorphic solution $x(z) = z^Q e^{P(z)} Y(z)$ on S, where $Y(z)$ has asymptotic series $Y(z) \sim U(z)$ on S.

Proof. On S write

$$P(z) = \sum_{m=0}^{p} \frac{A_m z^{m+1}}{m+1}, \quad Q = A_{-1}, \quad C(z) = B(z) - \frac{Q}{z} - P'(z) \sim \sum_{m \leq -2} A_m z^m,$$

and

$$Y(z) = \exp(D(z)), \quad D(z) = -\int_z^\infty C(t)\, dt \sim E(z) = \sum_{m \leq -2} \frac{A_m z^{m+1}}{m+1}.$$

Lemma 5.7.2 shows that $Y(z) = \exp(D(z))$ has an asymptotic series $Y(z) \sim U(z) = \sum_{m=0}^{\infty} u_m z^{-m}$ on S, where $u_m \in \mathbb{C}$, $u_0 = 1$, and $U(z)$ is the formal exponential of $E(z)$. Thus $Y(z)$ satisfies

$$\begin{aligned}
0 &= Y'(z) - D'(z)Y(z) = Y'(z) - C(z)Y(z) \\
&= Y'(z) - \left(B(z) - \frac{Q}{z} - P'(z)\right) Y(z) \\
&= Y'(z) + \frac{Q}{z} \cdot Y(z) + P'(z)Y(z) - B(z)Y(z) \\
&\sim U'(z) + \frac{Q}{z} \cdot U(z) + P'(z)U(z) - A(z)U(z).
\end{aligned}$$

Thus $x(z) = z^Q e^{P(z)} Y(z)$ solves (5.33) and $X(z) = z^Q e^{P(z)} U(z)$ is a formal solution of (5.32). □

The aim of the subsequent sections will be to prove a counterpart of Theorem 5.8.1 for the case of matrix linear differential equations (5.18). For the special case of (5.22), with $A(z)$ a bounded holomorphic matrix function in a sector, such a result is already provided by Theorem 5.7.2.

5.9 Reducing the dimension via eigenvalues

We start this section with the $\nu \times \nu$ equation

$$z^{1-\rho} x' = A(z)x, \quad A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m}, \tag{5.34}$$

and the associated formal equation

$$z^{1-\rho} x' = \widetilde{A}(z)x, \quad \widetilde{A}(z) = \sum_{m=0}^{\infty} A_m z^{-m}. \tag{5.35}$$

Here $\rho \in \mathbb{Z}$ and $A(z)$ is a $\nu \times \nu$ holomorphic matrix function, the asymptotic series in (5.34) being valid as $z \to \infty$ in a sector S given by $|z| > R > 0$, $-\infty < \alpha < \arg z < \beta < +\infty$. Then the cases where $\rho \leq 0$ or all A_m are the zero matrix are covered by Theorem 5.7.2. Assume for the rest of this section that $\rho \geq 1$ and $A_0 \neq (0)$: the equation is then said to have rank ρ.

Following Wasow [72, pp.52-55], assume for now that $\nu \geq 2$ and that $A_0 = \lim_{z\to\infty, z\in S} A(z)$ has the block form

$$A_0 = \begin{pmatrix} A_0^{11} & 0 \\ 0 & A_0^{22} \end{pmatrix}, \tag{5.36}$$

in which A_0^{11} and A_0^{22} are square matrices of dimensions μ and $\nu - \mu$ respectively, with no common eigenvalue. We seek a formal transformation

$$x = \widetilde{P}(z)y, \quad \widetilde{P}(z) = \sum_{m=0}^{\infty} P_m z^{-m}, \tag{5.37}$$

which turns the formal equation (5.35) into

$$z^{1-\rho} y' = \widetilde{B}(z)y, \quad \widetilde{B}(z) = \widetilde{P}(z)^{-1}\widetilde{A}(z)\widetilde{P}(z) - z^{1-\rho}\widetilde{P}(z)^{-1}\widetilde{P}'(z) = \sum_{m=0}^{\infty} B_m z^{-m}, \tag{5.38}$$

with each B_m a block diagonal matrix having the same block configuration as A_0.

The second equation of (5.38) can be written

$$z^{1-\rho}\widetilde{P}'(z) = \widetilde{A}(z)\widetilde{P}(z) - \widetilde{P}(z)\widetilde{B}(z). \tag{5.39}$$

Writing

$$\widetilde{A}(z) = \sum_{m\in\mathbb{Z}} A_m z^{-m}, \quad A_m = 0 \text{ for } m < 0, \tag{5.40}$$

with a similar convention for $\widetilde{B}$, P and P', gives the recurrence relation

$$-(m-\rho)P_{m-\rho} = \sum_{s\in\mathbb{Z}}(A_{m-s}P_s - P_sB_{m-s}) = \sum_{0\le s\le m}(A_{m-s}P_s - P_sB_{m-s}), \tag{5.41}$$

in which the sums on the right reduce because of (5.40). Since $\rho \ge 1$, the equation (5.41) is vacuous for $m < 0$, and for $m = 0$ it gives

$$A_0P_0 - P_0B_0 = \rho P_{-\rho} = (0). \tag{5.42}$$

For $m > 0$ we write (5.41) as

$$A_0P_m - P_mB_0 = \sum_{s=0}^{m-1}(P_sB_{m-s} - A_{m-s}P_s) - (m-\rho)P_{m-\rho}. \tag{5.43}$$

The choice $P_0 = I = I_\nu$ then gives, for $m > 0$, by (5.42), (5.43) and the fact that $\rho \ge 1$,

$$P_0 = I, \quad B_0 = A_0, \quad A_0P_m - P_mA_0 = B_m + H_m, \tag{5.44}$$

where H_m depends only on $A_0, \ldots, A_m$ (which are known) and those P_j and B_j with $0 \le j < m$.

We assert that these equations can be solved in such a way that, for each $m \ge 1$,

$$B_m = \begin{pmatrix} B_m^{11} & 0 \\ 0 & B_m^{22} \end{pmatrix}, \quad P_m = \begin{pmatrix} 0 & P_m^{12} \\ P_m^{21} & 0 \end{pmatrix}, \tag{5.45}$$

where B_m and P_m are block matrices in the same configuration as (5.36) (the first of these is clearly true for $m = 0$, but the second is not). For $m > 0$ write H_m in the block configuration of (5.36) as

$$H_m = \begin{pmatrix} H_m^{11} & H_m^{12} \\ H_m^{21} & H_m^{22} \end{pmatrix}.$$

Then we require, for $m > 0$,

$$\begin{aligned} B_m^{11} + H_m^{11} &= (0), \\ A_0^{11}P_m^{12} - P_m^{12}A_0^{22} &= H_m^{12}, \\ A_0^{22}P_m^{21} - P_m^{21}A_0^{11} &= H_m^{21}, \\ B_m^{22} + H_m^{22} &= (0). \end{aligned} \tag{5.46}$$

The first and last equations of (5.46) are automatically satisfied by setting $B_m^{11} = -H_m^{11}$ and $B_m^{22} = -H_m^{22}$. Because A_0^{11} and A_0^{22} have no common eigenvalue, Lemma 5.2.9 shows that the second and third equations are also solvable (and uniquely). This proves the following theorem.

Theorem 5.9.1 *Suppose that A_0 has the block form (5.36), where A_0^{11} and A_0^{22} are square matrices of dimensions μ and $\nu-\mu$ respectively, and with no common eigenvalue. Then there exists a formal transformation $x = \widetilde{P}(z)y = \sum_{m=0}^{\infty} P_m z^{-m} y$ with*

$$P_0 = I_\nu, \quad P_m = \begin{pmatrix} 0 & P_m^{12} \\ P_m^{21} & 0 \end{pmatrix} \quad (m \geq 1),$$

which transforms (5.35) to

$$z^{1-\rho} y' = \sum_{m=0}^{\infty} B_m z^{-m} y, \quad B_m = \begin{pmatrix} B_m^{11} & 0 \\ 0 & B_m^{22} \end{pmatrix}, \tag{5.47}$$

where $B_0 = A_0$ and B_m^{11} is $\mu \times \mu$, while B_m^{22} is $(\nu-\mu)\times(\nu-\mu)$.

□

The next issue is to resolve whether the same reduction is possible for *holomorphic* solutions of (5.34). To this end assume again that $A(z)$ has the asymptotic series in (5.34), in which A_0 has the block form (5.36), and write

$$P(z) = I_\nu + \widehat{P}(z), \quad B(z) = A_0 + \widehat{B}(z), \tag{5.48}$$

as well as

$$\begin{aligned} B &= B_0 + \widehat{B} = \begin{pmatrix} A_0^{11} + \widehat{B}^{11} & 0 \\ 0 & A_0^{22} + \widehat{B}^{22} \end{pmatrix}, \quad B_0 = A_0, \\ P &= I_\nu + \widehat{P} = \begin{pmatrix} I^{11} & \widehat{P}^{12} \\ \widehat{P}^{21} & I^{22} \end{pmatrix}, \\ A &= \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix}, \quad A_0^{12} = (0), \quad A_0^{21} = (0), \end{aligned} \tag{5.49}$$

with I^{11} and I^{22} identity matrices of appropriate dimension, and all of these matrices in the same block configuration as A_0. Then, by (5.38) and (5.39), the transformation $x = P(z)y$ turns (5.34) into $z^{1-\rho}y' = B(z)y$ if and only if P and B satisfy

$$\begin{aligned} z^{1-\rho} \begin{pmatrix} 0 & \left(\widehat{P}^{12}\right)' \\ \left(\widehat{P}^{21}\right)' & 0 \end{pmatrix} &= AP - PB \\ &= \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} \begin{pmatrix} I^{11} & \widehat{P}^{12} \\ \widehat{P}^{21} & I^{22} \end{pmatrix} \\ &\quad - \begin{pmatrix} I^{11} & \widehat{P}^{12} \\ \widehat{P}^{21} & I^{22} \end{pmatrix} \begin{pmatrix} A_0^{11} + \widehat{B}^{11} & 0 \\ 0 & A_0^{22} + \widehat{B}^{22} \end{pmatrix}. \end{aligned}$$

Expanding this out gives

$$\begin{aligned} (0) &= A^{11} + A^{12}\widehat{P}^{21} - (A_0^{11} + \widehat{B}^{11}), \\ z^{1-\rho}\left(\widehat{P}^{12}\right)' &= A^{11}\widehat{P}^{12} + A^{12} - \widehat{P}^{12}(A_0^{22} + \widehat{B}^{22}), \\ z^{1-\rho}\left(\widehat{P}^{21}\right)' &= A^{21} + A^{22}\widehat{P}^{21} - \widehat{P}^{21}(A_0^{11} + \widehat{B}^{11}), \\ (0) &= A^{21}\widehat{P}^{12} + A^{22} - (A_0^{22} + \widehat{B}^{22}). \end{aligned} \tag{5.50}$$

We eliminate $\widehat{B}^{11}$ and $\widehat{B}^{22}$ using the first and last equations of (5.50). The second and third equations then become

$$\begin{aligned} z^{1-\rho}\left(\widehat{P}^{12}\right)' &= A^{11}\widehat{P}^{12} + A^{12} - \widehat{P}^{12}(A^{21}\widehat{P}^{12} + A^{22}), \\ z^{1-\rho}\left(\widehat{P}^{21}\right)' &= A^{21} + A^{22}\widehat{P}^{21} - \widehat{P}^{21}(A^{11} + A^{12}\widehat{P}^{21}). \end{aligned} \tag{5.51}$$

Now the equations (5.48), (5.49) and (5.50) are satisfied when A, P and B are replaced by the formal series occurring in Theorem 5.9.1. Thus the equations (5.51) have a formal solution arising from the series $\sum_{m=0}^{\infty} P_m z^{-m}$ in Theorem 5.9.1. Suppose that the equations (5.51) have holomorphic solutions $\widehat{P}^{12}$ and $\widehat{P}^{21}$ on a sector $S^* \subseteq S$ for which

$$P(z) = I_\nu + \begin{pmatrix} 0 & \widehat{P}^{12}(z) \\ \widehat{P}^{21}(z) & 0 \end{pmatrix} \sim \sum_{m=0}^{\infty} P_m z^{-m}. \tag{5.52}$$

Then defining $\widehat{B}^{11}$ and $\widehat{B}^{22}$ using the first and fourth equations of (5.50) means that all four equations of (5.50) are satisfied and, with B and P defined by (5.48) and (5.49), the holomorphic change of variables $x = P(z)y$ transforms (5.34) into $z^{1-\rho}y' = B(z)y$, where B is a holomorphic block diagonal matrix on S^* given by

$$B(z) = \begin{pmatrix} B^{11}(z) & 0 \\ 0 & B^{22}(z) \end{pmatrix}.$$

Here (5.52) makes $P(z)$ invertible for large z, because P_0 is the identity. Moreover, B has an asymptotic series determined by the first and last equations of (5.50), and so the series $\sum_{m=0}^{\infty} B_m z^{-m}$ in Theorem 5.9.1 is an asymptotic series for B on S^*. Thus the key step is now to find holomorphic solutions $\widehat{P}^{12}$ and $\widehat{P}^{21}$ of (5.51) on a sector $S^* \subseteq S$ which satisfy (5.52).

Consider the first equation of (5.51), and write it in the form

$$z^{1-\rho}\left(\widehat{P}^{12}\right)' = A^{12} + A^{11}\widehat{P}^{12} - \widehat{P}^{12}A^{22} - \widehat{P}^{12}A^{21}\widehat{P}^{12}. \tag{5.53}$$

Now write the entries of $\widehat{P}^{12}$ as a column vector Y. Then (5.53) may be expressed as

$$z^{1-\rho}Y' = F(z, Y) = F_0(z) + F_1(z)Y + F_2(z, Y),$$

where the following conditions are satisfied: $F_0(z)$ and $F_2(z, Y)$ are column vectors; $F_0(z)$ is holomorphic in z on S, and independent of Y; the entries of $F_2(z, Y)$ are quadratic forms in the entries of Y, with coefficients which are holomorphic functions of z on S; the square matrix function $F_1(z)$ is holomorphic on S. Moreover, all the functions of z which appear as entries or coefficients in $F_0(z)$, $F_1(z)$ and $F_2(z, Y)$ have asymptotic series in S, and finite limits as $z \to \infty$ in S, because this is true of the entries of $A(z)$.

Now suppose that

$$F_1(\infty) = \lim_{z \to \infty, z \in S} F_1(z)$$

is not invertible. Then there exists a non-zero constant vector Y_0 such that $\lim_{z \to \infty} F_1(z)Y_0$ is the zero vector, and so there exists a non-zero constant matrix M_0 such that

$$\lim_{z \to \infty} (A^{11}M_0 - M_0A^{22}) = A_0^{11}M_0 - M_0A_0^{22}$$

is the zero matrix. This is impossible by Lemma 5.2.9, since A_0^{11} and A_0^{22} have no common eigenvalue, and so $F_1(\infty)$ is invertible.

To interpret the condition $\det F_1(\infty) \neq 0$, write $F_1(z) = (g_{jk}(z))$ and Y as the column vector $(Y_1, \ldots, Y_\tau)^T$, with $F(z, Y) = (F^1, \ldots, F^\tau)^T$. The jth entry of $F_1(z)Y$ is then $\sum_{k=1}^{\tau} g_{jk}(z)Y_k$, and since $F_2(z, Y)$ contains only quadratic terms in the Y_p, we get

$$\frac{\partial F^j}{\partial Y_k} = \frac{\partial}{\partial Y_k}\left(\sum_{k=1}^{\tau} g_{jk}(z)Y_k\right) = g_{jk}(z) + \text{terms involving the } Y_p.$$

It follows that

$$\lim_{z\to\infty, z\in S}\left(\frac{\partial F^j}{\partial Y_k}\right)_{Y=0} = F_1(\infty) \quad \text{is invertible.}$$

Furthermore, all these properties established for the first equation of (5.51) are shared by the second. Thus the existence of holomorphic matrix functions $\widehat{P}^{12}$ and $\widehat{P}^{21}$ which solve (5.51) and satisfy (5.52) on a sector $S^* \subseteq S$ is a consequence of the following theorem, which will be proved in the next subsection.

Theorem 5.9.2 *Let $\rho \in \mathbb{N}$ and suppose that in the differential equation*

$$z^{1-\rho}Y' = f(z, Y), \tag{5.54}$$

where Y and $f(z, Y)$ are N-dimensional column vectors, the function $f(z, Y)$ has the following properties on the sector S given by $|z| > R$, $|\arg z| < \alpha$, where $\alpha \leq \pi/2\rho$.
(i) If $Y = (Y_1, \ldots, Y_N)^T$ and $f(z, Y) = (f_1, \ldots, f_N)^T$, then each f_j is a polynomial in $Y_1, \ldots, Y_N$, with coefficients $a(z)$ which are holomorphic and bounded and each have an asymptotic series $a(z) \sim \sum_{m=0}^{\infty} a_m z^{-m}$ on S.
(ii) The matrix

$$\lim_{z\to\infty, z\in S}\left(\frac{\partial f_j}{\partial Y_k}\right)_{Y=0}$$

is invertible.
(iii) If the coefficients $a(z)$ of $f(z, Y)$ are replaced by their asymptotic series, then the equation (5.54) has a formal series solution

$$X(z) = \sum_{m=1}^{\infty} x_m z^{-m}, \tag{5.55}$$

where each x_m is a constant N-dimensional column vector.

Then in every sector S' given by $|\arg z| < \alpha' = \alpha - \varepsilon < \alpha$, the equation (5.54) has a holomorphic vector solution $Y = Y(z)$ satisfying

$$Y(z) \sim X(z) = \sum_{m=1}^{\infty} x_m z^{-m} \quad \text{as } z \to \infty \text{ in } S'. \tag{5.56}$$

The extension of Theorem 5.9.1 to encompass holomorphic solutions is then the following.

Theorem 5.9.3 *Let $\rho \in \mathbb{N}$ and let $A_0, A_1, \ldots$ be $\nu \times \nu$ constant matrices, and assume that there exists $\mu \in \{1, \ldots, \nu\}$ such that the eigenvalues of A_0 can be written as $\lambda_1, \ldots, \lambda_\nu$ in such a way that $\lambda_j \neq \lambda_k$ for $j \leq \mu$ and $k > \mu$. Then there exists a formal transformation $x = \sum_{m=0}^{\infty} P_m z^{-m} y$ which transforms (5.35) to (5.47), where P_0 is non-singular, while B_0 is similar to A_0, and B_m^{11} is $\mu \times \mu$ and B_m^{22} is $(\nu - \mu) \times (\nu - \mu)$.*

Furthermore, suppose that the $\nu \times \nu$ matrix function $A(z)$ is holomorphic, with asymptotic series $A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m}$, in a sector S given by $|z| > R > 0$, $\alpha < \arg z < \beta$, where $\beta - \alpha \leq \pi/\rho$. Then *in every sector S' given by $\alpha < \alpha' < \arg z < \beta' < \beta$, there exists a holomorphic matrix function $P(z)$ such that writing $x = P(z)y$ transforms the equation (5.34) to*

$$z^{1-\rho}y' = B(z)y, \quad B(z) = \begin{pmatrix} B^{11}(z) & 0 \\ 0 & B^{22}(z) \end{pmatrix}, \tag{5.57}$$

where B^{11} is $\mu \times \mu$ and B^{22} is $(\nu - \mu) \times (\nu - \mu)$. Moreover, P and B have asymptotic series

$$P(z) \sim \sum_{m=0}^{\infty} P_m z^{-m}, \quad B(z) \sim \sum_{m=0}^{\infty} B_m z^{-m}, \tag{5.58}$$

in S'. Finally, the w_j solve $z^{1-\rho}w_j' = B_{jj}(z)w_j$ in S', for $j = 1, 2$ if and only if $P(z)(w_1 \oplus w_2)$ solves $z^{1-\rho}x' = A(z)x$ in S', and the same correspondence holds for formal solutions.

Proof. If A_0 already has the block diagonal form (5.36), in which A_0^{11} and A_0^{22} have no eigenvalues in common, then the result follows from Theorems 5.9.1 and 5.9.2 and the discussion in between. In the general case, A_0 is similar to a Jordan matrix C_0 and, by Lemma 5.1.3, the Jordan blocks of C_0 can be permuted by a similarity transformation. Hence there exists a $\nu \times \nu$ constant matrix C such that

$$C^{-1}A_0C = J_0 = \begin{pmatrix} J_0^{11} & 0 \\ 0 & J_0^{22} \end{pmatrix},$$

in which J_0^{11} and J_0^{22} have no eigenvalues in common, and we may write

$$x = Cy, \quad z^{1-\rho}y' = C^{-1}z^{1-\rho}x' = C^{-1}A(z)Cy = J(z)y, \quad J(z) \sim J_0 + \sum_{m=1}^{\infty} C^{-1}A_mCz^{-m}.$$

Now applying the previous case, with A replaced by J, proves the theorem.

□

Remark. Suppose that the eigenvalues of A_0 are pairwise distinct. Since B_0 is similar to A_0, Theorem 5.9.3 may be used repeatedly, to split the system (5.35) into ν scalar equations, to each of which Theorem 5.8.1 may be applied, giving formal solutions and, in a suitable sector, holomorphic solutions with the formal solutions providing asymptotic series.

5.9.1 Proof of Theorem 5.9.2

Assume the hypotheses of Theorem 5.9.2 and write

$$b(z) = f(z, 0), \quad B(z) = \left(\frac{\partial f_j}{\partial Y_k}\right)_{Y=0}, \quad f(z, Y) = b(z) + B(z)Y + g(z, Y). \tag{5.59}$$

Here $b(z)$ and $g(z, Y)$ are column vectors and $B(z)$ is the Jacobian matrix of the f_j with respect to the variables Y_k, evaluated at $Y = 0$. Thus $b(z)$ is the part of $f(z, Y)$ which is independent of the Y_k, while $B(z)Y$ arises from the terms in $f(z, Y)$ which have total degree 1 in $Y_1, \ldots, Y_N$, and $g(z, Y)$ involves only terms of total degree at least 2.

By assumption (i), $B(z)$ has an asymptotic series $B(z) \sim \widetilde{B}(z) = \sum_{m=0}^{\infty} B_m z^{-m}$ on S, and assumption (ii) says that

$$B_0 = \lim_{z\to\infty, z\in S} B(z) = B(\infty)$$

is invertible, and so has only non-zero eigenvalues. The equation (5.54) becomes

$$z^{1-\rho}Y' = b(z) + B(z)Y + g(z, Y), \tag{5.60}$$

and it may be assumed that B_0 is in Jordan form. If this is not the case then there exists a constant invertible matrix M such that MB_0M^{-1} is in Jordan form and writing $W = MY$ transforms (5.60) to

$$z^{1-\rho}W' = Mb(z) + MB(z)Y + Mg(z, Y) = Mb(z) + MB(z)M^{-1}W + Mg(z, M^{-1}W).$$

Hence we may write

$$B_0 = \Lambda = D + H, \quad D = \mathrm{diag}\,\{\lambda_1, \ldots, \lambda_N\}, \quad H = H_1 \oplus \ldots \oplus H_s, \tag{5.61}$$

in which the λ_j are the eigenvalues of B_0, none of which are 0, and the H_j are upper triangular shifting matrices of appropriate dimensions. In particular, D commutes with H, because B_0 is in Jordan form.

Take α'' with $\alpha'' - \alpha$ small and positive and apply Theorem 5.7.1 to generate a holomorphic matrix function $\phi(z)$ on the sector S'' given by $|\arg z| < \alpha''$ with asymptotic series

$$\phi(z) \sim X(z) = \sum_{m=1}^{\infty} x_m z^{-m} \tag{5.62}$$

as $z \to \infty$ on S''. The fact that S'' is a slightly wider sector allows term by term differentiation of (5.62) on S.

Now write

$$Y = u + \phi(z), \tag{5.63}$$

so that (5.60) gives

$$z^{1-\rho}u' = b(z) + B(z)u + B(z)\phi(z) - z^{1-\rho}\phi'(z) + g(z, u + \phi(z)). \tag{5.64}$$

The aim is to show that (5.64) has a solution u with asymptotic series 0 on the smaller sector S', so that (5.63) gives a solution $Y = u + \phi$ of (5.60) with asymptotic series X.

Lemma 5.9.1 *The equation (5.64) may be written in the form*

$$z^{1-\rho}u' = \Lambda u + p(z, u), \quad u = (u_1, \ldots, u_N)^T, \tag{5.65}$$

in which $p(z, u)$ has the following properties.
(i) The entries of $p(z, u)$ are polynomials in the u_j, with coefficients which are holomorphic and bounded and have asymptotic series on S.
(ii) $p(z, 0) \sim 0$ as $z \to \infty$ in S.
(iii) If $u(z)$ is bounded on a subsector $\widetilde{S}$ of S, and if $m \in \mathbb{N}$ and $u(z) = O(|z|^{-m})$ as $z \to \infty$ in S, then $p(z, u(z))$ is bounded on $\widetilde{S}$ and satisfies $p(z, u(z)) = O(|z|^{-m-1})$ as $z \to \infty$ in $\widetilde{S}$.
(iv) To each $\gamma > 0$ corresponds $\varepsilon_0 > 0$ such that

$$\max\{\|u\|, \|w\|\} \leq \varepsilon_0 \Rightarrow \|p(z, u) - p(z, w)\| < \gamma\|u - w\| \quad \text{as } z \to \infty \text{ in } S. \tag{5.66}$$

Proof. The function $\phi(z)$ has on the slightly wider sector S'' the asymptotic series $X(z)$, which is a formal solution of (5.54), and hence of (5.60), and each entry of $g(z, Y)$ is is a polynomial in the entries of Y with coefficients having asymptotic series. This implies that, on S, with the symbol $\widehat{\ }$ denoting that a term in z is replaced by its asymptotic series,

$$\begin{aligned} c(z) &= b(z) + B(z)\phi(z) - z^{1-\rho}\phi'(z) + g(z, \phi(z)) \\ &\sim \widehat{b}(z) + \widehat{B}(z)X(z) - z^{1-\rho}X'(z) + \widehat{g}(z, X(z)) = \widehat{f}(z, X(z)) - z^{1-\rho}X'(z) = 0. \end{aligned}$$

Hence the equation (5.64) can be written in the form

$$z^{1-\rho}u' = B(z)u + g(z, u + \phi(z)) - g(z, \phi(z)) + c(z), \quad \text{where } c(z) \sim 0 \text{ on } S. \tag{5.67}$$

Now write ϕ as a column vector $\phi = (\phi_1, \ldots, \phi_N)^T$. Any term which appears in $g(z, Y)$ has form $a(z)Y_1^{p_1} \ldots Y_N^{p_N}$, where $p_1 + \ldots p_N \geq 2$ and $a(z)$ is holomorphic and bounded on S, with an asymptotic series there. For any of the finitely many terms $a(z)Y_1^{p_1} \ldots Y_N^{p_N}$ appearing in $g(z, Y)$ we can expand

$$a(z)(u_1 + \phi_1)^{p_1} \ldots (u_N + \phi_N)^{p_N} - a(z)\phi_1^{p_1} \ldots (\phi_N)^{p_N}$$

in terms of the u_j and ϕ_k. After cancellation, no term is independent of the u_j, and any term which has total degree 1 in the u_j has at least one ϕ_k as a factor. Moreover, $a(z)$ and the $\phi_k(z)$ have asymptotic series in S, and (5.62) implies that each $\phi_k(z)$ tends to 0 as $z \to \infty$ in S. It follows that we can write

$$g(z, u + \phi(z)) - g(z, \phi(z)) = B^*(z)u + h(z, u), \tag{5.68}$$

where $B^*(z)$ and $h(z, u)$ have coefficients which are bounded and have asymptotic series in S, while $h(z, u)$ has no terms of total degree less than 2 in the u_j, and $B^*(z) \to 0$ as $z \to \infty$ in S.

The equation (5.67) can now be written in the form

$$z^{1-\rho}u' = C(z)u + c(z) + h(z, u), \tag{5.69}$$

in which

$$\lim_{z \to \infty, z \in S} C(z) = \lim_{z \to \infty, z \in S} (B(z) + B^*(z)) = B_0 = \Lambda = D + H. \tag{5.70}$$

This gives (5.65), with

$$p(z, u) = (C(z) - \Lambda)u + c(z) + h(z, u),$$

and assertions (i), (ii) and (iii) hold, in view of (5.67), (5.69), (5.70) and the fact that $h(z, u)$ has no terms of total degree less than 2 in the u_j.

To prove (iv) write

$$p(z, u) - p(z, w) = h(z, u) - h(z, w) + (C(z) - \Lambda)(u - w), \quad w = (w_1, \ldots, w_N)^T, \tag{5.71}$$

and take $\gamma > 0$. We then have, by (5.70),

$$\|(C(z) - \Lambda)(u - w)\| < \frac{\gamma \|u - v\|}{2} \quad \text{as } z \to \infty \text{ in } S. \tag{5.72}$$

Furthermore, as observed following (5.68), $h(z, u)$ is a sum of finitely many terms $H(z)u_1^{q_1} \ldots u_N^{q_N}$, with bounded coefficients $H(z)$ which have asymptotic series in S, and with $q_1 + \ldots q_N \geq 2$. Writing $u_j = w_j + \sigma_j$ shows that $h(z, u) - h(z, w)$ is a sum of terms

$$H(z)\left((w_1 + \sigma_1)^{q_1} \ldots (w_N + \sigma_N)^{q_N} - w_1^{q_1} \ldots w_N^{q_N}\right),$$

each of which contains no terms independent of the σ_j. Combining this ob Hence there exists $\varepsilon_0 > 0$ such that (5.66) holds. $\square$

The next step is to write, using the fact that Λ is a constant matrix,

$$V(z) = \exp\left(\frac{z^\rho \Lambda}{\rho}\right), \quad V'(z) = z^{\rho-1}\Lambda V(z), \quad V(z)^{-1} = \exp\left(-\frac{z^\rho \Lambda}{\rho}\right). \tag{5.73}$$

It will suffice to find a solution $u \sim 0$ on S' of the integral equation

$$u(z) = \int^z V(z)V(t)^{-1}t^{\rho-1}p(t,u(t))\,dt, \tag{5.74}$$

because if V solves (5.74) then (5.73) gives (5.65) via

$$\begin{aligned} u'(z) &= z^{\rho-1}p(z,u) + \int^z V'(z)V(t)^{-1}t^{\rho-1}p(t,u(t))\,dt \\ &= z^{\rho-1}p(z,u) + z^{\rho-1}\Lambda \int^z V(z)V(t)^{-1}t^{\rho-1}p(t,u(t))\,dt \\ &= z^{\rho-1}p(z,u) + z^{\rho-1}\Lambda u(z). \end{aligned}$$

We use (5.61) to write (5.74) in the form

$$u(z) = \int^z \exp\left(\frac{z^\rho D}{\rho} - \frac{t^\rho D}{\rho}\right) \exp\left(\frac{z^\rho H}{\rho} - \frac{t^\rho H}{\rho}\right) t^{\rho-1}p(t,u(t))\,dt, \tag{5.75}$$

and employ the change of variables

$$\zeta = \frac{z^\rho}{\rho}, \quad \tau = \frac{t^\rho}{\rho}, \quad v(\zeta) = u(z), \quad p(t,u(t)) = q(\tau, v(\tau)), \tag{5.76}$$

noting that ζ maps the sector S into the sector Σ given by $|\arg w| < \beta = \rho\alpha \leq \pi/2$. Thus it now suffices to find, on a suitable sector, a holomorphic solution $v \sim 0$ of

$$v(\zeta) = \int^\zeta \exp\left((\zeta-\tau)D\right)\exp\left((\zeta-\tau)H\right) q(\tau, v(\tau))\,d\tau. \tag{5.77}$$

Since v and $q(\tau, v(\tau))$ are $1 \times N$, we will choose for the jth entry of $v(\zeta)$ a path $\delta_j(\zeta)$, terminating at ζ, and write (5.77) in the form

$$v(\zeta) = \int_{\Delta(\zeta)} \exp\left((\zeta-\tau)D\right)\exp\left((\zeta-\tau)H\right) q(\tau, v(\tau))\,d\tau, \tag{5.78}$$

where $\Delta(\zeta)$ denotes the collection of paths $\delta_j(\zeta)$. Here D is a diagonal matrix, and therefore so is $\exp(xD)$, while H commutes with D and is nilpotent.

The paths $\delta_j(\zeta)$ will now be chosen, and the aim is to do this so that $\exp(\lambda_j(\zeta-\tau))$ is small for τ on $\delta_j(\zeta)$, for each eigenvalue λ_j of B_0, each of which is non-zero by assumption. Take β' with $\beta - \beta'$ small and positive, such that there is no λ_j with $\mathrm{Re}\left(\lambda_j e^{\pm i\beta'}\right) = 0$. Let

$$\Sigma_0 = \{w \in \mathbb{C} : |\arg w| < \beta'\}, \quad \Sigma_1 = \{w \in \mathbb{C} : |\arg(w - \zeta_1)| < \beta'\}, \quad 2 \leq \zeta_1 \in \mathbb{R}, \tag{5.79}$$

so that $\Sigma_1 \subseteq \Sigma_0 \subseteq \Sigma$. Assume that ζ_1 is so large that $\zeta \in \Sigma_1$ gives $z \in S$. Provided β' was chosen close enough to β, we have $\zeta \in \Sigma_1$ for all sufficiently large $z \in S'$, and so it will be enough to find a solution $v \sim 0$ of (5.77) on Σ_1.

An eigenvalue λ_j will be called class I if $\mathrm{Re}\left(\lambda_j e^{i\theta}\right) < 0$ for $-\beta' < \theta < \beta'$, and class II otherwise. Suppose first that λ_j is class I. Then $\mathrm{Re}\left(\lambda_j e^{i\theta}\right) < 0$ for $-\beta' \leq \theta \leq \beta'$, by the choice of β', and so there exists $c_0 > 0$ such that $\mathrm{Re}\left(\lambda_j e^{i\theta}\right) < -c_0$ for $-\beta' \leq \theta \leq \beta'$. It follows that

$$\mathrm{Re}\,(\sigma\lambda_j) \leq -c_0|\sigma\lambda_j|$$

for all $\sigma \in \Sigma_0$. For $\zeta \in \Sigma_1$ we choose $\delta_j(\zeta)$ to be the straight line segment from ζ_1 to ζ. If $t \in \delta_j(\zeta)$ then $\zeta - \tau \in \Sigma_0 \cup \{0\}$, which implies that

$$\mathrm{Re}\,((\zeta-\tau)\lambda_j) \leq -c_0|(\zeta-\tau)\lambda_j| \quad \text{for } t \in \delta_j(\zeta). \tag{5.80}$$

Now suppose that λ_j is class II. Then there exists $\theta_j \in (-\beta', \beta')$ with $\mathrm{Re}\,\left(\lambda_j e^{i\theta}\right) > 0$, which gives $d > 0$ such that $\mathrm{Re}\,(\sigma\lambda_j) \geq d|\sigma|$ on the ray $\arg\sigma = \theta_j$. For $\zeta \in \Sigma_1$ we choose $\delta_j(\zeta) \subseteq \Sigma_0$ to be the half-line given by $\tau = \zeta + re^{i\theta_j}$, $r \geq 0$, which gives (5.80) again, after reducing c_0 if necessary. Here we choose the direction of travel to be from infinity to ζ, in accordance with (5.77).

Lemma 5.9.2 *There exists $K \geq 1$, depending only on the constant c_0 in (5.80) and the matrix Λ, with the following property. Let $d_0 > 0$ and let Σ_1 and ζ_1 be as in (5.79), and let $\chi(\zeta)$ be a holomorphic N-dimensional vector function on Σ_1, satisfying there $\|\chi(\zeta)\| \leq d_0|\zeta|^{-1}$. Then*

$$\psi(\zeta) = \int_{\Delta(\zeta)} \exp((\zeta-\tau)\Lambda)\chi(\tau)\,d\tau$$

is holomorphic on Σ_1, and satisfies $\|\psi(\zeta)\| \leq Kd_0|\zeta|^{-1}$ there.

We emphasise that K does not depend on ζ_1 here.

Proof. By considering $\chi(\zeta)/d_0$ and $\psi(\zeta)/d_0$, it may be assumed that $d_0 = 1$. For $\zeta \in \Sigma_1$ write

$$\psi(\zeta) = \int_{\Delta(\zeta)} \exp((\zeta-\tau)D)L(\zeta,\tau)\,d\tau, \quad L(\zeta,\tau) = \exp((\zeta-\tau)H)\chi(\tau).$$

Denote by $c_1, c_2, \ldots$ positive constants which depend at most on c_0 and Λ. Since H is nilpotent, $\exp(xH)$ is a matrix whose entries are polynomials in x, and so

$$\|\exp((\zeta-\tau)H)\| \leq c_1 + c_2|\zeta-\tau|^{c_3}. \tag{5.81}$$

Because D is a diagonal matrix, the jth entry of $\psi(z)$ is

$$\psi_j(\zeta) = \int_{\delta_j(\zeta)} e^{(\zeta-\tau)\lambda_j} L_j(\zeta,\tau)\,d\tau, \tag{5.82}$$

where $L_j(\zeta,\tau)$ is the jth entry of $L(\zeta,\tau)$.

Suppose that λ_j is class I, so that $\delta_j(\zeta)$ is the line segment from ζ_1 to ζ. Observe that in this case the initial point ζ_1 of $\delta_j(\zeta)$ does not lie in Σ_1, but all other points on $\delta_j(\zeta)$ do, and the existence of the integral is unaffected, because of the uniform bound for $\chi(\tau)$ as $\tau \to \zeta_1$ in Σ_1. Let $\delta_j^1(\zeta)$ be the part of $\delta_j(\zeta)$ on which $|\tau| \geq |\zeta|/2$. Then $\delta_j^1(\zeta)$ can be parametrised with respect to $s = |\zeta-\tau|$, giving an estimate $|d\tau| \leq c_4 ds$. Thus, by (5.80) and (5.81), the contribution of this part to $\psi_j(\zeta)$ has modulus at most

$$\begin{aligned} M_{j,1}(\zeta) &= \int_{\delta_j^1(\zeta)} e^{-c_5|\zeta-\tau|}(c_1 + c_2|\zeta-\tau|^{c_3})|\tau|^{-1}\,|d\tau| \\ &\leq c_6|\zeta|^{-1}\int_0^\infty e^{-c_5 s}(c_1 + c_2 s^{c_3})\,ds \leq c_7|\zeta|^{-1}. \end{aligned}$$

Next, let $\delta_j^2(\zeta)$ be the part of $\delta_j(\zeta)$ on which $|\tau| \leq |\zeta|/2$. Then $\delta_j^2(\zeta)$ has length at most $c_8|\zeta|$, while $|\zeta|/2 \leq |\zeta-\tau| \leq |\zeta|$ and $|\tau| \geq \zeta_1 \geq 2$ on $\delta_j(\zeta)$. We apply (5.80) and (5.81) again, and conclude that the contribution of this part to $\psi_j(\zeta)$ has modulus at most

$$M_{j,2}(\zeta) = \int_{\delta_j^2(\zeta)} e^{-c_9|\zeta|}(c_1 + c_2|\zeta|^{c_3})c_{10}\,|d\tau| \leq c_{11}|\zeta|e^{-c_9|\zeta|}(c_1 + c_2|\zeta|^{c_3}) \leq c_{12}|\zeta|^{-1}.$$

Suppose now that λ_j is class II, so that $\delta_j(\zeta)$ is the half-line $\tau = \zeta + re^{i\theta_j}$, $r \geq 0$, on which

$$|L_j(\zeta,\tau)| \leq (c_1 + c_2|\zeta-\tau|^{c_3})|\tau|^{-1} \leq (c_1 + c_2|\zeta-\tau|^{c_3})c_{13}|\zeta|^{-1}.$$

Again $\delta_j(\zeta)$ can be parametrised with respect to $s = |\zeta - \tau|$, giving an estimate $|d\tau| \leq c_{14}ds$. Thus (5.80) implies that $\psi_j(\zeta)$ in (5.82) has modulus at most

$$\begin{aligned} M_j(\zeta) &= c_{15}|\zeta|^{-1}\int_{\delta_j(\zeta)} e^{-c_{16}|\zeta-\tau|}(c_1 + c_2|\zeta-\tau|^{c_3})\,|d\tau| \\ &\leq c_{17}|\zeta|^{-1}\int_0^\infty e^{-c_{16}r}(c_1 + c_2 r^{c_3})\,dr = c_{18}|\zeta|^{-1}. \end{aligned}$$

This proves that the integral converges, with the required estimate.

To show that $\psi(\zeta)$ is holomorphic on Σ_1, fix $\zeta_2 \in \Sigma_1$. For ζ close to ζ_2, since $\delta_j(\zeta)$ lies in $\Sigma_1 \cup \{\zeta_1\}$, and $\chi(\tau)$ is bounded as $\tau \to \zeta_1$ in Σ_1, Cauchy's theorem implies that

$$\begin{aligned} \psi(\zeta) &= \int_{\Delta(\zeta_2)} \exp((\zeta-\tau)\Lambda)\chi(\tau)\,d\tau + \int_{\zeta_2}^{\zeta} \exp((\zeta-\tau)\Lambda)\chi(\tau)\,d\tau \\ &= \exp(\zeta\Lambda)\left(\int_{\Delta(\zeta_2)} \exp(-\tau\Lambda)\chi(\tau)\,d\tau + \int_{\zeta_2}^{\zeta} \exp(-\tau\Lambda)\chi(\tau)\,d\tau\right), \end{aligned}$$

from which the assertion evidently follows. □

The integral equation (5.77) will now be solved via a fairly standard iterative method, by setting

$$P(v) = P(v(\zeta)) = \int_{\Delta(\zeta)} \exp((\zeta-\tau)\Lambda)q(\tau, v(\tau))\,d\tau \tag{5.83}$$

and

$$v_0 = 0, \quad v_{n+1} = P(v_n). \tag{5.84}$$

Take $\gamma \in (0, 1/K)$, where K is as in Lemma 5.9.2, and let ε_0 be as in Lemma 5.9.1. Choose $d > 0$ such that (5.76) and Lemma 5.9.1 give

$$\|q(\zeta, 0)\| = \|p(z, 0)\| \leq \frac{d}{|\zeta|} \quad \text{for } \zeta \in \Sigma_1. \tag{5.85}$$

Assume that ζ_1 is so large that (5.66) holds on Σ_1 and

$$\frac{Kd}{(1-\gamma K)|\zeta|} < \varepsilon_0 \quad \text{for } \zeta \in \Sigma_1. \tag{5.86}$$

Lemma 5.9.2, (5.84) and (5.85) now imply that

$$v_1(\zeta) = P(v_0) = \int_{\Delta(\zeta)} \exp((\zeta-\tau)\Lambda)q(\tau, 0)\,d\tau$$

is holomorphic on Σ_1 with

$$\|v_1(\zeta)\| = \|v_1(\zeta) - v_0(\zeta)\| \leq \frac{Kd}{|\zeta|} \quad \text{for } \zeta \in \Sigma_1. \tag{5.87}$$

To take the iteration further, we assert that, for $\zeta \in \Sigma_1$ and $n = 0, 1, \ldots$,

$$\begin{aligned} \|v_{n+1}(\zeta) - v_n(\zeta)\| &\le \frac{\gamma^n K^{n+1} d}{|\zeta|} < (\gamma K)^n (1 - \gamma K)\varepsilon_0, \\ \|v_{n+1}(\zeta)\| &\le \frac{Kd}{(1-\gamma K)|\zeta|} < \varepsilon_0. \end{aligned} \tag{5.88}$$

This is true for $n = 0$, by (5.86) and (5.87). Assume next that $n > 0$ and that (5.88) holds with n replaced by any smaller non-negative integer, and write $u_j(z) = v_j(\zeta)$. Then we have

$$\max\{\|u_n(z), \|u_{n-1}(z)\|\} < \varepsilon_0.$$

For $\zeta \in \Sigma_1$ we then get, using (5.66) and (5.76),

$$\begin{aligned} \|q(\zeta, v_n(\zeta)) - q(\zeta, v_{n-1}(\zeta))\| &= \|p(z, u_n(z)) - p(z, u_{n-1}(z))\| \\ &\le \gamma \|u_n(z) - u_{n-1}(z)\| \\ &= \gamma \|v_n(\zeta) - v_{n-1}(\zeta)\| \\ &\le \frac{\gamma^n K^n d}{|\zeta|}, \end{aligned}$$

from which Lemma 5.9.2 and (5.86) give

$$\begin{aligned} \|v_{n+1}(\zeta) - v_n(\zeta)\| &= \|P(v_n) - P(v_{n-1})\| \\ &= \left\| \int_{\Delta(\zeta)} \exp((\zeta - \tau)\Lambda)(q(\tau, v_n(\tau)) - q(\tau, v_{n-1}(\tau))\, d\tau \right\| \\ &\le \frac{\gamma^n K^{n+1} d}{|\zeta|} < (\gamma K)^n (1 - \gamma K)\varepsilon_0, \end{aligned}$$

which proves the first inequality of (5.88). We also obtain, again in view of (5.86),

$$\|v_{n+1}(\zeta)\| \le \sum_{j=0}^{n} \|v_{j+1}(\zeta) - v_j(\zeta)\| \le \sum_{j=0}^{n} \frac{\gamma^j K^{j+1} d}{|\zeta|} \le \frac{Kd}{(1-\gamma K)|\zeta|} < \varepsilon_0,$$

which completes the induction.

Now (5.88) implies that

$$v_{n+1} = \sum_{j=0}^{n} (v_{j+1} - v_j)$$

converges uniformly on Σ_1 to some holomorphic v, and $\max\{\|v_n\|, \|v\|\} \le \varepsilon_0$ on Σ_1, for all $n \ge 0$. This yields, for $\zeta \in \Sigma_1$,

$$\begin{aligned} \|v(\zeta) - v_n(\zeta)\| &= \lim_{m\to\infty} \|v_{m+1}(\zeta) - v_n(\zeta)\| \le \lim_{m\to\infty} \sum_{j=n}^{m} \|v_{j+1}(\zeta) - v_j(\zeta)\| \\ &= \sum_{j=n}^{\infty} \|v_{j+1}(\zeta) - v_j(\zeta)\| \le \sum_{j=n}^{\infty} \frac{\gamma^j K^{j+1} d}{|\zeta|} = \frac{\gamma^n K^{n+1} d}{(1-\gamma K)|\zeta|}. \end{aligned} \tag{5.89}$$

Thus applying Lemmas 5.9.1 and 5.9.2 again yields

$$\|q(\zeta, v(\zeta)) - q(\zeta, v_n(\zeta))\| \le \frac{(\gamma K)^{n+1} d}{(1-\gamma K)|\zeta|}, \quad \|P(v) - P(v_n)\| \le \frac{K(\gamma K)^{n+1} d}{(1-\gamma K)|\zeta|},$$

so that

$$v_{n+1} = P(v_n) \to P(v), \quad v_{n+1} \to v, \quad v = P(v).$$

Hence v is a solution of the integral equation (5.77), and satisfies $v(\zeta) = O(|\zeta|^{-1})$ as $\zeta \to \infty$ in Σ_1, by the second estimate of (5.88).

It remains to show that $v \sim 0$ on Σ_1. To this end, suppose that m is a positive real number and $V(\zeta) = O(|\zeta|^{-m})$ as $\zeta \to \infty$ in Σ_1. It follows from Lemma 5.9.1, with $u(z) = v(\zeta) = O(|z|^{-\rho m})$, that

$$q(\zeta, v(\zeta)) = p(z, u(z)) = O(|z|^{-\rho m - 1}) = O(|\zeta|^{-m-1/\rho}) \tag{5.90}$$

as $\zeta \to \infty$ in Σ_1. Let $\zeta \in \Sigma_1$ be large, and consider the equation

$$v(\zeta) = \int_{\Delta(\zeta)} \exp((\zeta - \tau)D) \exp((\zeta - \tau)H) q(\tau, v(\tau)) \, d\tau, \tag{5.91}$$

which holds by (5.83). As in the proof of Lemma 5.80, the jth entry of the right-hand side of (5.91) is

$$\Psi_j(\zeta) = \int_{\delta_j(\zeta)} e^{(\zeta-\tau)\lambda_j} K_j(\zeta, \tau) \, d\tau, \tag{5.92}$$

where $K_j(\zeta, \tau)$ is the jth entry of $\exp((\zeta - \tau)H)q(\zeta, v(\zeta))$. Denote by $e_1, e_2, \ldots$ positive constants.

Suppose that λ_j is class I, and as before let $\delta_j^1(\zeta)$ be the part of $\delta_j(\zeta)$ on which $|\tau| \geq |\zeta|/2$. By (5.80), (5.81) and (5.90) and arguments similar to those in the proof of Lemma 5.9.2, the contribution of this part to $\Psi_j(\zeta)$ has modulus at most

$$\begin{aligned} N_{j,1}(\zeta) &= \int_{\delta_j^1(\zeta)} e^{-e_1|\zeta-\tau|}(e_2 + e_3|\zeta - \tau|^{e_4}) \, e_4|\tau|^{-m-1/\rho} \, |d\tau| \\ &\leq e_5|\zeta|^{-m-1/\rho} \int_0^\infty e^{-e_1 s}(e_2 + e_3 s^{e_4}) \, ds \leq e_6|\zeta|^{-m-1/\rho}. \end{aligned}$$

Next, let $\delta_j^2(\zeta)$ be the part of $\delta_j(\zeta)$ where $|\tau| \leq |\zeta|/2$, on which we then have $|\zeta|/2 \leq |\zeta - \tau| \leq |\zeta|$. We apply (5.80), (5.81) and (5.90) again, as well as Lemma 5.9.1(iii), and the contribution of this part to $\Psi_j(\zeta)$ has modulus at most

$$\begin{aligned} N_{j,2}(\zeta) &= \int_{\delta_j^2(\zeta)} e^{-e_1|\zeta-\tau|}(e_2 + e_3|\zeta - \tau|^{e_4}) \, e_7 \, |d\tau| \\ &\leq e^{-e_8|\zeta|}(e_9 + e_{10}|\zeta|^{e_{11}}) \int_{\delta_j^2(\zeta)} e_7 \, |d\tau| \\ &\leq e^{-e_8|\zeta|}(e_9 + e_{10}|\zeta|^{e_{11}}) e_{12}|\zeta| \leq e_{13}|\zeta|^{-m-1/\rho}, \end{aligned}$$

since ζ is large.

Now suppose that λ_j is class II. Then (5.80), (5.81) and (5.90) imply that $\Psi_j(\zeta)$ in (5.92) has modulus at most

$$\begin{aligned} N_j(\zeta) &= \int_{\delta_j(\zeta)} e^{-e_1|\zeta-\tau|}(e_2 + e_3|\zeta - \tau|^{e_4}) \, e_4|\tau|^{-m-1/\rho} \, |d\tau| \\ &\leq e_{14}|\zeta|^{-m-1/\rho} \int_0^\infty e^{-e_1 r}(e_2 + e_3 r^{e_4}) \, dr = e_{15}|\zeta|^{-m-1/\rho}. \end{aligned}$$

We have thus shown that an estimate $v(\zeta) = O(|\zeta|^{-m})$ as $\zeta \to \infty$ in Σ_1 can be improved to $v(\zeta) = O(|\zeta|^{-m-1/\rho})$ as $\zeta \to \infty$ in Σ_1. Since we already have such an estimate with $m = 1$, this completes the proof of Theorem 5.9.2. □

The above proof is based on [72, pp.65-75], but some simplifications have been made. In particular, the last part of the present proof avoids the need for repeated changes to the apex point ζ_1 of the sector Σ_1.

5.10 The shearing method

This is based on Balser's text [4, pp. 45-52], but with some modifications. The matrix differential equation (5.34) can be written in the form

$$zx' = \widehat{A}(z)x, \quad \widehat{A}(z) = z^{\rho}A(z), \quad \rho \in \mathbb{N} = \{1, 2, \ldots\}. \tag{5.93}$$

Here the holomorphic and formal cases will be treated simultaneously, so that A will either be a holomorphic $\mu \times \mu$ matrix function on a sector S, satisfying (in the sense of asymptotic series)

$$A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m} \quad \text{as } z \to \infty \text{ in } S, \tag{5.94}$$

or simply a formal series as on the right-hand side of (5.94), in which case we will still write $A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m}$. It will be assumed throughout this section that A_0 is not the zero matrix: the integer $\rho \in \mathbb{N}$ will then be called the *rank* of the equation, as in §5.9.

5.10.1 A transformation of the system

Given the system (5.93), with A satisfying (5.94), write $x = T(z)y$, where $T(z)$ is an invertible matrix. The equation (5.93) transforms to

$$zy' = \widehat{B}(z)y, \quad \widehat{B}(z) = z^{\rho}B(z) = T(z)^{-1}\widehat{A}(z)T(z) - zT(z)^{-1}T'(z), \tag{5.95}$$

so that

$$B(z) = T(z)^{-1}A(z)T(z) - z^{1-\rho}T(z)^{-1}T'(z). \tag{5.96}$$

In the particular case where T is a constant non-singular matrix, (5.96) takes the simple form $B(z) = T^{-1}A(z)T$.

Note that writing $U(z) = T(z)^{-1}$ in (5.96) gives

$$T(z)B(z)U(z) = A(z) - z^{1-\rho}T'(z)U(z)$$

and

$$A(z) = U(z)^{-1}B(z)U(z) + z^{1-\rho}T'(z)U(z) = U(z)^{-1}B(z)U(z) - z^{1-\rho}U(z)^{-1}U'(z),$$

since $I = TU$ gives $0 = T'U + TU'$. Thus A is recoverable from B and, in this sense, the transformation is reversible.

This transformation will be used in both the formal and holomorphic settings. In the formal case, T will be a formal matrix series in descending powers of $z^{1/p}$, with $p \in \mathbb{N}$ and $\det T(z)$ not the zero series, in which case $T(z)^{-1}$ is also a formal matrix series in descending powers of $z^{1/p}$, and so is $U(z)$. Furthermore, x is a formal solution of (5.93) if and only if $y = U(z)x$ is a solution of (5.95).

Turning to holomorphic solutions on a sector S, if T is a matrix function which is holomorphic and non-singular for large z in S, then x is a holomorphic solution of (5.93) if and only if y is a holomorphic solution of (5.95). Assume that one of T and $U = T^{-1}$ is represented on S by an asymptotic series in descending powers of some $z^{1/p}$, with determinant which is not the zero series, and that A has an asymptotic series in descending powers of z. Then T and U both have asymptotic series in descending powers of $z^{1/p}$, and so has B, by (5.96). Indeed, if we denote by $\widetilde{G}$ the asymptotic series for a matrix function G, then (5.96) translates to

$$\widetilde{B}(z) = \widetilde{U}(z)\widetilde{A}(z)\widetilde{T}(z) - z^{1-\rho}\widetilde{U}(z)\widetilde{T}'(z),$$

and holomorphic and formal solutions of (5.95) are obtained from those of (5.93) via premultiplying by U and $\widetilde{U}$ respectively.

The two systems will then be referred to as *equivalent* (with the caveat that $B(z)$ may involve fractional powers of z whereas $A(z)$ did not).

5.10.2 Systems in standard nilpotent form

Assume that the system (5.93) satisfies (5.94) with $A_0 \neq (0)$, but that the lead matrix A_0 in the formal/asymptotic series (5.94) is nilpotent, that is, A_0 satisfies $A_0^t = (0)$ for some $t \in \mathbb{N}$. The system will be said to be in *standard nilpotent form* if this A_0 is a block matrix of form

$$(0) \neq A_0 = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & M_\mu \end{pmatrix} = \mathrm{diag}\,(M_1, \dots, M_\mu). \tag{5.97}$$

Here the M_j are upper triangular shifting matrices

$$M_j = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 \end{pmatrix}$$

of dimensions s_j, with $s_1 \geq \ldots \geq s_\mu$ (and $\sum_{j=1}^{\mu} s_j = \nu$). The matrix A is then written in the same block configuration as A_0, that is,

$$A = \begin{pmatrix} A^{11} & A^{12} & \dots & A^{1\mu} \\ A^{21} & A^{22} & \dots & A^{2\mu} \\ \dots & \dots & \dots & \dots \\ A^{\mu 1} & \dots & A^{\mu(\mu-1)} & A^{\mu\mu} \end{pmatrix}. \tag{5.98}$$

It follows from (5.94) that each of these blocks either has an asymptotic series

$$A^{jk}(z) \sim \sum_{m=0}^{\infty} A_m^{jk} z^{-m}, \tag{5.99}$$

or is a formal series of this type. Here the block A^{jk} is $s_j \times s_k$ and so:
has at least as many rows as columns if $j \leq k$;
has at least as many columns as rows if $j \geq k$.

Lemma 5.10.1 *If the system (5.93) satisfies (5.94) with $A_0 \neq (0)$ nilpotent, then there exists a constant non-singular matrix S such that $x = Sy$ transforms (5.93) to a system (5.95) in standard nilpotent form.*

Proof. §5.10.1 shows that a transformation $x = Ty$, with T a constant non-singular matrix, replaces $A(z)$ by $T^{-1}A(z)T$. By applying this process repeatedly, and using Lemma 5.1.3, it may be assumed first that the lead matrix A_0 is in Jordan form, all its eigenvalues being 0 since A_0 is nilpotent, and second that the Jordan blocks have descending dimensions, that is, the Jordan form corresponds to standard nilpotent form.

5.10.3 Equations in normalised form

Assume that the system (5.93) satisfies (5.94) and choose a large positive integer M. The system will be said to be *normalised up to order M* if it is in standard nilpotent form as in §5.10.2 and the coefficients A_m^{jk} of the asymptotic series (5.99) for the blocks A^{jk} in (5.98) satisfy the following conditions for $1 \leq m \leq M$:

(i) for $j \geq k$ (i.e. blocks on or below the diagonal) all non-zero entries of the matrix A_m^{jk} lie in the first column;
(ii) for $j < k$ (i.e. blocks strictly above the diagonal) all non-zero entries of the matrix A_m^{jk} lie in the last row.

The next lemma says that every system (5.93) in standard nilpotent form can be transformed to a normalised system, for an arbitrarily large choice of M, and with the same lead matrix A_0.

Lemma 5.10.2 *Let the system (5.93) satisfy (5.94) and be in standard nilpotent form. Let $M \in \mathbb{N}$. Then there exists a transformation $x = T(z)y$ with*

$$T(z) = \sum_{m=1}^{M} T_m z^{-m}, \quad T_0 = I, \tag{5.100}$$

so that the transformed equation $zy' = \widehat{B}(z)y = z^\rho B(z)y$ is normalised up to order M and has $B_0 = A_0$ and the same rank ρ as (5.93).

Proof. With $T(z) = I + \delta(z)$ given by (5.100), where the coefficient matrices T_m are to be determined, it is clear that

$$\delta(\infty) = (0), \quad T(z)^{-1} = I - \delta(z) + \delta(z)^2 - \dots$$

for large $z \in S$, and so both T and T^{-1} are given by asymptotic series. Moreover, by (5.96), B is either holomorphic with an asymptotic series, or itself a formal series. Write

$$B(z) \sim \sum_{m=0}^{\infty} B_m z^{-m}, \quad T(z) = \sum_{m \in \mathbb{Z}} T_m z^{-m}, \quad T_0 = I, \quad T_m = (0) \quad \text{for} \quad m \notin \{0, \dots, M\}.$$

Now

$$z^{1-\rho} T'(z) = -\sum_{m \in \mathbb{Z}} m T_m z^{-m-\rho} = -\sum_{m \in \mathbb{Z}} (m-\rho) T_{m-\rho} z^{-m},$$

and (5.96) gives

$$T(z)B(z) = A(z)T(z) - z^{1-\rho}T'(z).$$

Then, for $m \geq 0$,comparing the coefficients of z^{-m} delivers

$$\sum_{p=0}^{m} T_p B_{m-p} = \sum_{p=0}^{m} A_{m-p} T_p + (m-\rho) T_{m-\rho}. \tag{5.101}$$

In particular this forces $B_0 = A_0$, since $T_0 = I$, while $-\rho < 0$ and $T_{-\rho} = (0)$. Thus (5.101) may be written for $m \geq 1$ as

$$A_0 T_m - T_m A_0 = B_m + R_m, \tag{5.102}$$

where R_m involves only the matrices A_j, which are known, and the previously determined matrices $B_0, \dots, B_{m-1}$ and $T_0, \dots, T_{m-1}$. For $m > M$ the equation (5.102) is clearly satisfied by writing $T_m = (0)$ and $B_m = -R_m$.

Now write

$$T_m = (T_m^{jk}), \quad B_m = (B_m^{jk}), \quad R_m = (R_m^{jk}), \tag{5.103}$$

using the same block configuration as appears in A_0 and (5.98). Because A_0 is given by (5.97), the equation (5.102) now gives

$$M_j T_m^{jk} - T_m^{jk} M_k = B_m^{jk} + R_m^{jk}, \quad 1 \leq m \leq M. \tag{5.104}$$

Suppose first that $j \geq k$, so that the block lies on or below the diagonal. Then by case (i) of Lemma 5.2.8 there exists a matrix B_m^{jk}, with all columns zero except possibly the first, such that (5.104) has a solution T_m^{jk}.

Now take $j < k$, and thus a block lying strictly above the diagonal. Then Lemma 5.2.8 gives a matrix B_m^{jk}, with all rows zero except possibly the last, such that (5.104) has a solution T_m^{jk}. Thus the system $zy' = z^\rho B(z)y$ is normalised up to order M, as required. □

5.10.4 The effect of shearing

A *shearing* is given by writing $x = T(z)y$ in (5.95) and (5.96), where

$$T(z) = \begin{pmatrix} z^{n_1} & 0 & \dots & 0 \\ 0 & z^{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & z^{n_\nu} \end{pmatrix} = \text{diag}\,(z^{n_1}, \dots, z^{n_\nu}), \quad n_j \in \mathbb{Q}. \tag{5.105}$$

Here it is clear that

$$T(z)^{-1} = \text{diag}\,(z^{-n_1}, \dots, z^{-n_\nu}),$$

and each of T and T^{-1} is a matrix rational function in some possibly non-integer power of z.

Now premultiplying by a diagonal matrix has the effect of multiplying rows, while postmultiplying by a diagonal matrix multiplies columns. Hence (5.105) gives

$$z^{1-\rho}T(z)^{-1}T'(z) = z^{-\rho} \begin{pmatrix} n_1 & 0 & \dots & 0 \\ 0 & n_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & n_\mu \end{pmatrix} = z^{-\rho}\text{diag}\,(n_1, \dots, n_\nu). \tag{5.106}$$

We write

$$A(z) = \left(a^{jk}(z)\right), \quad B(z) = \left(b^{jk}(z)\right). \tag{5.107}$$

Here a^{jk} will denote *entries*, whereas A^{jk} will denote *blocks*. In passing from A to $T^{-1}AT$ the kth column is multiplied by z^{n_k}, and the jth row by z^{-n_j}. Combining these observations with (5.96), (5.106) and (5.107) gives

$$b^{jk}(z) = a^{jk}(z)z^{n_k - n_j} - \delta_{jk}z^{-\rho}n_j, \tag{5.108}$$

with δ_{jk} the Kronecker symbol.

5.10.5 A simple shearing

The following is a special case of the situation in §5.10.4. A *simple shearing* is given by writing $x = T(z)y$, where

$$T(z) = T_n(z) = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & \dots & z^n & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & z^n \end{pmatrix} = \text{diag}\,(1, \dots, 1, z^n, \dots, z^n), \quad n \in \mathbb{Z}. \tag{5.109}$$

Here the last q diagonal entries of T are z^n, and the rest are 1. The equation (5.93) transforms as in (5.95), with B as in (5.96). Thus (5.109) gives, in view of (5.106),

$$z^{1-\rho}T(z)^{-1}T'(z) = z^{-\rho}\text{diag}\,(0, \dots, 0, n, \dots, n).$$

Also, again since premultiplying by a diagonal matrix has the effect of multiplying rows, while postmultiplying multiplies columns, passing from A to $T^{-1}AT$ multiplies the last q columns by z^n and the last q rows by z^{-n}. The effect on the matrix A as we pass to the matrix B in (5.96) is as follows, in which the bottom right quadrant is $q \times q$:

$$\left(\begin{array}{c|c} \text{no change} & \text{multiply by } z^n \\ \hline \text{multiply by } z^{-n} & \text{no change, except subtract } nz^{-\rho} \text{ on the diagonal} \end{array}\right). \tag{5.110}$$

5.10.6 Systems in reduced form

The system (5.93) is called *reduced up to order* M if it is in standard nilpotent form and the following conditions hold:
(a) for $j \neq k$ the matrix A_m^{jk} in (5.99), corresponding to the block form (5.98), vanishes for $1 \leq m \leq M$ (and indeed for $m = 0$ also because of the form (5.97) of A_0);
(b) for $1 \leq m \leq M$ each diagonal block A_m^{jj} can only have non-zero entries in its first column (while $A_0^{jj} = M_j$ because of (5.97)).

Clearly this is a stronger condition than being normalised, and this section will describe how a simple shearing may be used to turn a normalised system into a reduced system.

In the next lemma we assume that, for some large positive integer M, the system (5.93) has been normalised up to order M as in §5.10.3. Recall that this means first that it is in standard nilpotent form, i.e. that $A_0 \neq (0)$ is a block matrix as in (5.97), where the M_j are upper triangular shifting matrices of dimensions s_j, with $s_1 \geq \ldots \geq s_\mu$, and that when the matrix A is written in the same block configuration (5.98) as A_0, each block has a (formal or asymptotic) series (5.99) with matrix coefficients A_m^{jk}, such that for $1 \leq m \leq M$ the following conditions are satisfied:
(i) for $j \geq k$ (i.e. on or below the diagonal) all non-zero entries of the matrix A_m^{jk} lie in the first column;
(ii) for $j < k$ (i.e. strictly above the diagonal) all non-zero entries of the matrix A_m^{jk} lie in the last row.

Lemma 5.10.3 *Suppose that with the assumptions on A of the previous paragraph there exists $m_1 \in \{1, \ldots, M\}$ and a block $A_{m_1}^{jk} \neq (0)$ with $j \neq k$. Then the system $zx' = z^\rho A(z)x$ may be transformed via a simple shearing (5.109) to a system $zy' = z^\rho B(z)y$ for which B_0 is nilpotent, but superior to A_0 in the sense of §5.2.2.*

Proof. Assume first that there exist $m_1 \in \{1, \ldots, M\}$ and a block $A_{m_1}^{jk} \neq (0)$ with $j > k$ (i.e. below the diagonal). Take the *least* such m_1. Then take the largest τ for which there exists $k < \tau$ with $A_{m_1}^{\tau k} \neq (0)$. Because the system is normalised, all non-zero entries in these $A_{m_1}^{\tau k}$ are in the first column. We apply the simple shearing (5.109) with $n = -m_1$ and $q = s_\tau + \ldots + s_\mu$. Thus the diagonal entries of T in line with $M_1, \ldots, M_{\tau-1}$ are all 1, while those aligned with $M_\tau, \ldots, M_\mu$ are z^{-m_1}.

We assert that, by (5.110), the shearing produces a new system $zy' = z^\rho B(z)y$, such that A_0 and B_0 are related by (5.4) (and so the rank is unchanged). To see this, note first that, because m_1 and ρ are positive, all blocks of B_0 are the same as those of A_0, except for blocks corresponding to the bottom left quadrant of (5.110). Entries of A corresponding to this bottom left quadrant of (5.110) are those in the blocks A^{rs}, $r \geq \tau > s$, of the block form (5.98), and under the shearing they are multiplied by z^{m_1}. However, because m_1 is minimal, B still has a series in non-positive powers of z. Furthermore, the C_k in (5.4) are the matrices $A_{m_1}^{\tau k}$, for $1 \leq k < \tau$, all of which are such that all columns are zero bar the first, and at least one of which is not the zero matrix. Moreover, the maximality of τ ensures that the blocks of B_0 lying below these C_k are zero. Thus (5.4) holds with M the matrix diag $(M_{\tau+1}, \ldots, M_\mu)$ (and M not present if $\tau = \mu$). Finally, Lemma 5.2.6 implies that B_0 is nilpotent, but superior to A_0.

Now suppose that there exist $m_1 \in \{1, \ldots, M\}$ and a block $A^{jk}_{m_1} \neq (0)$ with $j < k$ (i.e. above the diagonal). Again take the *least* such m_1. Then take the largest τ for which there exists $j < \tau$ with $A^{j\tau}_{m_1} \neq (0)$. Because the system is normalised, all non-zero entries in these $A^{j\tau}_{m_1}$ are in the last row. Apply the simple shearing (5.109) with $q = s_\tau + \ldots + s_\mu$ as before, but taking $n = m_1$. This time the diagonal entries of T in line with $M_1, \ldots, M_{\tau-1}$ are all 1, while those aligned with $M_\tau, \ldots, M_\mu$ are z^{m_1}. The effect of the shearing is to make all blocks of B_0 the same as those of A_0, except for blocks corresponding to the upper right quadrant of (5.110), corresponding to the blocks A^{rs}, $r < \tau \leq s$, of (5.98). These blocks are multiplied by z^{m_1} but the minimality of m_1 again ensures that B has a series in non-positive powers of z. Moreover, the matrices A_0, B_0 satisfy (5.6), in which the D_j are the $A^{j\tau}_{m_1}$ with $1 \leq j < \tau$, and all blocks to the right of them are zero, by the maximality of τ. This time, Lemma 5.2.7 may be applied. □

The new system $zy' = z^\rho B(z)y$ may not be normalised, and B_0 may not be a direct sum of shifting matrices in standard nilpotent form as in (5.97). However, B_0 is similar to a matrix $C_0 = U^{-1}B_0U$ of form (5.97), with blocks of non-increasing dimension, but not necessarily with the same μ or s_j. Moreover, C_0 is still superior to A_0, because C_0^l and B_0^l have the same rank for each l. Here U is a constant matrix, and writing $y = Uv$ gives $zv' = z^\rho C(z)v$, where $C(z) = U^{-1}B(z)U$. This new system may be normalised up to order M using Lemma 5.10.2, which does not affect the lead matrix C_0. Lemma 5.10.3 may then be applied again but, as remarked in §5.2.2, it is not possible to produce superior matrices via this method an arbitrarily large number of times, because all of the matrices involved are nilpotent and so satisfy $N^\nu = (0)$. So eventually this must lead to a normalised system with a coefficient matrix D such that D_0 is in standard nilpotent form and $D^{jk}_m = (0)$ for all $j \neq k$ and $1 \leq m \leq M$. Thus the following lemma has been proved.

Lemma 5.10.4 *Let M be a large positive integer. Every system (5.93) which is normalised up to order M is equivalent via a transformation $x = H(z)y$, where $H(z)$ is a finite product of non-singular constant matrices, matrices $T(z)$ as in Lemma 5.10.2, and simple shearings as in Lemma 5.10.3, to a system which has the same rank ρ and is reduced up to order M.*

□

Combining Lemmas 5.10.2 and 5.10.4 then gives the following.

Lemma 5.10.5 *Let the system (5.93) satisfy (5.94) and be in standard nilpotent form. Let M be a large positive integer. Then (5.93) is equivalent via a transformation $x = H(z)y$, where $H(z)$ is a finite product of non-singular constant matrices, matrices $T(z)$ as in Lemma 5.10.2, and simple shearings as in Lemma 5.10.3, to a system which has the same rank ρ and is reduced up to order M.*

□

Note that the transformations $T(z)$ and simple shearings applied in Lemma 5.10.5 only involve integer powers of z.

5.10.7 Application of a block shearing to a reduced system

This section describes how shearing may be used so that either the rank ρ of the system is reduced, or a new system is generated, possibly with larger rank, but such that the lead matrix has at least two distinct eigenvalues, so that by §5.9 the equation can be split into two of lower order.

We assume as before that in the system (5.93) the matrix $A_0 \neq (0)$ has the block form (5.97), where the M_j are upper triangular shifting matrices of dimensions s_j, with $s_1 \geq \ldots \geq s_\mu$. The matrix

A is then written in the same block configuration (5.98) as A_0. Each of these blocks then has a (formal or asymptotic) series (5.99) with matrix coefficients A_m^{jk}, and by Lemma 5.10.5 we may assume that the system has been reduced up to some large order M. This means that for $j \neq k$ the matrix A_m^{jk} vanishes for $1 \leq m \leq M$ (and indeed for $m = 0$ also because of the form of A_0). Moreover, for $1 \leq m \leq M$ the diagonal block A_m^{jj} only has non-zero entries in its first column (while $A_0^{jj} = M_j$). Let

$$U = \{p/q : \, p, q \in \mathbb{N}, \, (p, q) = 1, \, 1 \leq p \leq q \leq s_1\}. \tag{5.111}$$

We will apply a block shearing given, for some $p/q \in U$, by

$$T = \begin{pmatrix} T_1 & 0 & \dots & 0 \\ 0 & T_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & T_\mu \end{pmatrix}, \quad T_j = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & z^{-p/q} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & z^{-(s_j-1)p/q} \end{pmatrix}. \tag{5.112}$$

Here T has the same block form as A_0, and it is not assumed that $\mu \geq 2$. In the case of a holomorphic coefficient matrix A on a sector S, we take an arbitrary branch of $z^{1/q}$. The entries of the original and new coefficient matrices A and B are then related by (5.107) and (5.108).

Consider first those j, k for which the entry in row j, column k of A_0 is 1. For these j, k we have $k = j + 1$ and, because these entries do not lie in the first column of any diagonal block,

$$a^{jk}(z) = 1 + O(|z|^{-M-1}),$$

either in the sense of formal series or, in the case of holomorphic coefficients, as $z \to \infty$ in the sector S. Moreover, with the notation (5.105), we have $n_k - n_j = -p/q$, by (5.112), and hence (5.108) yields

$$b^{jk}(z) = a^{jk}(z) z^{-p/q} = z^{-p/q} + O(|z|^{-p/q-M-1}). \tag{5.113}$$

Next, take any pair j, k for which the entry in row j, column k of A_0 is 0 and does not lie in the first column of any diagonal block. Then (5.108) and the fact that the system $zx' = z^\rho A(z)x$ is reduced give, since M is large and $p/q \leq 1 \leq \rho$,

$$a^{jk}(z) = O(|z|^{-M-1}), \quad |b^{jk}(z)| \leq O(|z|^{\nu p/q-M-1}) + O(|z|^{-\rho}) = O(|z|^{-p/q}). \tag{5.114}$$

We say that $p/q \in S$ is *admissible* if

$$b^{jk}(z) = O(|z|^{-p/q}) \quad \text{for} \quad 1 \leq j \leq \nu, \, 1 \leq k \leq \nu. \tag{5.115}$$

In view of (5.113) and (5.114), it is enough to check this holds for those coefficients $a^{jk}(z)$ arising from the first column of a diagonal block A^{rr}, $1 \leq r \leq \mu$. We label the entries of this column in descending order as

$$\alpha_{\gamma,r}(z), \quad 1 \leq \gamma \leq t = s_r \leq s_1.$$

For such an entry, reading along the corresponding row and up the corresponding column of T in (5.112) is equivalent to reading along row γ and up the first column of T_r; this shows that, with the terminology (5.105), the integers n_j and n_k are $-(\gamma - 1)p/q$ and 0 respectively. Thus, if we suppress the subscript r, (5.108) implies that the shearing replaces $\alpha_\gamma(z)$ by

$$\beta_\gamma(z) = \alpha_\gamma(z) z^{(\gamma-1)p/q}, \tag{5.116}$$

and so admissibility of p/q is equivalent to

$$\alpha_\gamma(z) z^{\gamma p/q} = O(1), \tag{5.117}$$

for every choice of γ (and r). The form of A_0 implies that we always have $\alpha_\gamma(z) = O(|z|^{-1})$, and since $\gamma \le s_1$ it follows that $p/q = 1/s_1$ is admissible.

Suppose that we apply this shearing with $p/q = 1$ and that 1 turns out to be admissible. Then the n_j are all integers in (5.105), and (5.108) and (5.115) show that the shearing has transformed the equation $zx' = z^\rho A(z)x$ into $zy' = z^\rho B(z)y$, with $B(z)$ given by a (formal or asymptotic) series in descending integer powers of z, such that $B(z) = O(|z|^{-1})$ as $z \to \infty$. In this case the rank of the equation has been reduced.

Assume now that 1 is not admissible, and let p/q be the *largest* admissible member of U.

Lemma 5.10.6 *There exists at least one term $\alpha_{\gamma,r}(z)$ for which $\lim_{z\to\infty, z\in S} \alpha_{\gamma,r}(z) z^{\gamma p/q}$ exists and is finite but non-zero.*

Proof. Let $U_1 = \{p'/q' \in U : p'/q' > p/q\}$. Then $1 \in U_1$. Let p''/q'' be the nearest member of U_1 to p/q. Then p''/q'' is not admissible and so by (5.117) there exists at least one γ (with a corresponding r) such that

$$\alpha_\gamma(z) z^{\gamma p''/q''} \to \infty;$$

this means as $z \to \infty$ in S or, in the formal solutions setting, that the series contains positive powers of z. It follows that, in the same sense,

$$\alpha_\gamma(z) z^{\gamma p'/q'} \to \infty \quad \text{for every } p'/q' \in U_1. \tag{5.118}$$

Fix this choice of γ (and r). Because $\alpha_\gamma(z)$ lies in the first column of a diagonal block of A, and because A_0 satisfies (5.97), there exist $c_\gamma \in \mathbb{C} \setminus \{0\}$ and $m_\gamma \in \mathbb{N}$ such that

$$\alpha_\gamma(z) \sim c_\gamma z^{-m_\gamma},$$

and since $1 \in U_1$ it must be the case that $\gamma > m_\gamma$.

Because $1 \le m_\gamma < \gamma \le s_r \le s_1$, we have $v = m_\gamma/\gamma \in U$. But

$$\alpha_\gamma(z) z^{\gamma v} = \alpha_\gamma(z) z^{m_\gamma} \to c_\gamma, \tag{5.119}$$

and so (5.118) implies that $v \notin U_1$, which implies that $p/q \ge v$. If $p/q > v$ then we have

$$\alpha_\gamma(z) z^{\gamma p/q} \to \infty,$$

by (5.119), contradicting (5.117). It follows that $p/q = v$, and (5.119) now proves the lemma. □

Still assuming that $p/q < 1$ is the maximal admissible member of U, recall that the original equation (5.93) was

$$zx' = \widehat{A}(z)x, \quad \widehat{A}(z) = z^\rho A(z), \quad \rho \ge 1, \quad A(z) \sim \widetilde{A}(z) = \sum_{m=0}^{\infty} A_m z^{-m},$$

with $\widetilde{A}(z)$ being either a formal series or an asymptotic series valid as $z \to \infty$ in a sector S. The transformed equation has the form

$$\begin{aligned}
zy' &= \widehat{B}(z)y, \\
\widehat{B}(z) &= z^\rho B(z) = T(z)^{-1}\widehat{A}(z)T(z) - zT(z)^{-1}T'(z), \\
B(z) &= T(z)^{-1}A(z)T(z) - z^{1-\rho}T(z)^{-1}T'(z),
\end{aligned}$$

as in (5.95) and (5.96), The fact that p/q is admissible implies by (5.115) that $B(z) = O(|z|^{-p/q})$. It follows using (5.108) and (5.112) that B has a (formal or asymptotic) series $\widetilde{B}(z)$ in powers of $z^{1/q}$ given by

$$B(z) \sim \widetilde{B}(z) = \sum_{m=p}^{\infty} B_m z^{-m/q}, \tag{5.120}$$

and the lead matrix B_p has the following properties. B_p can be written as a block matrix D_{jk} with blocks of the same dimensions as those of A in (5.98), and all blocks D_{jk} with $j \neq k$ vanish: this is by (5.114) and the fact that M is large and $p/q < 1 \leq \rho$. Moreover, by (5.113) and (5.114) the diagonal blocks D_{jj} are each of the form $D_{jj} = M_j + C_j$, where M_j is the same upper triangular shifting matrix as in A_0, and all entries not in the first column of C_j are 0. By the maximality of p/q, (5.116) and Lemma 5.10.6 show that at least one matrix C_j is non-zero.

Lemma 5.10.7 *The matrix B_p is not nilpotent.*

Proof. Assume that B_p is nilpotent. Because all non-diagonal blocks of B_p are (0), each of the blocks $D_{jj} = M_j + C_j$ must be nilpotent. But then Lemma 5.1.2 shows that C_j must vanish, which is false for at least one j. □

Lemma 5.10.8 *The matrix B_p has at least two distinct eigenvalues.*

Proof. Each of $T(z)^{-1}$, $T(z)$ and $T'(z)$ is a polynomial in $z^{1/q}$ or $z^{-1/q}$, and so the (formal or asymptotic) series satisfy

$$\widetilde{B}(z) = T(z)^{-1}\widetilde{A}(z)T(z) - z^{1-\rho}T(z)^{-1}T'(z) = T(z)^{-1}\widetilde{A}(z)T(z) - z^{-\rho}E,$$

where E is a constant diagonal matrix by (5.106). Let $T(ze^{2\pi i})$ denote the matrix resulting from formally replacing $z^{p/q}$ in $T(z)$ with $z^{p/q}e^{2\pi ip/q}$, with a similar convention for the other matrices. Thus (5.112) shows that $T(ze^{2\pi i}) = T(z)D$, where D is the diagonal matrix with entries $1, e^{-2\pi ip/q}, \ldots, e^{-(s_\mu - 1)2\pi ip/q}$. This yields

$$\begin{aligned}
\widetilde{B}(ze^{2\pi i}) &= T(ze^{2\pi i})^{-1}\widetilde{A}(ze^{2\pi i})T(ze^{2\pi i}) - (ze^{2\pi i})^{-\rho}E \\
&= T(ze^{2\pi i})^{-1}\widetilde{A}(z)T(ze^{2\pi i}) - z^{-\rho}E \\
&= D^{-1}T(z)^{-1}\widetilde{A}(z)T(z)D - z^{-\rho}E = D^{-1}\widetilde{B}(z)D.
\end{aligned}$$

It now follows from (5.120) that

$$B_p z^{-p/q}e^{-2\pi ip/q} = D^{-1}B_p z^{-p/q}D, \quad B_p = e^{2\pi ip/q}D^{-1}B_pD,$$

and so, for $\lambda \in \mathbb{C}$,

$$\begin{aligned}
\det(B_p - \lambda I) &= \det(e^{2\pi ip/q}D^{-1}B_pD - \lambda D^{-1}D) \\
&= \det(e^{2\pi ip/q}D^{-1}(B_p - e^{-2\pi ip/q}\lambda I)D) \\
&= e^{2\pi i\nu p/q}\det(B_p - e^{-2\pi ip/q}\lambda I).
\end{aligned}$$

Since B_p is not nilpotent, its characteristic equation has at least one non-zero root λ, and $e^{-2\pi ip/q}\lambda$ is another eigenvalue of B_p. □

The equation $zy' = z^\rho B(z)y$ may now be written in the form

$$z\frac{dy}{dz} = z^{\rho - p/q}F(z), \quad F(z) = z^{p/q}B(z) \sim \sum_{m=0}^{\infty} F_m z^{-m/q}, \quad F_0 = B_p.$$

Setting $w = z^{1/q}$, $z = w^q$ and $Y(w) = y(z) = y(w^q)$ gives

$$w\frac{dY}{dw} = qz\frac{dy}{dz} = qw^{q\rho-p}G(w), \quad G(w) = F(w^q) \sim \sum_{m=0}^{\infty} F_m w^{-m},$$

in which the lead matrix $F_0 = B_p$ has at least two distinct eigenvalues. This proves the following.

Theorem 5.10.1 *Every system (5.93) satisfying (5.94), in which A_0 is nilpotent, is equivalent via a transformation $x = H(z)y$, in which $H(z)$ is a finite product of non-singular constant matrices, matrices as in (5.100) and shearing matrices as in (5.105), to a system $zy' = \widehat{B}(z)y$, such that at least one of the following holds.*
(i) Both $H(z)$ and the (formal or asymptotic) series for $\widehat{B}(z)$ involve only integer powers of z, and the new system $zy' = \widehat{B}(z)y$ has rank less than ρ.
(ii) There exists $q \in \mathbb{N}$ such that writing $z = w^q$ and $Y(w) = y(z) = y(w^q)$ transforms the system $zy' = \widehat{B}(z)y$ to a system $wY' = \widehat{C}(w)Y$, where C has a formal series, or asymptotic series in an appropriate sector, in descending integer powers of w, in which the lead matrix C_0 has at least two distinct eigenvalues.

In the case of holomorphic coefficients and solutions the branch $w = z^{1/q}$ may be chosen arbitrarily. Obviously case (i) applies if A_0 is the zero matrix, because a power of z may be cancelled.

5.11 The main theorem on asymptotic integration

In the following theorem and proof a sector $S = S(R, \alpha, \beta)$ will be said to have opening $\beta - \alpha$, and $S' = S(R, \alpha', \beta')$ will be called a *proper* subsector of S if $\alpha < \alpha' < \beta' < \beta$.

Theorem 5.11.1 *Let $\rho \in \mathbb{Z}$ and let $A_0, A_1, \ldots$ be $\nu \times \nu$ constant matrices, with A_0 not the zero matrix. Then there exists $p \in \mathbb{N}$ such that the formal differential equation (5.35) has a formal solution*

$$x(z) = V(z)z^G e^{Q(z)}, \quad V(z) = \sum_{m=0}^{\infty} V_m z^{-m/p}, \tag{5.121}$$

satisfying the following:
(i) the V_m are $\nu \times \nu$ constant matrices and $\det V(z)$ is not the zero series;
(ii) G is a constant Jordan matrix of form $G = G_1 \oplus \ldots \oplus G_s$, where G_j is a $\mu_j \times \mu_j$ Jordan block and $\sum_{j=1}^{s} \mu_s = \nu$;
(iii) $Q(z)$ is a diagonal matrix of form

$$Q(z) = Q_1(z)I_{\mu_1} \oplus \ldots \oplus Q_s(z)I_{\mu_s},$$

with each $Q_j(z)$ a polynomial in $z^{1/p}$.

Furthermore, let $A(z)$ be a $\nu \times \nu$ matrix function which is holomorphic for all z in a sector

$$S = S(R, \alpha, \beta) = \{z \in \mathbb{C} : |z| > R, -\infty < \alpha < \arg z < \beta < +\infty\}$$

on the Riemann surface of $\log z$. Assume that $A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m}$ as $z \to \infty$ in S, in the sense of asymptotic series. Then for each $\theta \in (\alpha, \beta)$ there exists $r(\theta) > 0$ with the property that the equation (5.34) has a non-singular holomorphic matrix solution

$$x(z) = W(z)z^G e^{Q(z)} \tag{5.122}$$

in $S_\theta = S(R, \theta - r(\theta), \theta + r(\theta))$, such that $V(z)$ is an asymptotic series for $W(z)$ on S_θ, for some branch of $z^{1/p}$.

Proof. The non-singular nature of $x(z)$ follows from the fact that $\det V(z)$ is an asymptotic series for $\det W(z)$. The theorem is true if $\rho \leq 0$, by Theorem 5.7.2, and if $\nu = 1$, by Theorem 5.8.1; in both of these cases we have $p = 1$ and $W(z) \sim U(z)$ as $z \to \infty$ in the whole sector S. Assume that the theorem is false, and take the least $\nu \in \mathbb{N}$ having at least one pair $\{\nu, \rho\}$, with $\rho \geq 1$, for which the assertion of the theorem fails: then $\nu > 1$.

Claim: *with this value of ν, all assertions of the theorem hold if the lead matrix A_0 has at least two distinct eigenvalues.*

To prove this claim, observe first that the eigenvalues of A_0 give rise to μ as in the statement of Theorem 5.9.3, which then delivers a formal transformation $x = \sum_{m=0}^{\infty} P_m z^{-m} y$ sending (5.35) to (5.47), where P_0 is non-singular and B_0 is similar to A_0, while B_m^{11} is $\mu \times \mu$ and B_m^{22} is $(\nu - \mu) \times (\nu - \mu)$.

Next, suppose that $A(z)$ is holomorphic on the sector $S = S(R, \alpha, \beta)$, with $A(z) \sim \sum_{m=0}^{\infty} A_m z^{-m}$ as $z \to \infty$ in S, and take $\theta \in (\alpha, \beta)$. It may be assumed without loss of generality that $\beta - \alpha < \pi/\rho$. Now Theorem 5.9.3 gives a holomorphic matrix function $P(z)$ on a proper subsector $S' = S(R, \alpha', \beta')$ of S, with $\alpha' < \theta < \beta'$, such that writing $x = P(z)y$ transforms the equation (5.34) to (5.57), where B^{11} is $\mu \times \mu$ and B^{22} is $(\nu - \mu) \times (\nu - \mu)$, and P and B have asymptotic series (5.58) in S', in which P_0 is non-singular and B_0 is similar to A_0.

Since the theorem holds whenever the dimension of the equation is less than ν, the equations

$$z^{1-\rho} w' = \sum_{m=0}^{\infty} B_m^{jj} z^{-m} w, \quad j = 1, 2,$$

have formal solutions $x_j(z) = V_j(z) z^{G_j} e^{Q_j(z)}$ respectively, each of these satisfying conclusions (i) to (iii) of the theorem, for some $p = p_j \in \mathbb{N}$. By taking the least common multiple of p_1 and p_2, it may be assumed that $p_1 = p_2 = p$. Then

$$\left(\sum_{m=0}^{\infty} P_m z^{-m} \right) (V_1(z) \oplus V_2(z)) z^{G_1 \oplus G_2} e^{Q_1(z) \oplus Q_2(z)},$$

is the required formal solution of (5.37). Moreover, there exists $r(\theta) > 0$ such that the equations

$$z^{1-\rho} w' = B^{jj}(z) w, \quad j = 1, 2,$$

have holomorphic solutions $x_j(z) = W_j(z) z^{G_j} e^{Q_j(z)}$ respectively on $S_\theta = S(R, \theta - r(\theta), \theta + r(\theta))$, with $W_j(z) \sim V_j(z)$ there. Hence

$$P(z)(W_1(z) \oplus W_2(z)) z^{G_1 \oplus G_2} e^{Q_1(z) \oplus Q_2(z)},$$

is the required holomorphic solution of (5.34), since

$$P(z)(W_1(z) \oplus W_2(z)) \sim \left(\sum_{m=0}^{\infty} P_m z^{-m} \right) (V_1(z) \oplus V_2(z)),$$

using (5.58). This proves the claim.

With ν minimal as above, take the least integer ρ for which the theorem fails: then $\rho \geq 1$ by Theorem 5.7.2. With this choice of pair $\{\nu, q\}$, and the remaining hypotheses of the theorem, it may be assumed, by the claim above, that A_0 has just one eigenvalue λ.

Suppose first that this unique eigenvalue λ of A_0 satisfies $\lambda \neq 0$. Write

$$x = y \exp(\lambda z^\rho / \rho),$$

which gives

$$A(z)y\exp(\lambda z^\rho/\rho) = z^{1-\rho}x' = z^{1-\rho}y'\exp(\lambda z^\rho/\rho) + \lambda y\exp(\lambda z^\rho/\rho),$$

and transforms the equation (5.34) to

$$z^{1-\rho}y' = C(z)y, \quad C(z) = A(z) - \lambda I_\nu, \tag{5.123}$$

and its formal counterpart (5.37) similarly. Here the (formal or asymptotic) series expansion is

$$C(z) = A(z) - \lambda I_\nu \sim A_0 - \lambda I_\nu + \sum_{m=1}^{\infty} A_m z^{-m}.$$

If $C_0 = A_0 - \lambda I_\nu$ is the zero matrix then a power of z may be cancelled from the equation (5.123), so that ρ is reduced and the conclusion of the theorem holds for (5.123) and hence also for (5.34). If $C_0 = A_0 - \lambda I_\nu \neq (0)$ then C_0 is nilpotent, because λ is the only eigenvalue of A_0.

Thus it may be assumed henceforth that A_0 is nilpotent, but not the zero matrix. Now Theorem 5.10.1 delivers an invertible matrix $H(z)$, a finite product of non-singular constant matrices, matrices as in (5.100) and shearings as in (5.105), such that writing $x = H(z)y$ gives an equation $z^{1-\rho}y' = B(z)y$ and its formal counterpart $z^{1-\rho}y' = \widetilde{B}(z)y$. If it can be shown that all assertions of the theorem hold for the transformed equations, then premultiplying formal and holomorphic solutions by $H(z)$ gives all conclusions of the theorem for (5.34) and (5.37).

By Theorem 5.10.1 again, there are two possibilities for the equation $z^{1-\rho}y' = B(z)y$ and its formal counterpart. The first is that the new equations have rank $\rho' < \rho$, and both $H(z)$ and $\widetilde{B}(z)$ involve only integer powers of z. In this case all assertions of the theorem hold by the minimality of ρ.

The remaining possibility afforded by Theorem 5.10.1 is that there exists $s \in \mathbb{N}$ such that writing $z = w^s$ and $Y(w) = y(w^s)$ transforms the formal equation $z^{1-\rho}y' = \widetilde{B}(z)y$ into an equation

$$w^{1-\rho'}Y'(w) = \left(\sum_{m=0}^{\infty} C_m w^{-m}\right) Y(w),$$

in which the lead coefficient matrix C_0 has at least two eigenvalues. This equation then has, by the claim, a formal solution $U(w)w^F e^{P(w)}$ satisfying conclusions (i) to (iii) of the theorem, with $U(w)$ a formal series and $P(w)$ a polynomial matrix, both in powers of $w^{1/t}$ for some $t \in \mathbb{N}$. Hence $U(z^{1/s})z^{F/s}e^{P(z^{1/s})}$ is the required formal solution for $z^{1-\rho}y' = \widetilde{B}(z)y$, involving powers of $z^{1/st}$.

Moreover, with some arbitrary choice of holomorphic branch of $w = z^{1/s}$, the same change of variables transforms $z^{1-\rho}y' = B(z)y$ on a sector S^* of small opening to $w^{1-\rho'}Y'(w) = C(w)Y(w)$ on some sector S^{**}, with $C(w) \sim \sum_{m=0}^{\infty} C_m w^{-m}$ on S^{**}. This new equation has a holomorphic solution $U(w)w^F e^{P(w)}$ with $V(w) \sim U(w)$ on S^{**}, and so there exists a holomorphic solution $V(z^{1/s})z^{F/s}e^{P(z^{1/s})}$ of $z^{1-\rho}y' = B(z)y$ with $V(z^{1/s}) \sim U(z^{1/s})$ on S^*.

$\square$

Remark. If the eigenvalues of A_0 are pairwise distinct, the remark following Theorem 5.9.3 shows that we may take $p = 1$ in (5.121), since application of Theorem 5.9.3 does not introduce fractional powers of z.

5.11.1 Changing the branch of $z^{1/p}$

For an arbitrary choice of the branch $z^{1/2}$, the matrix function

$$x(z) = \begin{pmatrix} \exp\left(z^{1/2}\right) & \exp\left(-z^{1/2}\right) \\ 2^{-1}z^{-1/2}\exp\left(z^{1/2}\right) & -2^{-1}z^{-1/2}\exp\left(-z^{1/2}\right) \end{pmatrix}$$

is a locally holomorphic solution of

$$x' = B(z)x, \quad B(z) = \begin{pmatrix} 0 & 1 \\ 1/4z & -1/2z \end{pmatrix}.$$

This is easy to verify since, for $c = \pm 1$ and $f_c(z) = \exp(cz^{1/2})$, we have

$$f_c''(z) + \frac{1}{2z} \cdot f_c'(z) - \frac{1}{4z} \cdot f_c(z) = 0.$$

Changing the branch of $z^{1/2}$ interchanges the exponential parts in $x(z)$, and is equivalent to multiplying $x(z)$ on the right by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Suppose more generally that we have a formal solution $x(z) = Y(z^{1/p})z^F e^{P(z^{1/p})}$ of the matrix differential equation $zx' = A(z)x$ as in Theorem 5.11.1, and write

$$z = u^p, \quad x(z) = V(u) = Y(u)u^{pF}e^{P(u)}, \quad zx'(z) = zV'(u)(1/p)z^{1/p-1} = (1/p)uV'(u).$$

Thus V satisfies $uV'(u) = pA(u^p)V(u)$. Now let $c^p = 1$ and write $y(z) = V(cu)$, so that $y(z)$ is simply $x(z)$ with each occurrence of $z^{1/p}$ in the formal series Y and the polynomial P replaced by $cz^{1/p}$ (which is of course another branch of the p'th root of z). Then y satisfies

$$zy'(z) = zV'(cu)c(1/p)z^{1/p-1} = (1/p)(cu)V'(cu) = A((cu)^p)V(cu) = A(z)y(z).$$

Hence y solves the same equation as x, and so by Lemma 5.4.4 the exponential parts for y must be a permutation of those of x.

5.12 The case of scalar equations

Consider an nth order formal differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0 y = 0, \tag{5.124}$$

in which the coefficients a_j are formal series in descending integer powers of z (this phrase being used as in §5.3 to mean that each includes at most finitely many positive powers). Then by *canonical formal solutions* of (5.124) we mean expressions of the form, for some $p \in \mathbb{N}$,

$$f_j(z) = \exp(P_j(z^{1/p}))z^{\lambda_j} \sum_{m=0}^{n_j} (\log z)^m U_{j,m}(z) \tag{5.125}$$

which satisfy the equation (5.124) after formal differentiation and substitution, and for which the following conditions hold: the *exponential parts* $q_j(z) = P_j(z^{1/p})$ are polynomials in $z^{1/p}$; each λ_j is a complex number, while each n_j is a non-negative integer; $U_{j,m}(z)$ is a formal series in descending integer powers of $z^{1/p}$; the leading coefficient U_{j,n_j} is not the zero series. It is evident that solutions f_j given by (5.125) may always be normalised so that

$$\operatorname{Re} \lambda_j \in [0, 1/p). \tag{5.126}$$

Theorem 5.12.1 *Assume that the coefficients a_j in the formal differential equation (5.124) are formal series in descending integer powers of z. Then there exists $p \in \mathbb{N}$ with the property that (5.124) has a fundamental set of n linearly independent canonical formal solutions f_j satisfying (5.125). Moreover, these f_j have the property that*

$$\textit{if } 0 \leq m < n_j \textit{ then there exists } j' \neq j \textit{ with } (q_{j'}, \lambda_{j'}, n_{j'}) = (q_j, \lambda_j, m). \tag{5.127}$$

Proof. For any formal solution f of (5.124), the column vector $X = (f, f', \ldots, f^{(n-1)})^T$ is a vector solution of

$$x'(z) = A(z)x(z), \quad A(z) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-1} \end{pmatrix}. \tag{5.128}$$

The coefficients of $A(z)$ are formal series in descending integer powers of z. Thus Theorem 5.11.1 shows that there exists $p \in \mathbb{N}$ such that (5.128) has a principal formal matrix solution

$$x(z) = V(z) z^G e^{Q(z)}, \tag{5.129}$$

where $V(z)$ is an invertible matrix whose entries are formal series in descending integer powers of $z^{1/p}$, while $Q(z)$ is a diagonal matrix whose entries are polynomials in $z^{1/p}$, and G is a constant Jordan matrix which commutes with $Q(z)$. The $f_j(z)$ are then simply the entries from the first row of $x(z)$, and it follows from (5.128) that these satisfy (5.124) (and their eigenvalues may be normalised as in Lemma 5.4.3 so that the f_j satisfy (5.126)). Moreover, (5.128) shows that for $j = 1, \ldots, n-1$, the jth row of $x'(z)$ is the $(j+1)$th row of $x(z)$, and so each row of $x(z)$ is the derivative of the row above it. Furthermore, the f_j are linearly independent because otherwise $\det x(z)$ would vanish.

To see that (5.127) holds, assume that $0 \le m < n_j$. Since $f_j(z)$ is the jth entry in the first row of $x(z)$, the jth column of the block matrix z^G must contain a constant multiple of $z^{\lambda_j}(\log z)^{n_j}$ lying in some column of some block H of z^G, this block arising from a Jordan block of G with eigenvalue λ_j. By (5.129), this block H must have a (different) column in which the highest power of $\log z$ which occurs is $(\log z)^m$, this power occurring only once there. Hence there exists a column of z^G, say the kth, which contains a constant multiple of $z^{\lambda_j}(\log z)^m$, and for which all other entries are constant multiples of $z^{\lambda_j}(\log z)^{m'}$ with $m' < m$. Evidently we have $j \neq k$, but $P_j(z^{1/p}) = P_k(z^{1/p})$ by Theorem 5.11.1(ii), (iii). Now, since the kth column of $V(z)$ is not zero, the kth column of $V(z)z^G e^{Q(z)}$ has an entry which includes a non-trivial series in $z^{1/p}$ multiplied by $\exp(P_j(z^{1/p}))z^{\lambda_j}(\log z)^m$. On the other hand, no higher power of $\log z$ occurs in this kth column of $V(z)z^G e^{Q(z)}$. Since each entry of $V(z)z^G e^{Q(z)}$ is the derivative of that lying above it, the powers of $\log z$ which occur in these entries cannot increase as we follow the column downwards, and we must have a term involving $\exp(P_j(z^{1/p}))z^{\lambda_j}(\log z)^m$ in the first entry. This gives (5.127) with $j' = k$.

□

The solutions arising from Theorem 5.12.1 will be called *principal formal solutions* of (5.124), and an *admissible formal solution* of (5.124) is defined to be a linear combination over $\mathbb{C}$ of finitely many canonical formal solutions.

Lemma 5.12.1 *Assume that the the coefficients a_j in the formal differential equation (5.124) are formal series in descending integer powers of z. Then any admissible formal solution of (5.124) is a linear combination over $\mathbb{C}$ of the principal formal solutions given by Theorem 5.12.1, and any $n+1$ admissible formal solutions are linearly dependent.*

Proof. Let $f_1, \ldots, f_n$ be the principal formal solutions, and let g be any canonical formal solution. It suffices to prove that g is a linear combination of $f_1, \ldots, f_n$. Set $g_1 = g$ and $g_j = f_j$ for $j \ge 2$, and let $y(z)$ be the matrix whose jth column is $g_j(z), g_j'(z), \ldots, g_j^{(n-1)}(z)$. Then $y(z)$ is a basic formal matrix solution of (5.128) as in Definition 5.4.2. Thus Lemma 5.4.4 shows that $y(z) = X(z)C$, where C is a constant matrix and $X(z)$ is the principal formal matrix solution whose first row consists of $f_1(z), \ldots, f_n(z)$, and so g is a linear combination of $f_1, \ldots, f_n$ as required. □

Lemma 5.12.2 *Suppose that we have canonical formal solutions f_j as in (5.125) and (5.126), with the property that if $j \neq j'$ then $q_j \neq q_{j'}$ or $\lambda_j \neq \lambda_{j'}$. Then the f_j are linearly independent over $\mathbb{C}$.*

Proof. Let $g_1, \ldots, g_n$ be the principal formal solutions. Then each f_j is a linear combination of the g_k, and the hypotheses imply that the same g_k does not appear in the representation for two distinct f_j. Since the g_k are linearly independent, so are the f_j.

Alternatively, if a linear combination $\sum c_j f_j$, $c_j \in \mathbb{C}$, reduces to 0 then Lemma 5.3.2 forces $c_j U_{j,m} = 0$ for each j. □

The formal Wronskian of solutions $g_1, \ldots, g_n$ of (5.124) is defined as

$$W(g_1, \ldots, g_n) = \begin{vmatrix} g_1 & \ldots & g_n \\ g_1' & \ldots & g_n' \\ \vdots & & \\ g_1^{(n-1)} & \ldots & g_n^{(n-1)} \end{vmatrix}$$

and Leibnitz' rule gives Abel's identity $W' = -a_{k-1} W$. Here the principal solutions f_j given by (5.125) and (5.129) satisfy $W(z) = W(f_1, \ldots, f_n)(z) = \det V(z) z^{\operatorname{tr} G} e^{\operatorname{tr} Q(z)}$, by Lemma 5.2.5. Hence $W(z)$ has exponential part $P(z) = \sum_{j=1}^{n} P_j(z^{1/p})$.

Moreover, given any n admissible formal solutions $g_1, \ldots, g_n$, each g_j is a linear combination of the principal solutions $f_1, \ldots, f_n$. Hence there exists a constant matrix C such that

$$\begin{pmatrix} g_1 & \ldots & g_n \\ g_1' & \ldots & g_n' \\ \vdots & & \\ g_1^{(n-1)} & \ldots & g_n^{(n-1)} \end{pmatrix} = \begin{pmatrix} f_1 & \ldots & f_n \\ f_1' & \ldots & f_n' \\ \vdots & & \\ f_1^{(n-1)} & \ldots & f_n^{(n-1)} \end{pmatrix} \cdot C.$$

If the g_j are linearly dependent, then clearly their Wronskian vanishes identically. Furthermore, if $W(g_1, \ldots, g_n)$ vanishes identically, then the equation $W(g_1, \ldots, g_n) = W(f_1, \ldots, f_n) \det C$ forces $\det C = 0$, which implies the existence of a non-trivial constant column vector X with $CX = 0$, giving

$$(g_1, \ldots, g_n)X = (f_1, \ldots, f_n)CX = 0,$$

so that the g_j are linearly dependent.

The following lemma now follows via Lemma 5.4.4.

Lemma 5.12.3 *Given any n linearly independent canonical formal solutions of (5.124), their exponential parts $q_1, \ldots, q_n$ are given by a permutation of those of the principal formal solutions, and their formal Wronskian has exponential part $\sum_{j=1}^{n} q_j$.*

□

Lemma 5.12.4 *Assume that $W(f, g) = 0$, where f and g are given by*

$$f(z) = \exp(P(z^{1/p})) z^{\kappa} \sum_{j=0}^{m} (\log z)^j U_j(z), \quad g(z) = \exp(Q(z^{1/q})) z^{\lambda} \sum_{j=0}^{n} (\log z)^j V_j(z),$$

in which $p, q \in \mathbb{N}$, the U_j and V_j are formal series in descending powers of $z^{1/p}$ and $z^{1/q}$, with U_m, V_n not the zero series, while $\kappa, \lambda \in \mathbb{C}$ and P and Q are polynomials. Then f and g are linearly dependent.

In particular this holds if f and g are canonical formal solutions of an equation (5.124).

Proof. By taking the least common multiple it may be assumed that $p = q = 1$. Then the vanishing of $W(f, g)$ gives

$$\begin{aligned} 0 &= \exp(P(z) + Q(z)) z^{\lambda+\kappa} \sum_{j=0}^{m+n} (\log z)^j W_j(z), \\ W_{m+n}(z) &= \left(Q'(z) + \frac{\lambda}{z} + \frac{V_n'(z)}{V_n(z)} - P'(z) - \frac{\kappa}{z} - \frac{U_m'(z)}{U_m(z)} \right) U_m(z) V_n(z). \end{aligned} \tag{5.130}$$

Therefore $P - Q$ is constant and $\lambda - \kappa \in \mathbb{Z}$, and it may be assumed first that $P = Q$ and $\lambda = \kappa$, by incorporating an integer power of z into U_m, and second that $P = Q = 0$ and $\lambda = \kappa = 0$, by the standard property $W(fh, gh) = h^2 W(f, g)$ of the Wronskian.

It will now be proved by induction on $m + n$ that if

$$F(z) = \sum_{j=0}^{m} (\log z)^j U_j(z), \quad G(z) = \sum_{j=0}^{n} (\log z)^j V_j(z), \quad m, n \geq 0, \quad W(F, G) = 0,$$

then F and G are linearly dependent. This is clear if $m + n = 0$, since the formal series in descending powers of z form a field. Now (5.130) yields $U_m'/U_m = V_n'/V_n$, and so U_m/V_n is constant, and it may be assumed that $U_m = V_n = 1$, by the same property of the Wronskian as used earlier. It follows that

$$\begin{aligned} 0 &= ((\log z)^m + U_{m-1}(z)(\log z)^{m-1} + \ldots)((V_{n-1}'(z) + n/z)(\log z)^{n-1} + \ldots) \\ &\quad -((\log z)^n + V_{n-1}(z)(\log z)^{n-1} + \ldots)((U_{m-1}'(z) + m/z)(\log z)^{m-1} + \ldots). \end{aligned}$$

This delivers $V_{n-1}'(z) + n/z = U_{m-1}'(z) + m/z$, so that $m = n$ since $V_{n-1}'(z), U_{m-1}'(z)$ include no term in $1/z$. Now the fact that $0 = W(F, G) = W(F, G - F)$ allows $m + n$ to be reduced by at least 1, completing the induction. □

The final theorem of this section follows immediately from Theorems 5.11.1 and 5.12.1.

Theorem 5.12.2 *Suppose that $a_0, \ldots, a_{n-1}$ are holomorphic in a sector S given by $|z| > R$, $\alpha < \arg z < \beta \leq \alpha + 2\pi$, each with an asymptotic series in descending powers of z. Then (5.124) has n linearly independent principal formal solutions given by (5.125), and for each θ with $\alpha < \theta < \beta$ there exists $r(\theta) > 0$ such that (5.124) has n linearly independent holomorphic solutions*

$$g_j(z) = \exp(P_j(z^{1/p})) z^{\lambda_j} \sum_{m=0}^{n_j} (\log z)^m V_{j,m}(z)$$

with the property that $U_{j,m}(z)$ is an asymptotic series for $V_{j,m}(z)$ as $z \to \infty$ with $\theta - r(\theta) < \arg z < \theta + r(\theta)$.

□

5.12.1 Extending the sector of validity for holomorphic solutions

The following is one special case of the extension to wider sectors of asymptotic representations for solutions as in Theorem 5.12.2; for much more general results see [49].

Lemma 5.12.5 *Suppose that b_0, b_1 and b_2 are holomorphic functions on an annulus $R < |z| < \infty$, each with at most a pole at infinity, and that in the principal formal solutions of the equation*

$$y''' + b_2 y'' + b_1 y' + b_0 y = 0,$$

the exponential parts are P, $-P$ and 0, where $P(z) = a_M z^M + \ldots$ is a polynomial of positive degree M. Then there exists $p \in \mathbb{N}$ such that the principal formal solutions can be written in the form

$$F_j(z) = \exp(jP(z)) z^{\eta_j} U_j(z), \quad j = -1, 0, 1,$$

in which $\eta_j \in \mathbb{C}$ and $U_j(z)$ is a formal series in descending powers of $z^{1/p}$.

Furthermore, if $\varepsilon > 0$ and $\theta_0 \in \mathbb{R}$ satisfies $\mathrm{Re}\left(a_M e^{iM\theta_0}\right) = 0$, then there exist holomorphic solutions

$$G_j(z) = \exp(jP(z)) z^{\eta_j} V_j(z), \quad j = -1, 0, 1, \tag{5.131}$$

such that $V_j(z)$ has asymptotic series $U_j(z)$ as $z \to \infty$ with $|\arg z - \theta_0| < \pi/M - \varepsilon$.

Proof. Only the assertions concerning the G_j require proof, and it may be assumed that each $U_j(z)$ has the form

$$U_j(z) = \sum_{m=0}^{\infty} u_{j,m} z^{-m/p}, \quad u_{j,0} = 1.$$

Assume without loss of generality that $\theta_0 = 0$ and $\mathrm{Re}\left(a_M e^{iM\theta}\right) > 0$ for $0 < \theta < \pi/M$. By Theorem 5.12.2 there exist holomorphic solutions G_j as in (5.131) such that $V_j(z)$ has asymptotic series $U_j(z)$, and in particular $V_j(z) \to 1$, as $z \to \infty$ with $|\arg z| < r(0)$. For each $\phi \in (0, \pi/M)$ choose a corresponding $r(\phi)$: it may be assumed that $r(\phi) < \phi$. Compactness shows that there exist $N \in \mathbb{N}$ and $0 = \phi_0 < \phi_1 < \ldots < \phi_N$ such that the sector $0 \le \arg z \le \pi/M - \varepsilon$ is covered by the union of the sectors $|\arg z - \phi_\mu| < r(\phi_\mu)$. Here it can also be assumed that $\phi_\mu > r(0)/2$ for each $\mu \ge 1$.

Now suppose that $1 \le \mu \le N$ and, using Theorem 5.12.2 again, take holomorphic solutions

$$H_{k,\mu}(z) = \exp(kP(z)) z^{\eta_k} W_{k,\mu}(z), \quad k = -1, 0, 1,$$

such that $W_{k,\mu}(z)$ has asymptotic series $U_k(z)$ (and $W_{k,\mu}(z) \to 1$) as $z \to \infty$ with $|\arg z - \phi_\mu| < r(\phi_\mu)$. Since the G_j extend holomorphically into the sector $|\arg z| < \pi/M - \varepsilon$, there exist constants $c_{j,k,\mu}$ with

$$G_j = \sum_{k \in \{-1,0,1\}} c_{j,k,\mu} H_{k,\mu}.$$

Claim A: *Let $k > j \in \{-1, 0, 1\}$: then $c_{j,k,\mu}$ is 0 for each μ.*

To see this, take the largest $k > j$ for which there exists $\mu \in \{1, \ldots, N\}$ with $c_{j,k,\mu} \neq 0$, and choose such a μ. Then the holomorphic function

$$G_j(z) \exp(-kP(z)) z^{-\eta_k}$$

tends to $c_{j,k,\mu} \neq 0$ as $z \to \infty$ with $\arg z = \phi_\mu$, and to 0 as $z \to \infty$ with $\arg z = r(0)/2$, and is bounded as $z \to \infty$ in the sector between these rays, which contradicts the Phragmén-Lindelöf principle. This proves Claim A.

Claim A implies that, for $1 \le \mu \le N$,

$$G_j(z) \exp(-jP(z)) z^{-\eta_j}$$

tends to $c_{j,j,\mu}$ as $z \to \infty$ with $\arg z = \phi_\mu$, and to 1 as $z \to \infty$ with $\arg z = r(0)/2$, and is bounded in the sector between, and so the Phragmén-Lindelöf principle forces $c_{j,j,\mu} = 1$. Now write

$$\begin{aligned} G_j(z) &= \exp(jP(z))z^{\eta_j}V_j(z) = \exp(jP(z))z^{\eta_j}W_{j,\mu}(z) + \sum_{k<j} c_{j,k,\mu}H_{k,\mu}(z) \\ &= \exp(jP(z))z^{\eta_j}\left(W_{j,\mu}(z) + \sum_{k<j} c_{j,k,\mu}\exp((k-j)P(z))z^{\eta_k-\eta_j}W_{k,\mu}(z)\right). \end{aligned}$$

Here $U_j(z)$ is an asymptotic series for $W_{j,\mu}(z)$ as $z \to \infty$ with $|\arg z - \phi_\mu| < r(\phi_\mu)$, and so also for $V_j(z)$. A similar argument handles the sector $-\pi/M + \varepsilon \leq \arg z \leq 0$. $\square$

Chapter 6

Meromorphic flows

6.1 Introduction

The standard application of complex analysis to (incompressible, irrotational) fluid flow on a plane domain D, as given in many textbooks, goes as follows. The velocity of the fluid at $z \in D$

$$\dot{z} = \frac{dz}{dt} = \overline{g(z)}, \tag{6.1}$$

where $g = u + iv$ is analytic on D (with u, v real). In this model, if D is simply connected, the streamlines (trajectories) along which particles of fluid travel are found as follows. Let $G = P + iQ$ be analytic on D, with $G' = g$ and $Q = \operatorname{Im} G$. Then a streamline $z(t) = x(t) + iy(t)$ is determined by writing

$$g = u + iv = G' = P_x + iQ_x = Q_y + iQ_x, \quad \frac{dQ}{dt} = Q_x x_t + Q_y y_t = Q_x u - Q_y v = vu - uv = 0,$$

so that Q is constant on the streamline. Hence the trajectories in this model are determined by finding level curves of Q.

Consider next a meromorphic flow given by

$$\dot{z} = \frac{dz}{dt} = f(z), \tag{6.2}$$

in which the function f is meromorphic on a simply connected domain $D \subseteq \mathbb{C}$ (this will be the case throughout this chapter). Suitable references for these flows include [22, 23, 27, 30, 31, 50]. A trajectory for (6.2) will mean a continuous $z(t)$, defined on some maximal open interval of $\mathbb{R}$, with

$$z = z(t) \in D, \quad \frac{dz(t)}{dt} = f(z(t)) \in \mathbb{C}.$$

It will be shown in §6.3 that for $z_0 \in D$ with $f(z_0) \neq \infty$ there exists a unique trajectory with $z(0) = z_0$, and that $z(t)$ depends continuously (and indeed analytically for fixed t) on z_0.

6.1.1 A connection between (6.1) and (6.2)

If in (6.1) we set $f(z) = 1/g(z)$ then we obtain

$$\dot{z} = \frac{dz}{dt} = \overline{g(z)} = \frac{1}{\overline{f(z)}} = \frac{f(z)}{|f(z)|^2}. \tag{6.3}$$

This flow is therefore linked to (6.2), insofar as at every point the *direction* of travel is the same, although in general the speed is not. Indeed, given a trajectory $z(t)$ of (6.2) through the point $z(0)$, define $s = \phi(t)$ by

$$\phi(0) = 0, \quad \frac{ds}{dt} = \phi'(t) = |f(z(t))|^2.$$

Then ϕ is strictly increasing, with inverse function $t = \psi(s)$. Now set $w(s) = z(t)$, which gives $w(0) = z(0)$ and

$$w'(s) = z'(t)\,\psi'(s) = \frac{f(z(t))}{\phi'(t)} = \frac{f(z(t))}{|f(z(t))|^2} = \frac{f(w(s))}{|f(w(s))|^2},$$

so $w(s) = z(t) = z(\psi(s))$ is a trajectory of (6.3) which passes through the same points as $z(t)$, but at different speed.

6.2 Examples

6.2.1 Example I

Let $f(z) = z^2$ in (6.2). Then any trajectory $z(t)$ with $z(0) = 1/T \neq 0$ has

$$\frac{1}{z(t)} = \frac{1}{z(0)} - t = T - t.$$

If T is real and positive then $z(t)$ is real and tends to $+\infty$ as $t \to T-$, and to 0 as $t \to -\infty$. Thus $z(t)$ follows the positive real axis in the outward direction as t goes from $-\infty$ to T.

If T is real and negative then $z(t)$ is real and tends to $-\infty$ as $t \to T+$, and to 0 as $t \to +\infty$. Thus $z(t)$ follows the negative real axis in the inward direction as t goes from T to $+\infty$.

If T is non-real then $z(t)$ is defined for all real t, and as $t \to \pm\infty$ we have $1/z(t) \to \infty$ and $z(t) \to 0$.

6.2.2 Example II

Suppose that $f(z) = e^z$ in (6.2). Then integration gives

$$e^{-z(t)} = e^{-z(0)} - t$$

for any trajectory $z(t)$. If $T = e^{-z(0)}$ is real and positive (that is, if $\mathrm{Im}\, z(0) = k2\pi$ for some integer k) then $e^{-z(t)} \to 0$ and so $z(t) \to \infty$ as $t \to T-$. In this case the trajectory moves from left to right along the horizontal line $\mathrm{Im}\, z = k2\pi i$, and the maximal interval of definition of the trajectory is $(-\infty, T)$.

If $e^{-z(0)}$ is not real and positive then $z(t)$ is defined for all $t \in \mathbb{R}$. As $t \to \pm\infty$, the term $e^{-z(t)}$ tends to infinity, so that $z(t)$ tends to infinity in the left half-plane.

6.3 Existence and uniqueness

The following standard argument shows that for $z_0 \in D$ with $f(z_0) \neq \infty$ there is at most one trajectory $z(t)$ with $z(0) = z_0$. Take positive real numbers M and δ such that δ is small and $|z - z_0| \leq \delta$ gives $|f'(z)| \leq M$. Suppose that $\eta > 0$ and that $z_1(t)$ and $z_2(t)$ are trajectories which are both defined for $|t| \leq \eta$, and which satisfy $z_1(0) = z_2(0) = z_0$. It may be assumed that η is so small that $|t| \leq \eta$ gives $|z_j(t) - z_0| \leq \delta$. Suppose that $|t| \leq \lambda = \max\{\eta, 1/2M\}$, and that

$$|z_1(s) - z_2(s)| \leq |z_1(t) - z_2(t)| \quad \text{for } |s| \leq |t|.$$

Then

$$\begin{aligned}
|z_1(t) - z_2(t)| &= \left| \int_0^t f(z_1(s)) - f(z_2(s))\, ds \right| \\
&= \left| \int_0^t \int_{z_2(s)}^{z_1(s)} f'(u)\, du\, ds \right| \\
&\leq \int_0^t M|z_1(s) - z_2(s)|\, ds \\
&\leq \int_0^t M|z_1(t) - z_2(t)|\, ds \leq \frac{|z_1(t) - z_2(t)|}{2}.
\end{aligned}$$

This forces $z_1(t) = z_2(t)$ for $0 \leq |t| \leq \lambda$ and repetition of the same argument shows that the trajectories $z_j(t)$ are identical. Thus if $f(z_0) = 0$ then the only trajectory through z_0 is the trivial solution $z(t) \equiv z_0$.

When $z_0 \in D$ and $f(z_0) \neq 0, \infty$, the local existence and uniqueness of the trajectory through z_0 may be established by the following argument. Choose $r > 0$ with

$$\left| \frac{1}{f(z)} - \frac{1}{f(z_0)} \right| \leq \frac{1}{2} \left| \frac{1}{f(z_0)} \right|$$

on $D(z_0, r) = \{z \in \mathbb{C} : |z - z_0| < r\}$. Then

$$G(z) = \int_{z_0}^{z} \frac{1}{f(u)}\, du = \frac{z - z_0}{f(z_0)} + \int_{z_0}^{z} \frac{1}{f(u)} - \frac{1}{f(z_0)}\, du \tag{6.4}$$

satisfies

$$G(z_0) = 0 \quad \text{and} \quad |G(z_1) - G(z_2)| \geq \frac{|z_1 - z_2|}{2f(z_0)|}$$

on $D(z_0, r)$. Thus G is analytic and univalent on $D(z_0, r)$. Taking $z_2 = z_0$ and applying Rouché's theorem shows that $G(D(z_0, r))$ contains the disc $D(0, s)$, where $s = r/2|f(z_0)|$. Now the local change of variables $w = G(z)$ gives the flow $\dot{w} = 1$, which evidently has a unique trajectory with $w(0) = 0$ given by $w(t) = t$. Hence a trajectory of (6.2) satisfying $z(0) = z_0$, and defined at least for $-s < t < s$, is given uniquely by $z(t) = G^{-1}(t)$.

6.4 Dependence on initial conditions

Suppose that $z_0 \in D$ is such that the trajectory $z(t)$ exists and is injective for $0 \leq t \leq A$, with $A > 0$ and $z_0 = z(0)$. Then f has neither zeros nor poles on the curve $\gamma = \{z(t) : 0 \leq t \leq A\}$. Hence the function G of (6.4) is analytic on a simply connected domain Ω containing γ, and maps γ onto the real interval $[0, A]$, with $G(z(t)) = t$ and $G(z_0) = 0$. Here Ω may be formed as follows: let $\mathbb{C}^\infty = \mathbb{C} \cup \{\infty\}$ and map the complement of γ on the Riemann sphere conformally to $\{v \in \mathbb{C}^\infty : |v| > 1\}$ by $v = h(z)$, so that ∞ is mapped to ∞. Then take $S_1 > 1$ such that the images under h of all zeros of f in D lie in $X_1 = \{v \in \mathbb{C}^\infty : |v| \geq S_1\}$. Finally, let Ω be the complement of the closed connected subset $h^{-1}(X_1)$ of $\mathbb{C}^\infty$.

The next step is to choose a sub-domain $\Omega' \subseteq \Omega$ such that G is univalent on Ω'. By compactness and the argument of §6.3, there exists $r > 0$ such that, for each $t \in [0, A]$, the function G is univalent on $D(z(t), r)$, which lies in Ω. Uniform continuity gives $R > 0$ such that $|z(t) - z(t')| < r/2$ for all $t, t' \in [0, A]$ with $|t - t'| < R$, and for each $s \in [0, A]$ there exists $p(s) \in (0, r/2)$ with $|G(z) - G(z(s))| < R/2$ for all z in $D(z(s), p(s))$.

Let Ω' be the union of the discs $D(z(s), p(s))$, for $0 \leq s \leq A$. Then Ω' is a domain. If $z, z' \in \Omega'$ and $G(z) = G(z')$ then there exist $s, s' \in [0, A]$ with $z \in D(z(s), p(s))$ and $z' \in D(z(s'), p(s'))$. This leads to

$$|s - s'| = |G(z(s)) - G(z(s'))| = |G(z(s)) - G(z) + G(z) - G(z') + G(z') - G(z(s'))| < R$$

and so $z(s') \in D(z(s), r/2)$, and $z, z' \in D(z(s), r)$, giving $z = z'$. Thus G is univalent on Ω', as required.

Now for w close to z_0 the formula $\zeta_w(t) = G^{-1}(G(w)+t)$ defines a trajectory of (6.2) for $0 \leq t \leq A$, starting at w, and shows that $\zeta_w(t)$ is close to $z(t) = G^{-1}(t)$, uniformly for $0 \leq t \leq A$. Moreover, for fixed $t \in [0, A]$, the position $\zeta_w(t)$ depends analytically on w.

6.5 Re-scaling, conjugacy and simple zeros

Suppose that $z_0 \in D$ with $f(z_0) \neq 0, \infty$. Then for any $a, b \in \mathbb{C}$ with $a \neq 0$, a re-scaled flow may be defined by

$$w = az + b, \quad g(w) = af(z) = af((w-b)/a), \quad \dot{w} = a\dot{z} = af(z) = g(w). \tag{6.5}$$

Here any prescribed value may be assigned to $w_0 = az_0 + b$, and any prescribed non-zero value to $g(w_0)$.

Suppose next that f has a simple zero at $z_0 \in D$. Assume without loss of generality that $z_0 = 0$, and set $\alpha = f'(0) \neq 0$. Then writing

$$w = \psi(z), \quad \frac{\psi'(z)}{\psi(z)} = \frac{\alpha}{f(z)} = \frac{1}{z} + \ldots, \tag{6.6}$$

defines a conformal change of variables near the origin, and yields

$$\dot{w} = \psi'(z)\dot{z} = \psi'(z)f(z) = \alpha\psi(z) = \alpha w.$$

If $z(t)$ is a trajectory of (6.2) passing near to 0 then $w(t) = \psi(z(t))$ is a trajectory of

$$\dot{w} = \alpha w. \tag{6.7}$$

6.5.1 The case where α is real

In this case the flow (6.7) has a *node* at 0 (see [23]). The trajectory through a starting point $w_0 \neq 0$ satisfies $w(t) = w_0 e^{\alpha t}$ and is a ray, the direction of flow determined by the sign of α. Thus all trajectories of (6.2) in a punctured neighbourhood of 0 flow towards, or away from, 0.

6.5.2 The case where α is neither real nor purely imaginary

This case is referred to as a *focus*. The trajectory through a starting point $w_0 \neq 0$ still satisfies $w(t) = w_0 e^{\alpha t}$, but is a spiral. All trajectories of (6.2) in a punctured neighbourhood of 0 either spiral into, or away from, the fixpoint at the origin.

A focus or node is called attracting (or a sink) if $\mathrm{Re}\,\alpha < 0$, and repelling (or a source) when $\mathrm{Re}\,\alpha > 0$.

6.5.3 The case where α is purely imaginary

Here the flows (6.2) and (6.7) have a *centre* at 0. The trajectory of (6.7) through a starting point $w_0 \neq 0$ satisfies $w(t) = w_0 e^{\alpha t}$, but is this time a circle, and all trajectories of (6.2) in a punctured neighbourhood of 0 are periodic and flow around the fixpoint at the origin.

6.6 The behaviour near poles

Suppose that $f(z) \sim c(z-z_0)^{-m}$ as $z \to z_0$, for some $c \neq 0$ and $m \geq 0$. Define a conformal mapping $w = \phi(z)$ near z_0 by writing

$$\phi(z)^{m+1} = \int_{z_0}^{z} \frac{1}{f(u)}\, du = \frac{(z-z_0)^{m+1}}{(m+1)c} + \dots .$$

This gives

$$(m+1)\dot{w} = w^{-m}, \quad w^{m+1}(t) = w^{m+1}(0) + t. \tag{6.8}$$

The equation (6.8) has $m+1$ disjoint trajectories tending to 0 in increasing time, determined by choosing $w^{m+1}(0) \in (-\infty, 0) \subseteq \mathbb{R}$. Thus (6.2) has $m+1$ trajectories tending to z_0 in increasing time (each taking finite time to do so).

Suppose next that D contains an annulus $R < |z| < \infty$ and f has a pole of order $n \geq 2$ at infinity. Setting $w = 1/z$ gives $\dot{w} = g(w) = -f(z)/z^2$, so that g has a pole of order $n-2$ at 0 and (6.2) has $n-1$ trajectories tending to infinity in finite increasing time: this is a result of King and Needham [50, Theorem 5].

6.7 Periodic cycles and their stability

Suppose that $z_0 \in D$ and $f(z_0) \neq 0, \infty$ and that the trajectory through z_0 satisfies $z_0 = z(0) = z(T)$ for some (minimal) positive T. Then z_0 lies on a periodic cycle, and its trajectory describes a Jordan curve Γ in D, which has

$$\int_{\Gamma} \frac{1}{f(u)}\, du = T. \tag{6.9}$$

It will be shown that if z_1 lies close enough to z_0 then z_1 also lies on a periodic cycle of period T.

The following approach is used in [30, Theorem 2]. For z close to z_0 let $\zeta_z(t)$ be the trajectory with $\zeta_z(0) = z$. Then $\zeta_z(T)$ depends analytically on z; to see this, split Γ into two injective sub-trajectories, each taking time $T/2$ to describe, and use the method of §6.3 and the chain rule (since $\zeta_z(T/2)$ depends analytically on z). But if z lies on Γ then $\zeta_z(T) - z = 0$. So $\zeta_z(T) = z$ for all z close to z_0, by the identity theorem. Continuous dependence on initial conditions and (6.9) imply that the period is the same.

An alternative proof proceeds as follows. Let δ be small and positive and take the pre-image $L = L_\delta(z_0)$ of the real interval $[-\delta, \delta]$ under the function $i\int_{z_0}^{z} 1/f(u)\, du$. Then any trajectory which meets L does so non-tangentially. For $z \in L$, close to z_0, follow the trajectory through z until the first point z' at which it meets L again, as it must by continuous dependence on initial conditions, and suppose that $z \neq z'$. Joining z' to z by a sub-arc of L gives a simple closed curve Γ' for which $\int_{\Gamma'} 1/f(u)\, du$ is non-real. But Cauchy's theorem gives $\int_{\Gamma'} 1/f(u)\, du = \int_{\Gamma} 1/f(u)\, du = T \in \mathbb{R}$.

6.7.1 Periodic cycles not enclosing poles

The following argument is adapted from [22]. Suppose again that $z_0 \in D$ and $f(z_0) \neq 0, \infty$ and the trajectory through z_0 satisfies $z_0 = z(0) = z(T)$ for some (minimal) positive T, so that $z(t)$ describes a Jordan curve Γ in D as t goes from 0 to T. Assume that the interior domain of Γ lies in D but contains no poles of f.

By (6.9) and Cauchy's theorem, Γ must enclose at least one zero of f, without loss of generality at 0. Let $z = g(v)$ be the Riemann mapping from $D(0,1)$ to the interior domain of Γ, with $g(0) = 0$. Set

$$G(z) = \exp\left(\frac{2\pi i}{T} \int_{z_0}^{z} \frac{1}{f(u)}\, du\right)$$

near z_0. By compactness and the discussion in §6.7, there exists $\delta_1 > 0$ such that if the distance from w to Γ is less than δ_1 then w lies on a periodic cycle of period T. Moreover, there exists $\delta_2 > 0$ such that if $|w - z_0| < \delta_2$ then the trajectory of (6.2) through w always has distance less than δ_1 from Γ. Take a small positive δ, so small that the pre-image $L = L_\delta(z_0)$ of the real interval $(-\delta, \delta)$ under the function $i \int_{z_0}^{z} 1/f(u)\, du$ lies within the disc of centre δ_2 and radius z_0.

Let Ω be the union of all trajectories which meet L, each of these having period T. Then Ω is open, and doubly connected, since any point lying between two of these periodic trajectories must also lie on a periodic trajectory, which must in turn meet L. The function G clearly continues analytically throughout Ω. Moreover, G maps the trajectory through $w \in \Omega$ injectively onto a circle of centre 0, its radius determined by the real part of

$$\frac{2\pi i}{T} \int_{z_0}^{w} \frac{1}{f(u)}\, du,$$

and hence by the point at which the trajectory meets L. Thus G extends to be analytic and univalent on Ω, mapping Ω onto an open annulus containing the unit circle.

Moreover, $|G(g(u))|$ is defined and tends to 1 as $|u| \to 1-$. If $|u_0| = 1$ then reflection gives an extension of $H = G \circ g$ to a disc $D(u_0, \sigma_0)$ with $u_0 > 0$. A compactness argument and the fact that the intersection of two discs is connected extends H analytically to an annulus Ω_1 given by $1/R < |u| < R$, where $R > 1$. This extension has the property that if $u^* = 1/\overline{u}$ is the reflection of u across the unit circle, then $H(u^*)$ is the reflection of $H(u)$. Thus H is univalent for $1/R < |u| < R$, because it is univalent for $1/R < |u| < 1$. As u crosses the unit circle, so does $H(u)$, and therefore $G^{-1}(H(u))$ crosses Γ. Hence g may be extended analytically to $D(0, R)$ by writing $g = G^{-1} \circ H$ on Ω_1, and $g(u)$ lies outside Γ for $1 < |u| < R$. This property, coupled with the fact that $G^{-1} \circ H$ is univalent on Ω_1, ensures that g is analytic and univalent on $V = D(0, R)$.

Now consider the equation

$$\dot{v} = \frac{f(g(v))}{g'(v)} = \sigma(v) = v\rho(v) \tag{6.10}$$

on V. Here ρ is analytic on V since $f(g(0)) = 0$ and f is analytic on $g(V)$ (because Γ encloses no poles of f). The unit circle is a periodic trajectory of this flow, since $g(v) \in \Gamma$ for $|v| = 1$ and $z = g(v)$ gives $\dot{z} = f(z)$. This means that for $|v| = 1$ the vector $\sigma(v)$ must be perpendicular to the vector v, and so $\rho(v)$ must be purely imaginary. But then $\rho(v)$ is constant on V by the maximum principle for harmonic functions. Thus the flow (6.10) reduces to $\dot{v} = \lambda v$, where $\rho(v) \equiv \lambda \in i\mathbb{R} \setminus \{0\}$. Using the Taylor expansion of f and g about 0 shows that $\lambda = f'(0)$.

If $0 < r < R$ then the circle $|v| = r$ is mapped by $z = g(v)$ to a Jordan curve Γ_r, and

$$\frac{2\pi i}{\lambda} = \int_{|v|=r} \frac{1}{\lambda v}\, dv = \int_{|v|=r} \frac{g'(v)}{f(g(v))}\, dv = \int_{\Gamma_r} \frac{1}{f(z)}\, dz.$$

Setting $r = 1$ shows that $\lambda = f'(0) = 2\pi i/T$. Thus each circle $|v| = r \in (0, 1]$ is a cycle of (6.10) with period $T = 2\pi i/f'(0)$, and every point in $g(V)$ lies on a periodic cycle of (6.2) with the same period. In particular this is true for all points inside Γ, and all points close enough to Γ. Also 0 is the only zero of f in $g(V)$, because of the equation $f(g(v)) = \lambda v g'(v)$.

Now let P be the union of $\{0\}$ and all periodic trajectories γ which enclose 0 (that is, have non-zero winding number about 0) but enclose no poles of f. Then P is open, by the above argument (or by stability of periodic cycles), and is a domain since the interior of each such γ contains a neighbourhood of 0.

In fact, P is simply connected, for the following reason. Let Λ be a Jordan curve in P. For each $z \in \Lambda$ there exists a cycle $\gamma_z \subseteq P$ which encloses z, and if $z' \in \Lambda$ lies close enough to z then z' also lies inside Γ_z. Compactness gives finitely many cycles $\gamma_{z_j} \subseteq P$, each enclosing 0, such that every $z \in \Lambda$

lies inside at least one of them. But these cycles either coincide or are disjoint, and so one of them, γ say, must enclose all the others. But then the interior of γ lies in P, and so does that of Λ.

Lemma 6.7.1 ([22]) *Suppose that $D = \mathbb{C}$ and every $z_0 \in \mathbb{C} \setminus \{0\}$ lies on a periodic cycle enclosing 0. Then $f(z) = \alpha z$ for some $\alpha \in i\mathbb{R} \setminus \{0\}$.*

Proof. The above argument shows that 0 is the only zero of f, and all the cycles have the same minimal period T. The function

$$F(z) = \exp\left(\frac{2\pi i}{T}\int_1^z \frac{1}{f(u)}\,du\right)$$

is analytic on the plane, and univalent on, and so inside, each periodic cycle. Thus F is a univalent entire function and so linear, and so are F/F' and f.

6.8 An example

Following [22], consider the flow

$$\left(\frac{1}{w^n} + \frac{\lambda}{w}\right)\dot{w} = 1 \tag{6.11}$$

on $\mathbb{C}$, where $\lambda \in \mathbb{C}$. To determine trajectories for (6.11) set

$$u = \frac{w^{1-n}}{1-n}, \qquad \frac{\dot{u}}{u} = (1-n)\frac{\dot{w}}{w},$$

so that u satisfies, near infinity,

$$((1-n)u + \lambda)\frac{\dot{u}}{u} = 1-n, \quad \left(u + \frac{\lambda}{1-n}\right)\frac{\dot{u}}{u} = 1. \tag{6.12}$$

Let R, S/R and T/S be large and positive and consider first a trajectory $u(t)$ which has

$$|u(0)| \geq T, \quad \mathrm{Re}\left(u(0) + \frac{\lambda}{1-n}\right) \geq 0. \tag{6.13}$$

Write

$$v = u + \frac{\lambda}{1-n}\log u \quad \text{on} \quad D_R^+ = \{u \in \mathbb{C} : |u| > R,\, -\pi < \arg u < \pi\}. \tag{6.14}$$

Then the trajectory $u(t)$ has $u(0) \in D_R^+$ and $|\arg u(0)| < \pi/2 + \delta$, where $\delta > 0$ can be chosen arbitrarily small, subject to T being large enough. But then $|\arg v(0)| < \pi/2 + 2\delta$ and $\dot{v} = 1$, and so $v(t) = v(0) + t$ has $S \leq |v(t)| \to +\infty$ and $|\arg v(t)| < \pi/2 + 2\delta$ for $t \geq 0$. Since $\mathrm{Re}\, v$ is bounded above as $u \to \partial D_R^+$, the trajectory for u stays in D_R^+ and also tends to infinity, with $\arg u(t) \to 0$ as $t \to +\infty$.

Now suppose that

$$|u(0)| \geq T, \quad \mathrm{Re}\left(u(0) + \frac{\lambda}{1-n}\right) \leq 0. \tag{6.15}$$

This time writing

$$v = u + \frac{\lambda}{1-n}\log u \quad \text{on} \quad D_R^- = \{u \in \mathbb{C} : |u| > R,\, 0 < \arg u < 2\pi\}$$

gives $\dot{v} = 1$ again, and shows that, as $t \to -\infty$, both $v(t)$ and $u(t)$ tend to infinity, with $\arg u(t) \to \pi$.

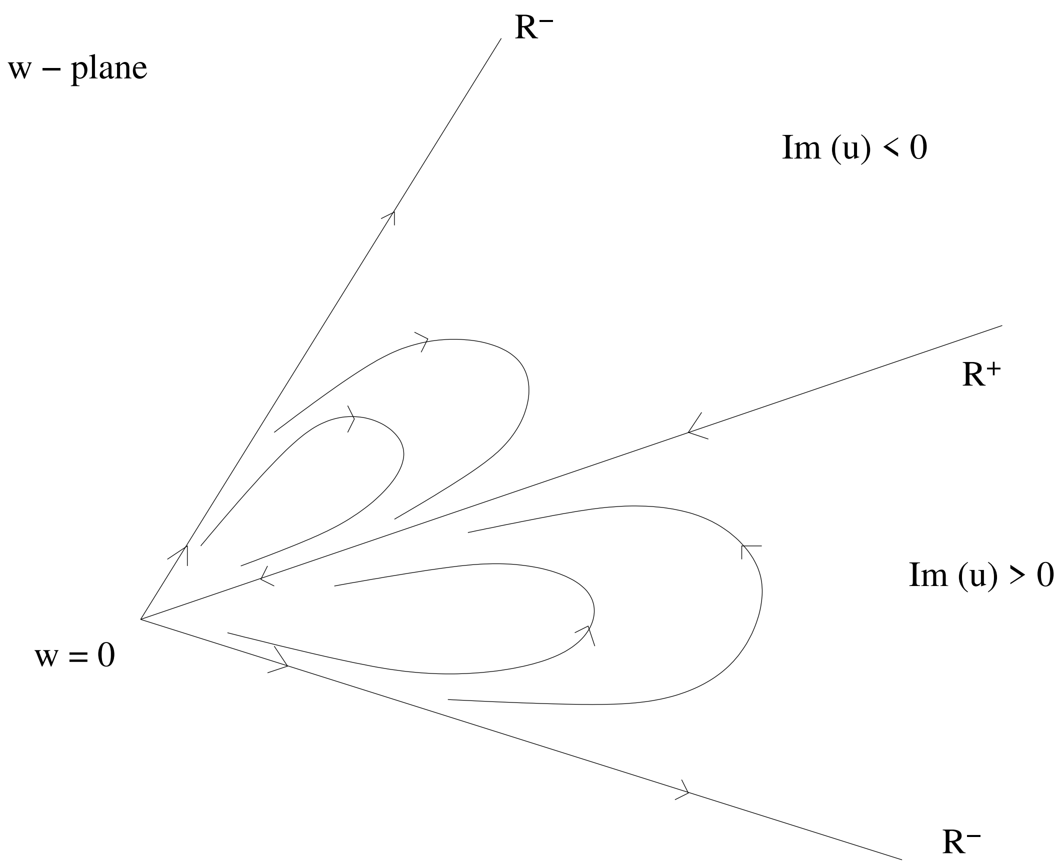

Figure 6.1: Trajectories of (6.11) near the origin

Now take any trajectory $u(t)$ which has $|\mathrm{Im}\, u(0)| \geq T$. Then $u(0) \in D_R^+$. Continue $u(t)$ in the directions of both increasing and decreasing t, as far as is possible while keeping $u \in D_R^+$, and define v by (6.14). Since this gives $|\mathrm{Im}\, v(t)| = |\mathrm{Im}\, v(0)| \geq S$, whereas $\mathrm{Im}\, v$ is bounded on ∂D_R^+, this continuation never causes u to exit D_R^+, and $u(t) \to \infty$ as $t \to \pm\infty$. Again $\arg u(t) \to 0$ as $t \to +\infty$, and $\arg u(t) \to \pi$ as $t \to -\infty$.

In summary, $u(t) \to \infty$ and $w(t) \to 0$ as $t \to +\infty$ when (6.13) holds, and $u(t) \to \infty$ and $w(t) \to 0$ as $t \to -\infty$ when (6.15) holds. Every trajectory for (6.11) with $|w(0)|$ small enough is such that at least one of these is satisfied, and there are infinitely many trajectories for which both hold.

Now let $s > 0$ be small and take a trajectory $w(t)$ of (6.11) for which $|w(0)| = s$, so that $|u(0)|$ is large. If, at time $t = 0$,

$$\mathrm{Re}\left(\frac{\dot{u}}{u}\right) = \frac{d}{dt}\left(\log|u(t)|\right) \geq 0,$$

then (6.12) implies that (6.13) holds, and so $u(t) \to \infty$ and $w(t) \to 0$ as $t \to +\infty$. Similarly, any trajectory $w(t)$ of (6.11) which has $|w(0)| = s$ and $|w|$ non-decreasing at time $t = 0$ is such that (6.15) holds, so that $u(t) \to \infty$ and $w(t) \to 0$ as $t \to -\infty$. Therefore every trajectory of (6.11) which meets $|w| = s$ tends to 0 as $t \to +\infty$ or as $t \to -\infty$ or both, depending on the sign of $\mathrm{Re}\,(\dot{u}/u)$ (and hence of $\mathrm{Re}\,(\dot{w}/w)$).

This gives rise to "elliptic sectors" in the terminology of [23]. Divide up a neighbourhood of $w = 0$ into sectors on which $\mathrm{Im}\, u$ is alternately positive and negative; each is bounded by rays R^+, R^- on which u is real and positive, negative respectively. If $|\mathrm{Im}\, u(0)| \geq T$ then $w(t) \to 0$ as $t \to \pm\infty$, with $w(t) \to R^+$ as $t \to +\infty$ and $w(t) \to R^-$ as $t \to -\infty$. If $|u(0)|$ is large enough but $|\mathrm{Im}\, u(0)| < T$ then one of (6.13) and (6.15) is satisfied, and $u(0)$ is close to R^+ or R^-, and the trajectory tends to zero in increasing or decreasing time.

6.9 Multiple zeros

Assume that f has a zero of multiplicity $n \geq 2$ at the origin. Then the following argument from [22] shows that (6.2) is conjugate near 0 to an equation of form (6.11). In a neighbourhood of 0 write

$$\frac{1}{f(z)} = \frac{b_n}{z^n} + \ldots + \frac{b_2}{z^2} + \frac{\lambda}{z} + q(z), \quad b_n \neq 0, \quad q(0) \in \mathbb{C}, \tag{6.16}$$

and

$$f_1(z) = \frac{b_n}{(1-n)z^{n-1}} + \ldots - \frac{b_2}{z} + \int_0^z q(u)\,du = \frac{g_1(z)}{z^{n-1}}, \quad f_2(w) = \frac{1}{(1-n)w^{n-1}} = \frac{g_2(w)}{w^{n-1}}, \tag{6.17}$$

so that $f_1'(z) = 1/f(z) - \lambda/z$ and $f_2'(w) = 1/w^n$ by (6.16). Choose μ_0 so that

$$\frac{g_1(0)}{g_2(0)} = b_n = \frac{1}{\mu_0^{n-1}}, \tag{6.18}$$

using (6.17). For v near to 0 and μ close to μ_0, set

$$H(v, \mu) = \frac{g_2(v\mu)}{\mu^{n-1}} + \lambda v^{n-1} \log \mu - g_1(v). \tag{6.19}$$

Here

$$H(0, \mu_0) = \frac{g_2(0)}{\mu_0^{n-1}} - g_1(0) = 0$$

by (6.18), and

$$\frac{\partial H}{\partial \mu}(0,\mu_0) = \frac{(1-n)g_2(0)}{\mu_0^n} \neq 0.$$

Thus the implicit function theorem (Lemma 6.11.1) gives a function $\psi(v)$ with $\psi(0) = \mu_0 \neq 0$ such that ψ is analytic near 0 and satisfies there

$$H(v,\psi(v)) = 0. \tag{6.20}$$

Set $\phi(v) = v\psi(v)$, so that ϕ is conformal in a neighbourhood of 0, with $\phi(0) = 0$. Then (6.19) and (6.20) yield, near 0,

$$\begin{aligned} 0 &= \frac{g_2(v\psi(v))}{\psi(v)^{n-1}} + \lambda v^{n-1}\log\psi(v) - g_1(v) \\ &= \frac{v^{n-1}g_2(\phi(v))}{\phi(v)^{n-1}} + \lambda v^{n-1}\log\frac{\phi(v)}{v} - g_1(v), \end{aligned}$$

and so

$$\begin{aligned} 0 &= \frac{g_2(\phi(v))}{\phi(v)^{n-1}} + \lambda\log\frac{\phi(v)}{v} - \frac{g_1(v)}{v^{n-1}}, \\ &= f_2(\phi(v)) + \lambda\log\frac{\phi(v)}{v} - f_1(v). \end{aligned}$$

Differentiating now yields

$$\begin{aligned} 0 &= \frac{\phi'(v)}{\phi(v)^n} + \lambda\left(\frac{\phi'(v)}{\phi(v)} - \frac{1}{v}\right) - \frac{1}{f(v)} + \frac{\lambda}{v} \\ &= \frac{\phi'(v)}{\phi(v)^n} + \lambda\frac{\phi'(v)}{\phi(v)} - \frac{1}{f(v)}. \end{aligned}$$

Given a trajectory $z(t)$ of (6.2) near 0 write $w(t) = \phi(z(t))$ so that

$$\dot{w}\left(\frac{1}{w^n} + \frac{\lambda}{w}\right) = \dot{z}\phi'(z)\left(\frac{1}{\phi(z)^n} + \frac{\lambda}{\phi(z)}\right) = \frac{\dot{z}}{f(z)} = 1,$$

which makes $w(t)$ a trajectory of (6.11).

Lemma 6.9.1 *Suppose that f has a zero of multiplicity $n \geq 2$ at $z_0 \in \mathbb{C}$, and let $\delta > 0$. Then there exists a Jordan curve $C \subseteq D(z_0,\delta)$ which surrounds z_0 and has the following properties: any trajectory $z(t)$ of (6.2) which passes from outside C to inside in increasing time tends to z_0 as $t \to +\infty$; any trajectory $z(t)$ which passes from inside C to outside in increasing time tends to z_0 as $t \to -\infty$. Furthermore, there exists at least one trajectory $z(t)$ which remains inside C and tends to z_0 as $t \to \pm\infty$.*

The lemma is proved by assuming that $z_0 = 0$ and taking C to be the pre-image under ϕ of the circle $|w| = s$, for some small positive s; the asserted properties all hold by §6.8.

6.10 Limit points of trajectories

Lemma 6.10.1 *Let the function f be meromorphic and non-constant on a simply connected domain $D \subseteq \mathbb{C}$, with finitely many zeros in D, or finitely many poles in D. Let $z(t)$ be a non-periodic trajectory of (6.2), with maximal interval of definition $(a_0,b_0) \subseteq \mathbb{R}$. Suppose that $z_0 \in D$ is a limit point of $z(t)$ as $t \to b_0$, that is, there exist $s_n \in (a_0,b_0)$ with $s_n \to b_0$ such that $z(s_n) \to z_0$. Then $f(z_0) \in \{0,\infty\}$ and $\lim_{t\to b_0-} z(t) = z_0$.*

Proof. It suffices to show that $f(z_0) \in \{0,\infty\}$; once this is proved, it must be the case that $\lim_{t\to b_0-} z(t) = z_0$, since otherwise there exists $z_0' \in D$ with $f(z_0') \notin \{0,\infty\}$ such that z_0' is a limit point of $z(t)$ as $t \to b_0$.

Assume then that $z(t)$ and z_0 are as in the hypotheses, but that $f(z_0) \neq 0, \infty$. Observe that $z(t)$, being non-periodic, must be injective for $a_0 < t < b_0$. By employing a linear re-scaling $w = az + b$, $g(w) = af(z)$, it may be assumed that $z_0 = 0$ and $f(z_0) = i$.

If $z(t) \to 0$ as $t \to b_0$ with $t \in (a_0, b_0)$, then so does $u(t) = \phi(z(t))$, where $\phi(z) = \int_0^z 1/f(s)\,ds$ near 0. But then $\dot{u} = 1$, so that $b_0 < +\infty$ and $u(t)$ extends beyond $t = b_0$, as does $z(t)$, contrary to assumption. Hence there exists an arbitrarily small positive σ such that $z(t)$ enters and leaves the disc $D(0,\sigma) = \{z \in \mathbb{C} : |z| < \sigma\}$ infinitely often as $t \to b_0$. Because σ is small and $f(0) = i$, there exists $\tau > 0$ such that any trajectory which meets $D(0,\sigma)$ crosses the real interval $I = (-2\sigma, 2\sigma)$ non-tangentially from below to above in increasing time, and exits $D(0,2\sigma)$ after leaving I, taking at least time τ to do so: thus $b_0 = +\infty$.

It is now possible to choose a sequence (t_n), with $a_0 < t_n < t_{n+1} < \infty$, such that $z(t_n)$ and $z(t_{n+1})$ both lie in $I \setminus \{0\}$ but $z(t) \notin I$ for $t_n < t < t_{n+1}$. Then $t_{n+1} \geq t_n + \tau$, so $t_n \to \infty$ and $\liminf_{n\to\infty} |z(t_n)| = 0$. Since the trajectory is non-periodic, $z(t_n) \neq z(t_{n+1})$. Let J_n be the open real interval with end-points $z(t_n)$ and $z(t_{n+1})$, let K_n be the arc $\{z(t) : t_n \leq t \leq t_{n+1}\}$, and let L_n be the Jordan curve formed from J_n and K_n.

Let P_n be the component of $I \setminus \{z(t_n)\}$ containing $z(t_{n+1})$, and Q_n the component of $I \setminus \{z(t_{n+1})\}$ containing $z(t_n)$. Choose u_n and v_n with $u_n - t_{n+1}$ and $t_n - v_n$ small and positive. Then $z(u_n)$ lies in a component $\Omega_{1,n}$ of $(\mathbb{C} \cup \{\infty\}) \setminus L_n$, as do all points lying just above the open interval P_n. Similarly, $z(v_n)$ and all points lying just below Q_n belong to the same component $\Omega_{2,n}$ of $(\mathbb{C} \cup \{\infty\}) \setminus L_n$. The fact that $J_n = P_n \cap Q_n$ gives $\Omega_{1,n} \neq \Omega_{2,n}$. All points $z(t)$ with $t > t_{n+1}$ also lie in $\Omega_{1,n}$, because $z(t)$ cannot meet K_n for $t > t_{n+1}$ and cannot cross J_n from above as t increases. This gives $z(t_m) \notin Q_n$ for all $m > n+1$, because the contrary case leads to $z(v_m) \in \Omega_{2,n}$. It follows that the sequence $z(t_n)$ is monotone, with $z(t_n) \to 0$ as $n \to \infty$.

Now the integral of $1/f(z)$ along K_n is real, by (6.2), but that along J_n is not, and so choosing n large enough makes

$$I_n = \mathrm{Im}\left(\int_{L_n} \frac{1}{f(z)}\,dz\right)$$

arbitrarily small but non-zero. Thus the lemma is proved if f has finitely many zeros in D, or if the trajectory $z(t)$ remains within a compact subset of D, because in these cases there are only finitely many zeros of f which may lie inside L_n, and so only finitely many possible values of I_n, by the residue theorem.

Assume now that f has infinitely many zeros, and hence finitely many poles, in D. Since $f(z_0) \neq 0, \infty$ by assumption, it may be assumed further that the trajectory $z(t)$ does not remain within any compact subset of D as $t \to +\infty$. Thus $\Omega_{2,n}$ must be the bounded component of $(\mathbb{C} \cup \{\infty\}) \setminus L_n$, and these bounded components satisfy $\Omega_{2,n} \subseteq \Omega_{2,n+1}$. Let Λ_n be the domain obtained by deleting from $\Omega_{2,n+1}$ all points in the closure of $\Omega_{2,n}$, and let $n_0 \in \mathbb{N}$ be so large that for $n \geq n_0$ there are no poles of f in Λ_n. Let $n \geq n_0$ and let Σ_n be the set of $w \in \Lambda_n$ with $f(w) \neq 0$. For $w \in \Sigma_n$, follow the trajectory ζ_w through w in decreasing time. The resulting path σ_w cannot exit $\Omega_{2,n+1}$, and so remains within a compact subset of D. Thus by the argument of the previous paragraph, with time reversed, σ_w must either cross J_n, or be periodic, or tend to a zero of f in Λ_n. Here the set of $w \in \Sigma_n$ corresponding to each of these finitely many possibilities is open, by §6.4, §6.5 and §6.7, as well as Lemma 6.9.1. But there are points $w \in \Lambda_n$, close to the trajectory $z(t)$, for which σ_w does cross J_n, and so by connectedness the same is true for all $w \in \Sigma_n$. A similar argument shows that for every $w \in \Sigma_n$ the trajectory ζ_w exits $\Omega_{2,n+1}$ through J_{n+1} in increasing time. However, if v is a zero of f in Λ_n, then §6.5 and §6.7, together with Lemma 6.9.1, show that there exists $w \neq v$, close to v, such that either ζ_w is periodic or ζ_w tends to w in increasing or decreasing time. Hence f has no zeros in Λ_n. It follows that there are only finitely many possible values for I_n, and this is a contradiction. □

Suppose now that f is non-constant and meromorphic in $\mathbb{C}$ in (6.2), with finitely many poles. If $\gamma(t)$ is a simple trajectory for (6.2), with maximal interval of definition $(\alpha, \beta) \subseteq \mathbb{R}$, then it follows from Lemma 6.10.1, with $D = \mathbb{C}$, that the initial and final end-points $\gamma^- = \lim_{t \to \alpha+} \gamma(t) \in \mathbb{C} \cup \{\infty\}$ and $\gamma^+ = \lim_{t \to \beta-} \gamma(t) \in \mathbb{C} \cup \{\infty\}$ both exist, and may coincide.

If $\gamma^+ = z_0 \in \mathbb{C}$ and $f(z_0) \neq \infty$ then z_0 must be a sink or a multiple zero of f (see §6.6), and the trajectory takes infinite time to reach z_0 (that is, $\beta = +\infty$). To see this, take $C > 0$ and $m \in \mathbb{N}$ with $|f(z)| \leq C|z - z_0|^m$ as $z \to z_0$. Let n be large and consider any $z(t)$ such that $|z(t_n) - z_0| = 2^{-n}$ and $|z(t_{n+1}) - z_0| = 2^{-n-1}$ and $2^{-n-1} \leq |z(t) - z_0| \leq 2^{-n}$ for $t_n \leq t \leq t_{n+1}$. This yields

$$2^{-n-1} \leq |z(t_{n+1}) - z(t_n)| = \left| \int_{t_n}^{t_{n+1}} f(z(t))\, dt \right| \leq (t_{n+1} - t_n) C 2^{-nm}$$

and so $t_{n+1} - t_n \geq C^{-1} 2^{(m-1)n-1} \geq 1/2C$. Similar remarks apply if $\gamma^- \in \mathbb{C}$.

6.11 The analytic implicit function theorem

Lemma 6.11.1 *Let the function $P(w, z)$ be C^1 on a neighbourhood of $(w_0, z_0) \in \mathbb{C}^2$ and satisfy the following: for each w near to w_0, the functions $P(w, z)$ and $P_w(w, z)$ are analytic functions of z on a neighbourhood of z_0; for each z near to z_0, the function $q(w) = P(w, z)$ is an analytic function of w on a neighbourhood of w_0.*

Assume that $P(w_0, z_0) = 0$, and that $P_w(w_0, z_0) \neq 0$. Then there exists an analytic function $\phi(z)$ on a neighbourhood of z_0, with $\phi(z_0) = w_0$, such that $P(\phi(z), z) = 0$ near z_0.

Proof. It may be assumed that $w_0 = z_0 = 0$. The function $g(w) = P(w, 0)$ is analytic near 0 with $g(0) = 0$ and $g'(0) = P_w(0, 0) \neq 0$. Thus g has a simple zero at 0 and, if ε is small and positive,

$$\frac{1}{2\pi i} \int_{|w| = \varepsilon} \frac{g'(w)}{g(w)}\, dw = 1,$$

with all integrations once counter-clockwise. In particular, $g(w) = P(w, 0) \neq 0$ for $|w| = \varepsilon$. Hence if $|z|$ is small enough then $P(w, z) \neq 0$ for $|w| = \varepsilon$, since P is C^1, and

$$\frac{1}{2\pi i} \int_{|w| = \varepsilon} \frac{P_w(w, z)}{P(w, z)}\, dw = 1,$$

by continuity of the integral and the argument principle applied to $q(w) = P(w, z)$. Thus, again if $|z|$ is small enough, the equation $P(w, z) = 0$ has a unique root $w = \phi(z) \in D(0, \varepsilon)$, and the residue theorem gives

$$\phi(z) = \frac{1}{2\pi i} \int_{|w| = \varepsilon} \frac{w P_w(w, z)}{P(w, z)}\, dw,$$

so that $\phi(z)$ is analytic near 0.

Chapter 7

Univalent functions and the hyperbolic metric

7.1 Basic results on univalent functions

7.1.1 The area theorem

Let $g(z) = 1/z + \sum_{n=1}^{\infty} b_n z^n$ be analytic and univalent in $0 < |z| < 1$. Then

$$\sum_{n=1}^{\infty} n|b_n|^2 \leq 1. \tag{7.1}$$

Proof. We can assume that $b_1 = a$ is real and non-negative, because if g is univalent and $|\alpha| = 1$ then $\alpha g(z\alpha) = 1/z + \alpha^2 b_1 z + \ldots$ is also univalent, and this does not change $|b_n|$. Our assumptions imply that the power series has radius of convergence at least 1. We extend g to a one-one meromorphic function on $D(0,1)$ by setting $g(0) = \infty$. If $0 < r < 1$ then

$$J_r(t) = g(re^{it}), \quad 0 \leq t \leq 2\pi,$$

is a simple closed curve. For finite w not on J_r, the winding number satisfies

$$n(J_r, w) = \frac{1}{2\pi i} \int_{J_r} \frac{1}{u - w} du = \frac{1}{2\pi i} \int_{|z|=r} \frac{g'(z)}{g(z) - w} dz$$

and by the argument principle this is zero or non-zero, depending on whether or not g takes the value w in $0 < |z| < r$.

For $0 < r < 1$ let $A(r)$ be the area of the set of finite complex values not taken by g in $D(0,r)$: this is the same as the area enclosed by J_r. For $0 < s < S < 1$ we have

$$A(s) - A(S) = \int_{s \leq |z| \leq S} |g'(z)|^2 r dr d\theta$$

because the integral on the RHS (computed using polar coordinates) is the area of the image under g of $s \leq |z| \leq S$. Differentiating gives

$$A'(r) = -\int_{|z|=r} |g'(z)|^2 r d\theta.$$

We write

$$g'(z) = -z^{-2} + \sum_{n=1}^{\infty} n b_n z^{n-1}, \quad \overline{g'(z)} = -(\overline{z})^{-2} + \sum_{n=1}^{\infty} n\overline{b_n}(\overline{z})^{n-1},$$

and use the elementary fact that, for $j, k \in \mathbb{Z}$,

$$\int_{|z|=r} z^j \overline{z}^k d\theta$$

is 0 unless $j = k$, in which case the integral is $2\pi r^{2j}$. Thus we get

$$-A'(r) = 2\pi \left(r^{-3} + \sum_{n=1}^{\infty} n^2 |b_n|^2 r^{2n-1} \right),$$

and by integration there is a constant C such that

$$A(r) = C + \pi r^{-2} - \pi \sum_{n=1}^{\infty} n|b_n|^2 r^{2n}, \quad 0 < r < 1.$$

We assert that $C = 0$. Suppose first that $g(z) = 1/z + az$, still with $a > 0$. With this assumption,

$$g(re^{i\theta}) = (1/r + ar)\cos\theta - i(1/r - ar)\sin\theta$$

describes an ellipse E_r enclosing an area

$$\pi(1/r + ar)(1/r - ar) = \pi(r^{-2} - |b_1|^2 r^2) = \pi r^{-2} + O(r^2).$$

In the general case, as $|z| = r \to 0$ we have

$$g(z) = 1/z + b_1 z + O(r^2)$$

and so the distance from J_r to the ellipse E_r is $O(r^2)$. Since E_r has length $O(1/r)$, the difference between $A(r)$ and the area enclosed by $E(r)$ is $O(r)$ as $r \to 0$, and this gives $A(r) = \pi r^{-2} + O(r)$ and so $C = 0$. Using the fact that $A(r) \geq 0$, and letting $r \to 1$, we deduce the lemma.

7.1.2 The class S

Suppose that h is analytic and univalent in $D(0, 1)$. Then $h'(0) \neq 0$ and the function

$$H(z) = \frac{h(z) - h(0)}{h'(0)}, \quad H(0) = 0, \quad H'(0) = 1,$$

is also analytic and univalent in $D(0, 1)$. This normalization gives us the class S of functions

$$f(z) = z + a_2 z^2 + a_3 z^3 + \ldots = z + \sum_{n=2}^{\infty} a_n z^n$$

which are analytic and univalent in $D(0, 1)$.

7.1.3 Bieberbach's theorem

Let $f \in S$. Then $|a_2| \leq 2$. Further, equality holds if and only if f is a Koebe function

$$f(z) = k_\theta(z) = \frac{z}{(1 - ze^{i\theta})^2} = z + 2z^2 e^{i\theta} + \ldots, \tag{7.2}$$

for some real θ.

Proof. Take $f \in S$, and write

$$f(z^2) = z^2(1 + a_2 z^2 + a_3 z^4 + \ldots) = z^2 G(z)$$

so that, since $G(z) \neq 0$ in $D(0,1)$, the function F given by

$$F(z) = zG(z)^{1/2} = z(1 + \frac{1}{2} a_2 z^2 + \ldots)$$

is analytic in $D(0,1)$. We claim that F is univalent on $D(0,1)$. To see this, suppose that $F(u) = \pm F(v)$. Then $f(u^2) = F(u)^2 = F(v)^2 = f(v^2)$ and so $u^2 = v^2, u = \pm v$. But $v = -u \neq 0$ gives $F(v) = -F(u) \neq 0$, since the power series for F has only odd powers, and so $F(u) = F(v)$ forces $u = v$.

Now we know that F is univalent on $D(0,1)$, we consider

$$g(z) = \frac{1}{F(z)} = \frac{1}{z} - \frac{1}{2} a_2 z + \ldots = \frac{1}{z} + \sum_{n=1}^{\infty} b_n z^n,$$

which is analytic and univalent on $0 < |z| < 1$. From (7.1) we get $\sum_{n=1}^{\infty} n|b_n|^2 \leq 1$ and so in particular $|b_1| = \frac{1}{2}|a_2| \leq 1$. If $|a_2| = 2$ then we must have $|b_1| = 1$ and $b_n = 0$ for $n \geq 2$ and so, for some real θ,

$$g(z) = \frac{1}{z} - ze^{i\theta} = \frac{1 - z^2 e^{i\theta}}{z} \quad , \quad F(z) = \frac{z}{1 - z^2 e^{i\theta}} \quad , \quad f(z) = F(z^{1/2})^2 = \frac{z}{(1 - ze^{i\theta})^2}.$$

7.1.4 Koebe quarter theorem

Suppose that $f \in S$ and that f does not take the finite value w in $D(0,1)$. Then $|w| \geq 1/4$. If $|w| = 1/4$ then f is given by (7.2), with $w = -\frac{1}{4}e^{-i\theta}$ for some real θ.

Proof. Assume $f(z) \neq w$. Then

$$\frac{wf}{w - f} = -w + \frac{w^2}{w - f} = z + (a_2 + 1/w)z^2 + \ldots$$

is also in S. This gives, by Bieberbach's theorem,

$$|a_2 + 1/w| \leq 2, \quad |1/w| \leq 2 + |a_2| \leq 4.$$

Also if $|1/w| = 4$ then $1/w = -4e^{i\theta}$ for some real θ. Since $|a_2| \leq 2$ and $|a_2 + 1/w| \leq 2$ we must have $a_2 = 2e^{i\theta}$ and so f is given by (7.2).

Note that

$$k_\theta(z) = e^{-i\theta} k_0(ze^{i\theta}).$$

Also

$$k_0(z) = \frac{z}{(1-z)^2} = \frac{1}{4}\left(\frac{1+z}{1-z}\right)^2 - \frac{1}{4},$$

and this maps $D(0,1)$ univalently onto the region obtained by deleting from the complex plane the half-line $\{w \in \mathbb{R} : w = x \leq -\frac{1}{4}\}$.

7.1.5 The distance to the boundary

Let $f \in S$. By the Koebe quarter theorem we know that the distance from 0 to the boundary of $f(D(0,1))$ is at least $\frac{1}{4}$. On the other hand, this distance is at most 1, for otherwise the inverse function F is defined and analytic on a disk $D(0,R)$ with $R > 1$, and Schwarz' lemma applied to $h(z) = F(Rz)$ gives $1 = |1/f'(0)| = |F'(0)| \le 1/R$.

Suppose now that a is any point in $D(0,1)$, and that g is analytic and univalent on $D(0,1)$. Set

$$G(z) = g\left(\frac{z+a}{1+\overline{a}z}\right).$$

Then $G'(0) = (1-|a|^2)g'(a)$ and

$$H(z) = \frac{G(z) - G(0)}{(1-|a|^2)g'(a)}$$

is in S. Thus the distance from 0 to the boundary of $H(D(0,1))$ is at least $\frac{1}{4}$ and at most 1. This gives

$$\frac{1}{4}(1-|a|^2)|g'(a)| \le \operatorname{dist}\{g(a), \partial(g(D(0,1)))\} \le (1-|a|^2)|g'(a)|. \tag{7.3}$$

7.1.6 Koebe distortion theorem

Let $f \in S$. Then for $|z_0| = r < 1$ we have

$$\frac{1-r}{(1+r)^3} \le |f'(z_0)| \le \frac{1+r}{(1-r)^3}. \tag{7.4}$$

Proof. Set

$$g(z) = f\left(\frac{z+z_0}{1+\overline{z_0}z}\right) = b_0 + b_1 z + b_2 z^2 + \dots.$$

Then g is analytic and univalent in $D(0,1)$ and

$$b_0 = f(z_0), \quad b_1 = g'(0) = f'(z_0)(1-r^2), \quad 2b_2 = g''(0) = (1-r^2)^2 f''(z_0) - 2\overline{z_0}(1-r^2)f'(z_0).$$

Applying Bieberbach's theorem to $(g(z) - g(0))/g'(0)$ we get $|g''(0)| \le 4|g'(0)|$, and so

$$f''(z_0)/f'(z_0) - 2\overline{z_0}(1-r^2)^{-1}$$

has modulus at most $4(1-r^2)^{-1}$. Multiplying through by $z_0 r^{-1}$ we get

$$\left|\frac{z_0 f''(z_0)}{r f'(z_0)} - \frac{2r}{1-r^2}\right| \le \frac{4}{1-r^2}. \tag{7.5}$$

If we write $G = \log f'(z), \zeta = \log z, \rho = |z|$ then the Cauchy-Riemann equations give

$$\frac{\partial \log|f'(z)|}{\partial \log \rho} = \operatorname{Re}\left(\frac{dG}{d\zeta}\right) = \operatorname{Re}\left(\frac{zf''(z)}{f'(z)}\right).$$

Thus

$$\frac{\partial \log|f'(z)|}{\partial \rho} = \rho^{-1}\operatorname{Re}\left(\frac{zf''(z)}{f'(z)}\right)$$

and so (7.5) tells us that

$$\frac{2\rho - 4}{1-\rho^2} = \frac{2\rho}{1-\rho^2} - \frac{4}{1-\rho^2} \le \frac{\partial \log|f'(z)|}{\partial \rho} \le \frac{2\rho}{1-\rho^2} + \frac{4}{1-\rho^2} = \frac{2\rho+4}{1-\rho^2}.$$

Integrating from 0 to r with respect to ρ using partial fractions, and then taking exponentials, we get (7.4).

7.2 The hyperbolic metric

We begin with a refinement of the standard Schwarz lemma.

7.2.1 The Schwarz-Pick lemma

Let $f : D(0,1) \to D(0,1)$ be analytic, and let $a \in D(0,1)$. Then we have

$$\left| \frac{f(z) - f(a)}{1 - \overline{f(a)} f(z)} \right| \le \left| \frac{z-a}{1 - \overline{a} z} \right| \tag{7.6}$$

and

$$\frac{|f'(z)|}{1 - |f(z)|^2} \le \frac{1}{1 - |z|^2} \tag{7.7}$$

for all z in $D(0,1)$. If there exists z in $D(0,1)$ for which equality holds in (7.7), or $z \in D(0,1) \setminus \{a\}$ for which equality holds in (7.6), then f is a conformal map of (i.e. a one-one analytic function from) the unit disc $D(0,1)$ onto itself.

The conformal maps f of $D(0,1)$ onto itself have the form

$$f(z) = e^{i\theta} \frac{z-a}{1-\overline{a}z} \tag{7.8}$$

for some constants θ, a with θ real and $|a| < 1$. For such f, equality holds in both (7.6) and (7.7).

Proof. It is easy to check that f of the form (7.8) is a conformal map of $D(0,1)$ onto itself: f is Möbius and so one-one, and $f(a) = 0$, and $|f(z)| = 1$ for $|z| = 1$. We denote the collection of mappings of form (7.8) by A. It is easy to check that A is a group under composition.

Next let f map $D(0,1)$ analytically into itself, and let $a \in D(0,1)$, and define G by

$$G(z) = G_1(z)G_2(z), \quad G_1(z) = \frac{f(z) - f(a)}{1 - \overline{f(a)} f(z)} \quad , \quad G_2(z) = \frac{1 - \overline{a}z}{z-a}.$$

Then G has a removable singularity at a and so is analytic in $D(0,1)$. Further, we have $|G_1(z)| \le 1$ on $D(0,1)$, while $|G_2(z)| \to 1$ as $|z| \to 1$. So the maximum principle gives $|G(z)| \le 1$ on $D(0,1)$.

There are now two possibilities. The first is that G is a constant of modulus 1, so that equality holds in (7.6). Further, we can solve for f, and since A is a group it follows that f is in A, and is a conformal map of $D(0,1)$ onto itself. Finally, since

$$|f'(a)| = \lim_{z \to a} \left| \frac{f(z) - f(a)}{z-a} \right|, \tag{7.9}$$

we get equality in (7.7).

In the converse direction, suppose that f is a conformal map of $D(0,1)$ onto itself. Then $|f(z)| \to 1$ as $|z| \to 1$, and so $|G(z)| \to 1$ as $|z| \to 1$. Since f is one-one, $G(z)$ is non-zero on $D(0,1)$ (the zeros cancel out) and so $|G(z)| = 1$ on $D(0,1)$, by the maximum principle applied to G and $1/G$. It follows that $f \in A$ and that equality holds in (7.6) and (7.7).

Finally, suppose that f is not a conformal map of $D(0,1)$ onto itself, and take $a \in D(0,1)$. Then $|G(z)| < 1$ for all z in $D(0,1)$, so that we have strict inequality in (7.6), for $z \ne a$. Further, a lies in a compact set K_a on which $|G(z)| \le k_a < 1$, and (7.9) gives us (7.7) for $z = a$, with strict inequality. Since a is arbitrary the proof is complete.

7.2.2 Lemma

Let γ be a smooth contour joining a to b, with $|a| \leq |b|$, and let $f(z) = g(|z|) \geq 0$ be a function of $|z|$ which is upper semi-continuous on γ. Then

$$\int_\gamma f(z)|dz| \geq \int_{|a|}^{|b|} g(t)dt. \tag{7.10}$$

Proof. Suppose first that g is continuous. Take $\delta > 0$ and a partition $|a| = x_0 < x_1 < \ldots < x_n = |b|$ such that

$$\max\{g(t) : x_{j-1} \leq t \leq x_j\} - \delta < m_j = \min\{g(t) : x_{j-1} \leq t \leq x_j\}$$

for each j. Then for each j there is a sub-path γ_j of length at least $x_j - x_{j-1}$ and lying in $x_{j-1} \leq |z| \leq x_j$, on which $f(z) \geq m_j$. Thus

$$\int_\gamma f(z)|dz| \geq \sum_{j=1}^n m_j(x_j - x_{j-1}) \geq \int_{|a|}^{|b|} (g(t) - \delta)dt.$$

This proves (7.10) when g is continuous. In the general case take continuous $g_n \downarrow g$ so that

$$\int_\gamma g_n(|z|)|dz| = \int_A^B g_n(|\gamma(s)|)|\gamma'(s)|ds \to \int_A^B g(|\gamma(s)|)|\gamma'(s)|ds = \int_\gamma g(|z|)|dz|,$$

by the monotone convergence theorem applied to $g_1 - g_n$, and

$$\int_\gamma g_n(|z|)|dz| \geq \int_{|a|}^{|b|} g_n(t)dt \geq \int_{|a|}^{|b|} g(t)dt.$$

7.2.3 The hyperbolic metric in the disc

Let γ be a piecewise smooth contour in the unit disc $D(0,1)$. The hyperbolic (non-Euclidean) length of γ is defined to be

$$L_\gamma = \int_\gamma \frac{2|dz|}{1 - |z|^2},$$

in which $|dz|$ indicates that the integration is with respect to arc length (sometimes the factor 2 is omitted).

If f is a conformal map of $D(0,1)$ onto itself then the hyperbolic length of $f(\gamma)$ is

$$\int_{f(\gamma)} \frac{2|dw|}{1 - |w|^2} = \int_\gamma \frac{2|f'(z)||dz|}{1 - |f(z)|^2} = \int_\gamma \frac{2|dz|}{1 - |z|^2} = L_\gamma$$

using the fact that we have equality in (7.7). Thus the hyperbolic length is invariant under f.

Now suppose that γ joins 0 to $r \in (0,1)$. Then Lemma 7.2.2 gives

$$L_\gamma \geq \int_0^r \frac{2dx}{1 - x^2} = \log\left(\frac{1+r}{1-r}\right).$$

In particular, the shortest path (in terms of hyperbolic length) from 0 to r is the straight line segment.

If z_1, z_2 are in $D(0,1)$ we now define the hyperbolic distance $[z_1, z_2]$ to be the infimum of L_γ over all piecewise smooth contours γ joining z_1 to z_2 through $D(0,1)$. The distance is not altered if we apply a conformal map f of $D(0,1)$ onto itself. We can choose f so that $f(z_1) = 0, f(z_2) = r > 0$, and the shortest path between these two points is then the straight line S from 0 to r. Hence the shortest path from z_1 to z_2 is the arc $f^{-1}(S)$, which is either a straight line through 0 or (since f is Möbius) a circular arc which meets the circle $|z| = 1$ at right-angles. In particular

$$[0,r] = \log\left(\frac{1+r}{1-r}\right).$$

7.2.4 The hyperbolic metric on a simply connected domain

Let D be a simply connected domain in the complex plane, not the whole plane. Then by the Riemann mapping theorem, there exists an analytic function H mapping D one-one onto $D(0,1)$. We can thus define the hyperbolic distance between w_1, w_2 in D to be the hyperbolic distance between $H(w_1), H(w_2)$ in $D(0,1)$. This does not depend on which H we choose, because if G is another conformal map of D onto $D(0,1)$ then $H \circ G^{-1}$ is a conformal map of $D(0,1)$ onto itself, so that $[H(w_1), H(w_2)] = [G(w_1), G(w_2)]$.

The next lemma gives a useful estimate for the hyperbolic metric on a simply connected domain. It is related in style and applicability to §15.1.6.

7.2.5 Lemma

Let D be a simply connected domain in the finite plane, not containing the origin, and let $w_1, w_2 \in D$. For $t > 0$ let $t\theta(t)$ be the length of the longest open arc of the intersection of D and the circle $|w| = t$. Then

$$[w_1, w_2]_D \geq \int_{|w_1|}^{|w_2|} \frac{dt}{t\theta(t)}. \tag{7.11}$$

Proof. Let h map $D(0,1)$ analytically and univalently onto D, with $h(z_j) = w_j$. Let γ be the hyperbolic geodesic (shortest path with respect to the hyperbolic metric) from z_1 to z_2. Then, with $\Gamma = h(\gamma)$, the estimate (7.3) gives

$$[w_1, w_2]_D = [z_1, z_2] = \int_\gamma \frac{2|dz|}{1-|z|^2} = \int_\Gamma \frac{2|dw|}{|h'(z)|(1-|z|^2)} \geq \int_\Gamma \frac{|dw|}{2\mathrm{dist}\{w, \partial D\}} \geq \int_\Gamma \frac{|dw|}{|w|\theta(|w|)}$$

since w can be joined to a point of ∂D by a circular arc of length at most $|w|\theta(|w|)/2$. Since D is open, $1/t\theta(t)$ is upper semi-continuous (see §15.1.3) and Lemma 7.2.2 now gives (7.11).

Chapter 8

Harmonic and subharmonic functions

8.1 Harmonic functions and Poisson's formula

8.1.1 Lemma

If the real-valued function $u(x, y)$ has continuous first and second partial derivatives on a domain D in $\mathbb{R}^2$ then $u_{xy} = u_{yx}$.

To prove this, take any closed rectangle $I = [a, b] \times [c, d]$ and use Fubini's theorem to get

$$\int_I u_{xy}\, dx\, dy = \int_a^b \int_c^d u_{xy}\, dy\, dx = \int_a^b u_x(x, d) - u_x(x, c)\, dx = u(b, d) - u(a, d) - u(b, c) + u(a, c).$$

But

$$\int_I u_{yx}\, dx\, dy = \int_c^d \int_a^b u_{yx}\, dx\, dy = \int_c^d u_y(b, y) - u_y(a, y)\, dy = u(b, d) - u(b, c) - u(a, d) + u(a, c).$$

Since the integrals are always the same the functions must agree: if not then by continuity we have without loss of generality $u_{xy} > u_{yx}$ on some rectangle.

8.1.2 Harmonic functions

Let D be a domain in $\mathbb{C}$ (or $\mathbb{R}^2$: we shall use these interchangeably). A function $u : D \to \mathbb{R}$ is called harmonic if u has continuous first and second partial derivatives and satisfies Laplace's equation

$$\Delta u = \nabla^2 u = u_{xx} + u_{yy} = 0.$$

By the Cauchy-Riemann equations, if $f = u + iv$ (u, v real) is analytic then u, v are harmonic.

Also, if $u(z) = u(x, y)$ is harmonic then Lemma 8.1.1 shows that $f = u_x - iu_y$ is analytic. If, in addition, D is simply connected then

$$F = u(a) + \int_a^z f(w)\, dw = U + iV$$

is analytic on D, and $f = F' = U_x + iV_x$ so $u_x = U_x$. Also $U_y = -V_x = u_y$. Thus $U = u$ and V is called a harmonic conjugate of u.

Note that if u is harmonic and h is analytic then the composition $u \circ h$ is (locally) the real part of an analytic function and so harmonic.

8.1.3 Identity theorem for harmonic functions

Suppose that u is harmonic on the domain D in $\mathbb{C}$ and constant on a non-empty subdomain G of D. Then u is constant on D.

Proof. The function $u_x - iu_y$ is analytic on D and 0 on G and so 0 on D.

8.2 Boundary behaviour of harmonic functions

8.2.1 Example

The following example shows that a bounded harmonic function need not have limits at every boundary point. Let $D = D(0,1)$ and define u on D by $u(z) = \arg\left(\frac{1+z}{1-z}\right)$. Then for $|w| = 1, \mathrm{Im}(w) > 0$ we have $\lim_{z\to w, z\in D} u(z) = \pi/2$ and for $|w| = 1, \mathrm{Im}(w) < 0$ we have $\lim_{z\to w, z\in D} u(z) = -\pi/2$. For $w = \pm 1$, the limit $\lim_{z\to w, z\in D} u(z)$ does not exist, although $u(x) = 0$ for real x.

8.2.2 Poisson's formula

Let $F(w)$ be a measurable function defined on $|w| = 1$ and taking values in $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, \infty\}$. For $|z| < 1$ and $|w| = 1$ set

$$K(z,w) = \mathrm{Re}\left(\frac{w+z}{w-z}\right) = \frac{1-|z|^2}{|w-z|^2}$$

(the Poisson kernel) and

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} K(z, e^{it})F(e^{it})\,dt.$$

The function u has the following properties.

(i) If $F(e^{it}) \in L^1([0,2\pi])$ (i.e. $\int_0^{2\pi} |F(e^{it})|\,dt < \infty$) then u is harmonic in $|z| < 1$.

(ii) If $|F(e^{it})| \leq M_0 < \infty$ for all t in $[0,2\pi]$ then $|u(z)| \leq M_0$ on $|z| < 1$.

(iii) If $F(e^{it}) \in L^1([0,2\pi])$ and $|w| = 1$ and F is finite and continuous at w then as $z \to w$ we have $u(z) \to F(w)$.

(iv) If F is non-negative and $F(e^{it})$ is not in $L^1([0,2\pi])$ then $u(z) \equiv \infty$.

Proof. Suppose first that $F \in L^1$. Set

$$Q(z,F) = \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{it}+z}{e^{it}-z}F(e^{it})\,dt = \frac{1}{2\pi}\int_0^{2\pi} -F(e^{it})\,dt + \frac{1}{2\pi}\int_0^{2\pi} \frac{2e^{it}}{e^{it}-z}F(e^{it})\,dt =$$

$$= \frac{1}{2\pi}\int_0^{2\pi} -F(e^{it})\,dt + 2M(z),$$

where

$$M(z) = \frac{1}{2\pi i}\int_{|w|=1} \frac{1}{w-z}F(w)\,dw.$$

The integral $M(z)$ exists because, for fixed z, the term $(w-z)^{-1}$ is bounded. Now, as $h \to 0$,

$$\frac{M(z+h)-M(z)}{h} = \frac{1}{2\pi i}\int_{|w|=1} \frac{1}{(w-z)(w-z-h)} F(w)\,dw \to \frac{1}{2\pi i}\int_{|w|=1} \frac{1}{(w-z)^2} F(w)\,dw,$$

by the dominated convergence theorem since, for fixed z and small h, the term $(w-z)^{-1}(w-z-h)^{-1}$ is uniformly bounded. Thus $Q(z,F)$ is analytic on $|z|<1$ and $u(z) = \mathrm{Re}(Q(z,F))$ is harmonic. This proves (i). Note that if we choose $F=1$ then $u(z) = Q(z,F) \equiv 1$ on $|z|<1$ by the residue theorem. Since the Poisson kernel is positive, this proves (ii).

Next we prove (iii). Assume that $F(v)$ is finite, $|v|=1$ and F is continuous at v. Take $\varepsilon > 0$ and choose $\delta > 0$ so that $|F(e^{it}) - F(v)| < \varepsilon/2$ for all $t \in T_0 = \{s \in [0,2\pi] : |e^{is} - v| < \delta\}$. Then

$$u(z) - F(v) = \frac{1}{2\pi}\int_0^{2\pi} K(z,e^{it})(F(e^{it}) - F(v))\,dt.$$

Now, since $K \geq 0$,

$$\frac{1}{2\pi}\int_{T_0} K(z,e^{it})(F(e^{it}) - F(v))\,dt$$

has modulus at most

$$(\varepsilon/2)\frac{1}{2\pi}\int_{T_0} K(z,e^{it})\,dt \leq (\varepsilon/2)\frac{1}{2\pi}\int_{[0,2\pi]} K(z,e^{it})\,dt = (\varepsilon/2).$$

On the other hand

$$\frac{1}{2\pi}\int_{[0,2\pi]\setminus T_0} K(z,e^{it})(F(e^{it}) - F(v))\,dt$$

has modulus at most

$$\left(\frac{1}{2\pi}\int_{[0,2\pi]} |F(e^{it})|\,dt + |F(v)|\right) \sup\{K(z,e^{it}) : t \in [0,2\pi] \setminus T_0\} \to 0$$

as $z \to v$, since for $t \notin T_0$ we have $|z - e^{it}| \geq \delta - |z-v|$ and so $K(z,e^{it}) \to 0$ as $z \to v$, uniformly on $[0,2\pi] \setminus T_0$.

Finally, to prove (iv) suppose that F is non-negative and $\int_{[0,2\pi]} F(e^{it})\,dt = \infty$. Since the Poisson kernel is positive and, for fixed z, bounded below on $[0,2\pi]$, we get $u \equiv \infty$ on $|z|<1$.

8.2.3 Corollary

Let F be continuous on the circle $|z - z_0| = R > 0$. Then, with $w = z_0 + Re^{it}$,

$$u(z) = P(z,F,z_0,R) = \frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - |z-z_0|^2}{|Re^{it} - (z-z_0)|^2} F(z_0 + Re^{it})\,dt = \frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - |z-z_0|^2}{|w-z|^2} F(w)\,dt$$

is harmonic on $D = D(z_0,R)$ and $u(z) \to F(w)$ as $z \to w \in \partial D$.

To prove this just put $z = z_0 + R\zeta$ and $u(z) = v(\zeta)$, where $v(\zeta)$ is the Poisson integral of $G(e^{it}) = F(w)$.

8.2.4 Maximum principle: first version

Suppose that u is harmonic on the bounded domain D in $\mathbb{C}$ and continuous on the closure of D, and that $u(z) \leq M$ on ∂D. Then $u(z) \leq M$ on D.

Proof. Suppose $u(z_0) > M$ for some z_0 in D. Then, if $t > 0$ is small enough, the function $v(z) = u(z) + t(x^2 + y^2)$ is such that $v(z_0) > \max\{v(z) : z \in \partial D\}$ and so v has a local maximum at some $z_1 \in D$. But at z_1 this gives $v_x = v_y = 0$ and $v_{xx} + v_{yy} = 4t > 0$ so that at least one of v_{xx}, v_{yy} must be positive. This is a contradiction.

We will subsequently see another way to prove this, via the mean value property.

8.2.5 The mean value property

Let D be a domain in $\mathbb{C}$. We say that a function $u : D \to \mathbb{C}$ has the mean value property if each z_0 in D has $r_0 > 0$ such that

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it})\, dt, \quad (0 < r \leq r_0).$$

Obviously these functions on D form a vector space.

If u is harmonic on D then u has the mean value property. To see this, take a disc $D(z_0, R) \subseteq D$ on which $u = \mathrm{Re}(f)$ with f analytic. Then Cauchy's integral formula

$$f(z_0) = \frac{1}{2\pi i} \int_{|z-z_0|=r} \frac{f(z)}{z - z_0}\, dz = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{it})\, dt, \quad (0 < r < R), \tag{8.1}$$

implies that f has the mean value property and so has u.

8.3 Subharmonic functions

Let D be a domain in $\mathbb{C}$. A function $u : D \to [-\infty, \infty)$ is subharmonic if:

(i) u is upper semi-continuous (upper semi-continuous) in D (see §1.5);

(ii) u has the sub-mean-value-property, that to each z_0 in D corresponds $r_0 > 0$ such that

$$u(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it})\, dt, \quad (0 < r \leq r_0).$$

The integral exists because u is measurable, by (i), and bounded above on the circle, again by (i).

8.3.1 Examples of subharmonic functions

Obviously harmonic functions are subharmonic.

Suppose that f is analytic on the domain D: set $u = \log|f|$. If $a \in D$ and $f(a) \neq 0$ then we can define a branch g of $\log f$ on a neighbourhood A of a and $u = \mathrm{Re}(g)$ is harmonic on A. If $f(a) = 0$ then $u(z) \to u(a) = -\infty$ as $z \to a$. Thus u is subharmonic on D.

It is easy to check from (8.1) that $|f(z)|$ is also subharmonic on D. Thus if $p > 0$ then $|f(z)|^p$ is also subharmonic (it is clearly upper semi-continuous and we need only check the sub-mean value property: this is obvious if $f(a) = 0$ and, if $f(a) \neq 0$, we write $|f(z)|^p = |f(z)^p|$ locally).

Further, if u, v are subharmonic then so are $u + v, \max\{u, v\}$, and so subharmonic functions are more "flexible" than analytic or harmonic functions.

Thus the maximum of a finite family of subharmonic functions is subharmonic. However the sup of an infinite family need not be: for example, let $u_n(z) = (1/n)\log|z|$. Then the sup is $-\infty$ at 0, and is 0 for $0 < |z| < 1$, and so is not upper semi-continuous.

8.3.2 Maximum principle: second version

Let D be a domain in $\mathbb{C}$ and let u be subharmonic on D. If u has a maximum in D then u is constant on D.

Proof. Assume that $u(z) \leq u(z_0) = M$ on D. If $M = -\infty$ then the result is obvious. Assume now that $M \in \mathbb{R}$. If $u(z_1) = M$ then since

$$M = u(z_1) \leq \frac{1}{2\pi}\int_0^{2\pi} u(z_1 + re^{it})\, dt \leq M$$

for small positive r we must have $u \equiv M$ on $|z - z_1| = r$, by the fact that u is upper semi-continuous. This is because if $u(z_1 + re^{is}) < M$ we get $u(z_1 + re^{it}) < M' < M$ for t close to s and this makes the integral less than M. So the set $\{z \in D : u(z) = M\}$ is non-empty and open. The set $\{z \in D : u(z) < M\}$ is open since u is upper semi-continuous. By connectedness, the second set must be empty.

8.3.3 Maximum principle: third version

Let D be a domain in $\mathbb{C}$ and define $\partial_\infty D$ to be the collection of all boundary points of D in $\mathbb{C}^*$, with respect to the spherical metric. Thus $\partial_\infty D$ is the finite boundary ∂D plus, if D is unbounded, the point ∞. Then $\partial_\infty D$ is compact in $\mathbb{C}^*$. If u is subharmonic in D and

$$\limsup_{z\to\zeta, z\in D} u(z) \leq M \in [-\infty, \infty)$$

for every $\zeta \in \partial_\infty D$, then either $u(z) \equiv M$ on D, or $u(z) < M$ for all z in D.

Proof. The first assertion is obvious since $\partial_\infty D$ is closed and $\mathbb{C}^*$ is compact. Set $L = \sup\{u(z) : z \in D\}$ and take $z_n \in D$ with $u(z_n) \to L$. Assume without loss of generality that z_n converges to the point z^* in $D \cup \partial_\infty D$. Now if $L > M$ then $z^* \in D$ and we get $u(z) \equiv L$ on D by Lemma 8.3.2, an obvious contradiction. So $L \leq M$. Furthermore, either $u < M$ on D or Lemma 8.3.2 gives $u \equiv M$ on D.

8.3.4 Lemma

Let u be subharmonic and bounded above on the domain D in $\mathbb{C}$. For $w \in \partial_\infty D$, set

$$\phi(w) = \limsup_{z\to w, z\in D} u(z).$$

Then the function $v(z)$ defined by $v(z) = u(z)$ if $z \in D$ and $v(z) = \phi(z)$ if $z \in \partial_\infty D$ is upper semi-continuous on $D \cup \partial_\infty D$.

Proof. We only need consider w on the boundary. Suppose $\phi(w) < s < t$. Then there is some spherical disc $D_q(w,r) = \{z \in \mathbb{C}^* : q(z,w) < r\}$ such that

$$u(z) < s, \quad (z \in D \cap D_q(w,r)).$$

But then if $x \in \partial_\infty D \cap D_q(w,r)$ we have $\phi(x) \leq s < t$. So $v(x) < t$ for all $x \in D \cup \partial_\infty D$ which are sufficiently close to w.

8.3.5 Theorem (comparison with a Poisson integral)

Let u be subharmonic on the disc $D(z_0,R)$. Let $v(w)$ be upper semi-continuous on $|w - z_0| = R$, taking values in $[-\infty,\infty)$, with

$$\limsup_{z\to w, z\in D(z_0,R)} u(z) \leq v(w), \quad (|w - z_0| = R).$$

Then for $z \in D(z_0,R)$ we have

$$u(z) \leq P(z,v) = P(z,v,z_0,R) = \frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - |z-z_0|^2}{|Re^{it} - (z-z_0)|^2} v(z_0 + Re^{it})\,dt. \tag{8.2}$$

If $|w - z_0| = R$ then

$$\limsup_{z\to w, z\in D(z_0,R)} P(z,v) \leq v(w), \tag{8.3}$$

and if $u \not\equiv -\infty$ on $D(z_0,R)$ then $P(z,v)$ is harmonic there.

If u is harmonic in $D(z_0,R)$ and continuous on $|z - z_0| \leq R$ then setting $v = u$ gives equality in (8.2), so that u is the Poisson integral of its boundary values.

Note that if u is subharmonic in a domain containing the set $\{z \in \mathbb{C} : |z - z_0| \leq R\}$ then we may take $v = u$, since u is upper semi-continuous. Further, since the circle $|z - z_0| = R$ is compact, v is bounded above there.

Proof. To prove the theorem take a sequence of continuous functions f_n on $|z - z_0| = R$, decreasing pointwise to v. Such a sequence exists by Theorem 1.5.1. Let $u_n(z) = P(z,f_n) = P(z,f_n,z_0,R)$ be as defined by (8.2). Then u_n is harmonic on $D(z_0,R)$ and $u_n(z) \to f_n(w)$ as $z \to w \in \partial D(z_0,R)$ with $z \in D(z_0,R)$.

Hence $u - u_n$ is subharmonic in $D(z_0,R)$ and since $v \leq f_n$ we have

$$\limsup_{z\to w, z\in D(z_0,R)} (u - u_n)(z) \leq 0, \quad (w \in \partial D(z_0,R)).$$

Thus the maximum principle 8.3.3 gives, for $z \in D(z_0,R)$,

$$u(z) \leq u_n(z) = P(z,f_n).$$

Let $M = \max\{f_1(w) : |w - z_0| = R\}$. Then $M - f_{n+1} \geq M - f_n \geq 0$ and the monotone convergence theorem gives

$$P(z, M - f_n) \to P(z, M - v), \quad P(z,f_n) \to P(z,v)$$

for every $z \in D(z_0,R)$. Thus $u(z) \leq P(z,v)$, which is (8.2).

Further, if $|w - z_0| = R$ and $z \to w$ with $z \in D(z_0,R)$ then

$$\limsup P(z,v) \leq \limsup P(z,f_n) = \limsup u_n(z) = f_n(w) \to v(w),$$

which gives (8.3).

To prove that $P(z,v)$ is harmonic if $u \not\equiv -\infty$ we can assume without loss of generality that $v \le 0$ (since v is bounded above on the compact set $|w-z_0| = R$). We then apply (after rescaling) the Poisson formula 8.2.2: if $\int_{[0,2\pi]} v(z_0 + Re^{it})\,dt = -\infty$ then $P(z,v) \equiv -\infty$ and $u(z) \equiv -\infty$ on $D(z_0,R)$. On the other hand if $v(z_0+Re^{it}) \in L^1([0,2\pi])$ then $P(z,v)$ is harmonic.

Finally, if u is harmonic on $D(z_0,R)$ and continuous on the closure we set $v = u$ and apply the above to u and $-u$ to get $u = P(z,u)$ on $D(z_0,R)$.

8.3.6 Corollary

Suppose that $u : D \to \mathbb{R}$ is continuous and has the mean value property 8.2.5 on the domain D in $\mathbb{C}$. Then u is harmonic on D.

Note that the hypotheses are equivalent to u and $-u$ both being subharmonic on D.

Proof. Take any disc $D(z_0,R)$ whose closure lies in D. Form the Poisson integral $\tilde{u}$ on $D(z_0,R)$ with boundary values $u(z_0+Re^{it})$. Then $\tilde{u}$ is harmonic on $D(z_0,R)$. Since u and $-u$ are subharmonic, Theorem 8.3.5 gives

$$u(z) \le \tilde{u}(z), \quad -u(z) \le -\tilde{u}(z), \quad (z \in D(z_0,R))$$

and so $u = \tilde{u}$ on $D(z_0,R)$.

The following result addresses the issue of on how large a set a non-constant subharmonic function can be $-\infty$.

8.3.7 Theorem

Let u be subharmonic on a domain D in $\mathbb{C}$ and let $0 < s < r$ and $D(z_0,r) \subseteq D$. Suppose that $u(z) \equiv -\infty$ on a subset of the circle $S(z_0,s)$ of positive angular measure. Then $u(z) \equiv -\infty$ on D.

In particular if $u \equiv -\infty$ on $D(z_0,r)$ then $u \equiv -\infty$ on D.

Proof. Since u is bounded above on $S(z_0,s)$ we get $\int_0^{2\pi} u(z_0+se^{i\theta})d\theta = -\infty$. Thus the Poisson integral of u is identically $-\infty$ on $D(z_0,s)$ and Theorem 8.3.5 gives $u(z) \equiv -\infty$ on $D(z_0,s)$.

Now let F be the set of $w \in D$ such that $u \equiv -\infty$ on a neighbourhood of w. Obviously F is open, and we will show that F is also closed (in D) so that the result follows by connectivity. Let $w_n \in F$ and $w_n \to w \in D$. Then for arbitrarily small t we have $u(z) \equiv -\infty$ on a subset of $S(w,t)$ of positive measure. Hence $u(z) \equiv -\infty$ on $D(w,t)$ and $w \in F$.

8.3.8 Lemma (Poisson modification of a subharmonic function)

Let u be subharmonic on the domain D in $\mathbb{C}$ and let the closure of $D(z_0,R)$ be contained in D. Define $U(z) = u(z)$ on $D \setminus D(z_0,R)$, and on $D(z_0,R)$ let U be the Poisson integral of $u(z_0+Re^{it})$ i.e.

$$U(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - |z-z_0|^2}{|Re^{it} - (z-z_0)|^2} u(z_0+Re^{it})\,dt.$$

Then U is subharmonic with $U \ge u$ on D. If $u \not\equiv -\infty$ on D then U is harmonic on $D(z_0,R)$.

Note that in particular this gives

$$u(z_0) \le \frac{1}{2\pi}\int_0^{2\pi} u(z_0+Re^{it})\,dt$$

for *every* $R > 0$ such that the closure of $D(z_0, R)$ is contained in D, and not just for $0 < r \le r_0$ as in the definition 8.3 of a subharmonic function.

Proof. We already know that $u \le U$, by Theorem 8.3.5. Thus we only need check that U is upper semi-continuous and has the sub-mean value property at all z_1 with $|z_1 - z_0| = R$. First, for small r,

$$U(z_1) = u(z_1) \le \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{it})\,dt \le \frac{1}{2\pi}\int_0^{2\pi} U(z_0 + re^{it})\,dt.$$

Next,

$$\limsup_{z\to z_1, z\in D(z_0,R)} U(z) \le u(z_1)$$

by Theorem 8.3.5 (see (8.3), while

$$\limsup_{z\to z_1, z\notin D(z_0,R)} U(z) = \limsup_{z\to z_1, z\notin D(z_0,R)} u(z) \le u(z_1)$$

since u is upper semi-continuous.

Hence U is subharmonic on D. Finally if U is not harmonic on $D(z_0, R)$ then $U \equiv -\infty$ there and the same is true of u on $D(z_0, R)$ and hence on D.

8.3.9 Harnack's inequality

Let u be harmonic and non-negative on $|z - z_0| \le R$. If $|z - z_0| = r < R$ then

$$\left(\frac{R-r}{R+r}\right) u(z_0) \le u(z) \le \left(\frac{R+r}{R-r}\right) u(z_0).$$

This follows at once from Poisson's formula.

8.3.10 Harnack's theorem

Let D be a domain in $\mathbb{C}$. Let u_n be harmonic functions on D with $u_1 \le u_2 \le u_3 \le \dots$. Let $v(z) = \lim_{n\to\infty} u_n(z)$. Then either $v \equiv \infty$ on D, or v is harmonic on D, in which case $u_n \to v$ locally uniformly on D.

Proof. Suppose first that $v(w) < \infty$. Take a disc $D(w, 4R) \subseteq D$. Then we assert that $u_n \to v$ uniformly on $D(w, 2R)$. Take $\delta > 0$. Then there exists N such that for all $n \ge m \ge N$ we have

$$0 \le u_n(w) - u_m(w) < \delta$$

and so Harnack's inequality applied on $|z - w| \le 3R$ gives

$$|u_n(z) - u_m(z)| = u_n(z) - u_m(z) < 5\delta$$

for all z in $D(w, 2R)$. Letting $n \to \infty$ we see that $v(z)$ is finite and $|u_m(z) - v(z)| \le 5\delta$ on $D(w, 2R)$. Hence $u_m \to v$ uniformly, and so v is continuous, on $D(w, 2R)$. Now on $D(w, R)$ we have, denoting the Poisson integral by $P(z, u)$,

$$v(z) = \lim u_n(z) = \lim P(z, u_n) = P(z, v)$$

by uniform convergence. Thus v is harmonic on $D(w, R)$.

Now suppose that $v(w) = \infty$. Fix m. For $M \in (0, \infty)$ we have $u_n(w) - u_m(w) > M$ for large n. This time Harnack's inequality gives $u_n(z) - u_m(z) > M/5$ on $D(w, 2R)$, and we have thus shown that $u_n \to \infty$ uniformly on $D(w, R)$.

The sets $\{w : v(w) < \infty\}$ and $\{w : v(w) = \infty\}$ are thus open, and by connectedness one of them is empty.

Chapter 9

Perron's method

9.1 The Dirichlet problem

Let D be a domain in $\mathbb{C}$ and let f be a bounded real-valued function on $X = \partial_\infty D$ (the boundary with respect to the extended plane). The Dirichlet problem is to find, if possible, a harmonic function $h = h_f$ on D such that

$$\lim_{z \to w, z \in D} h(z) = f(w) \quad \text{for every } w \text{ in } X. \tag{9.1}$$

When D is a disc, and f is continuous, this is achieved by means of the Poisson integral (see 8.2.2 and 8.2.3).

9.1.1 The Perron family and Perron function

Let D be a domain in $\mathbb{C}$ and let f be a bounded real function on $X = \partial_\infty D$. The *Perron family* $U(f)$ is the collection of all subharmonic functions u on D such that for every $w \in X = \partial_\infty D$ we have

$$\limsup_{z \to w, z \in D} u(z) \leq f(w).$$

The *Perron function* v_f is then defined by

$$v(z) = v_f(z) = \sup\{u(z) : u \in U(f)\}.$$

Obviously if $f \leq g$ then $v_f \leq v_g$.

If a function h satisfying (9.1) exists then $h = v_f$: to see this, note first that $h \in U(f)$, so that $h \leq v_f$. Further, for every $u \in U(f)$, we have $\limsup_{z \to w}(u(z) - h(z)) \leq 0$ for every $w \in X$, and so $u \leq h$ on D, by the maximum principle. Thus $v_f \leq h$. So if the Dirichlet problem is solvable, then the solution h_f equals the Perron function v_f. Most of this section will be concerned with the converse direction: that is, proving that if $f : X \to \mathbb{R}$ is continuous and X is sufficiently regular, then $h = v_f$ does indeed satisfy (9.1). However, we first look at an example.

9.1.2 Example

This example shows that the Dirichlet problem is not always solvable. Let $D_1 = D(0,1) \setminus \{0\}$ and let $f(x) = 0$ for $|x| = 1$, with $f(0) = 1$. Now let $v \in U(f)$ and set $u = \max\{v, 0\}$. Then $u \in U(f)$ and $0 \leq u(z) \leq 1$ on D_1 by the maximum principle. Thus for $0 < t < 1$ we get

$$u(z) \leq w(z) = \frac{\log 1/|z|}{\log 1/t}, \quad (t < |z| < 1).$$

This is because $u - w$ is subharmonic on $t < |z| < 1$ and at most 0 on the boundary. Fixing z and letting $t \to 0$ we see that $u(z) \equiv 0$. This implies that $v_f \equiv 0$. Hence the Dirichlet problem for f and D_1 cannot have a solution h, because if it did we would have $h = v_f \equiv 0$ by 9.1.1.

9.1.3 Lemma

Let f be a bounded real-valued function on X with $|f| \leq M$ on X, and let v_f be its Perron function. Then the following are true:

(i) we have $|v_f| \leq M$;

(ii) the function $v = v_f$ is harmonic on D.

Proof. First, $v \geq -M$ since $-M \in U(f)$. Further, each u in $U(f)$ has $\limsup_{z \to w, z \in D} u(z) \leq M$ and so $u \leq M$ by the maximum principle.

To prove that v is harmonic, we take a disc $D_1 = D(z_0, R)$ whose closure lies in D, and we make the following observations. First, the maximum of finitely many elements of $U(f)$ is subharmonic on D and is an element of $U(f)$. Second, if $u_0 \in U(f)$ then there exists an element U_0 of $U(f)$ which is harmonic on D_1 and satisfies $u_0 \leq U_0$ on D. To see this, let $U_0(z) = u_0(z)$ on $D \setminus D_1$, but for z in D_1 let $U_0(z)$ equal the Poisson integral $P(z, u_0)$, where

$$P(z, g) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - |z - z_0|^2}{|Re^{it} - (z - z_0)|^2} g(z_0 + Re^{it})\, dt.$$

Then by Lemma 8.3.8, U_0 is subharmonic with $U_0 \geq u_0$ on D. Also U_0 is equal to u_0 outside D_1, so that $U_0 \in U(f)$.

So we start by taking $v_n \in U(f)$ such that $v_n(z_0) \to v(z_0)$ and setting

$$u_n(z) = \max\{v_1(z), \ldots, v_n(z)\}.$$

This gives a sequence (u_n) in $U(f)$ such that $u_n \leq u_{n+1}$ on D and $u_n(z_0) \to v(z_0)$. Next, let U_n be u_n but with its values in D_1 replaced by the Poisson integral $P(z, u_n)$, so that $U_n \in U(f)$. Since $u_n(z_0) \to v(z_0)$ and $u_n \leq U_n \leq v$, we have $U_n(z_0) \to v(z_0)$. We also claim that that $U_n \leq U_{n+1}$ on D: this is clear on $D \setminus D_1$ and in D_1 we just compare the Poisson integrals. By doing this we have found a non-decreasing sequence (U_n) in $U(f)$ such that U_n is harmonic on D_1 and $U_n(z_0) \to v(z_0)$. Since $U_n \leq M$ on D_1, Harnack's theorem 8.3.10 gives us a harmonic function u on D_1 such that $U_n \to u$, and clearly $u(z_0) = v(z_0)$. The idea now is to show that $u = v$ on all of D_1, so that v is harmonic on D_1 and hence on D.

To do this take any other point $z_1 \in D_1$. The same construction gives $W_n \in U(f)$ such that $W_n \leq W_{n+1}$ on D and $W_n(z_1) \to v(z_1)$ and W_n is harmonic on D_1. We then combine U_n and W_n: let $h_n(z) = \max\{U_n(z), W_n(z)\}$ and define H_n to be $P(z, h_n)$ on D_1 and $h_n(z)$ outside D_1, so that $H_n \geq h_n$. Again we have $H_n \leq H_{n+1}$ on D, because we clearly have $h_n \leq h_{n+1}$ and in D_1 we compare Poisson integrals again. Also, the function H_n is in $U(f)$ and is harmonic on D_1. Thus, on D_1,

$$U_n(z) \leq H_n(z) \leq v(z) \leq M, \quad W_n(z) \leq H_n(z) \leq v(z) \leq M,$$

and so $H_n(z_0) \to v(z_0)$ and $H_n(z_1) \to v(z_1)$.

By Harnack's theorem 8.3.10 there is a harmonic function h on D_1 such that $H_n \to h$. We also have $u \leq h$ on D_1, since $U_n \leq H_n$. But $u(z_0) = h(z_0) = v(z_0)$ and so $u = h$ on D_1, by the maximum principle, since $(u - h)(z) \leq (u - h)(z_0)$ on D_1. This gives

$$v(z_1) = h(z_1) = u(z_1)$$

and, since z_1 is arbitrary, $v = u = h$ on D_1.

9.1.4 The barrier

Following Ransford [61], let D be a domain in $\mathbb{C}$ and let $x_0 \in X = \partial_\infty D$. A barrier (for D) at x_0 is a subharmonic function b defined on $D \cap N$, where N is an open neighbourhood of x_0, such that

$$b(z) < 0, \quad (z \in D \cap N), \qquad \lim_{z \to x_0, z \in D \cap N} b(z) = 0.$$

If the barrier exists then x_0 is called a regular boundary point, and D is called regular if all its boundary points are regular.

Note that if G is a subdomain of D and x is a boundary point of both D and G, and is regular for D, then x is regular for G.

Note also that simply connected proper subdomains D of $\mathbb{C}$ are regular, as we can write $b(z) = \log |F(z)|$, where $F : D \to D(0,1)$ is the analytic bijection between D and $D(0,1)$ arising from the Riemann mapping theorem.

The following is an example of a non-regular boundary point. Let $D = D(0,1) \setminus \{0\}$, and let $x_0 = 0$. If a barrier b exists at x_0 then because $b(z) < 0$ on $D \cap N$ we get $b(z) \leq t < 0$ on a circle $|z| = s$ with s small and positive, by Lemma 1.5.2. We also have $b(z) \to 0$ as $z \to 0$. By taking $b(z)/|t|$ we can assume that $t = -1$. But then the function $w(z) = 1 + b(sz)$ belongs to the family $U(f)$ from Example 9.1.2 (since $w(z) \leq 0$ for $|z| = 1$ and $\lim_{z \to 0} w(z) \leq 1$), which forces $w \leq v_f = 0$ and hence $b(z) \to -1$ as $z \to 0$, a contradiction.

The next lemma is also from [61] and will be used to prove the boundary properties of the Perron function.

9.1.5 Bouligand's lemma

Let x_0 be a regular boundary point of D and let N_0 be a spherical disc centred at x_0. Let $\delta > 0$. Then there exists a function w subharmonic on D such that

$$w(z) < 0 \quad (z \in D), \quad w(z) = -1 \quad (z \in D \setminus N_0), \quad \liminf_{z \to x_0, z \in D} w(z) \geq -\delta.$$

Thus w is negative on D and -1 away from x_0, but not too negative near x_0.

Proof. The idea of the proof is to modify a barrier function by subtracting a Poisson integral. Assume without loss of generality that $0 < \delta < 1$. Choose a neighbourhood N and a barrier function b as in the definition of barrier. Choose an open disc G centred at x_0 with closure satisfying $Cl(G) \subseteq N \cap N_0$. (If x_0 is finite then G is a Euclidean disc, while if $x_0 = \infty$ then G is a set $\{z \in \mathbb{C}^* : |z| > R\}$). Let $E = \partial G \cap D$. Then E is a relatively open subset of ∂G. Choose a compact subset K of E so that $L = E \setminus K$ has angular measure $2\pi\sigma < \delta$. Again, L is relatively open.

We can use the Poisson integral formula to make a harmonic function u on $G \cap \mathbb{C}$, which satisfies $0 \leq u \leq 1$ and is such that $u(z) \to 1$ as $z \to \eta \in L$ and $u(z) \to \sigma$ as $z \to x_0$. If x_0 is finite we just use the Poisson integral formula on G with boundary values 1 on L and 0 on $\partial G \setminus L$, while if $x_0 = \infty$ we have to first use a map $z \to 1/z$.

Now let $\sup\{b(z) : z \in K\} = -m$. Then $-m < 0$ by Lemma 1.5.2, because K is a compact subset of $D \cap N$ and b is negative and subharmonic, and so upper semi-continuous, on $D \cap N$. We may assume that $m = 1$. For $\eta \in K \subseteq E = \partial G \cap D$ we have

$$\limsup_{z \to \eta, z \in D \cap G} (b(z) - u(z)) \leq b(\eta) \leq -1$$

since b is upper semi-continuous and $u \geq 0$. On the other hand if $\eta \in L = E \setminus K$ we have

$$\limsup_{z \to \eta, z \in D \cap G} (b(z) - u(z)) \leq -1$$

since $b < 0$ and $u(z) \to 1$ as z approaches η. This implies that, for every $\eta \in E = \partial G \cap D$,

$$\limsup_{z \to \eta, z \in D \cap G} (b(z) - u(z)) \leq -1.$$

So we define

$$w(z) = \max\{-1, b(z) - u(z)\} \quad (z \in D \cap G), \quad w(z) = -1 \quad (z \in D \setminus G).$$

Then w is subharmonic in D. Since $b < 0$ and $u \geq 0$ on $D \cap G$ we have $w < 0$, and for $w \in D \setminus N_0$ we have $w = -1$. Also as $z \to x_0$ with $z \in D$ then $z \in G$ and $b(z) \to 0$ and $u(z) \to \sigma < \delta$ so $w(z) \to -\sigma > -\delta$.

9.1.6 Lemma

Let f and g be bounded real functions on X. Then $v_f + v_g \leq v_{f+g}$ on D. In particular, $v_f(z) \leq -v_{-f}(z)$ on D.

Proof. Let $u_f \in U(f)$ and $u_g \in U(g)$. Then $u_f + u_g \in U(f+g)$ and so

$$u_f(z) + u_g(z) \leq v_{f+g}(z)$$

on D. Now take the suprema over $U(f)$ and $U(g)$.

9.1.7 Theorem

Let x_0 be a regular boundary point of D and let f be bounded on X. Then

$$M_0 = \liminf_{x \to x_0} f(x) \leq \liminf_{z \to x_0, z \in D} v_f(z) \leq \limsup_{z \to x_0, z \in D} v_f(z) \leq M_1 = \limsup_{x \to x_0} f(x). \tag{9.2}$$

In particular, if f is continuous at x_0 then $v_f(z) \to f(x_0)$ as $z \to x_0$ with $z \in D$. Hence if f is continuous on X and D is regular then v_f solves the Dirichlet problem for f on D.

Proof. Let $M = \{\sup |f(x)| : x \in X\}$. Then $M + M_0 \geq 0$. Let $\delta > 0$ and take a neighbourhood N_0 of x_0 such that $f(x) > M_0 - \delta$ on $X \cap N_0$. Take a spherical disc N centred at x_0, whose closure lies in N_0. Define w as in Bouligand's lemma, using δ and the disc N.

Set

$$u(z) = M_0 - \delta + (M + M_0)w(z).$$

Then u is subharmonic on D. Let $x \in X$. If $x \in N_0$ then as $z \to x$ with $z \in D$ we have, since $w < 0$,

$$\limsup u(z) \leq M_0 - \delta \leq f(x).$$

On the other hand if $x \notin N_0$ then as $z \to x_0$ with $z \in D$ we have $z \notin N$ and so $w(z) \leq -1$ and

$$\limsup u(z) \leq -\delta - M < f(x).$$

Hence $u(z) \leq v_f(z)$. But then

$$\liminf_{z \to x_0, z \in D} v_f(z) \geq \liminf_{z \to x_0, z \in D} u(z) \geq M_0 - \delta - \delta(M + M_0).$$

Since δ is arbitrary we get

$$\liminf_{z \to x_0, z \in D} v_f(z) \geq M_0.$$

to Applying the same argument to $-f$ gives

$$\liminf_{z\to x_0, z\in D} v_{-f}(z) \geq -M_1$$

and so, using Lemma 9.1.6,

$$\limsup_{z\to x_0, z\in D} v_f(z) \leq \limsup_{z\to x_0, z\in D} -v_{-f}(z) \leq M_1.$$

9.1.8 A sufficient condition for existence of the barrier

Let $x_0 \in X$. Let E be the component of X which contains x_0 (i.e. the union of all connected subsets of X which contain x_0) and suppose that $E \neq \{x_0\}$. Then x_0 is regular. In particular if there exists a path in X joining x_0 to $x_1 \neq x_0$ then x_0 is regular.

Proof. Suppose first that $x_0 = \infty$ and choose $x_1 \in E\setminus\{\infty\}$. If γ is a closed PSC in D then the winding number $n(\gamma, z)$ is integer-valued and continuous on $(\mathbb{C} \cup \{\infty\}) \setminus \gamma$: here we set $n(\gamma, \infty) = 0$. Thus $n(\gamma, x_1) = 0$ (because otherwise E would be partitioned into relatively open sets $\{z \in E : n(\gamma, z) = 0\}$ and $\{z \in E : n(\gamma, z) \neq 0\}$, the second non-empty by assumption and the first non-empty since ∞ belongs to it). Assume without loss of generality that $x_1 = 0$.

Thus we can define an analytic branch of $\log z = u(z) + iv(z)$ on D and

$$b(z) = -\mathrm{Re}\left(\frac{1}{\log z}\right) = \frac{-u}{u^2 + v^2}$$

is harmonic on D and has $b(z) < 0$ on $D \cap \{z : |z| > 1\}$ and $|b(z)| \leq 1/|u(z)| \to 0$ as $z \to \infty$.

If x_0 is finite then without loss of generality $x_0 = 0$ and we first apply the transformation $z \to 1/z$.

9.2 Convexity and subharmonic functions

9.2.1 Theorem

Let u be subharmonic in $a < |z| < b$. For $a < r < b$ set

$$I(r, u) = \frac{1}{2\pi}\int_0^{2\pi} u(re^{i\tau})d\tau.$$

Then $I(r, u)$ is a convex function of $\log r$ on (a, b) i.e.

$$I(s, u) \leq \frac{\log t/s}{\log t/r} I(r, u) + \frac{\log r/s}{\log r/t} I(t, u) \tag{9.3}$$

for $a < r < s < t < b$.

If $a = 0$ and $I(r, u)$ is bounded above as $r \to 0+$, then $I(r, u)$ is non-decreasing on $(0, b)$.

Proof. Let $a < r < s < t < b$. Take continuous f_n such that $f_{n+1} \leq f_n$ and $f_n \to u$ pointwise on the union of the circles $|z| = r, |z| = t$ (on which u is upper semi-continuous: we can do this by 1.5.1). Let D be the annulus $r < |z| < t$. If $\zeta \in \partial D$ then since u is upper semi-continuous we get $\limsup_{z\to\zeta, z\in D} u(z) \leq f_n(\zeta)$. Thus $u \in U(f_n)$ in the terminology of Perron's method. Solving the Dirichlet problem for f_n gives functions u_n harmonic on $r < |z| < t$ and continuous on $r \leq |z| \leq t$ and equal to f_n on the boundary circles. Further, $u \leq u_n$ for $r \leq |z| \leq t$.

Now, on (r, t), we have, with $v = u_n$,

$$\frac{d^2 I(s,v)}{d(\log s)^2} = \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial^2 v}{\partial \sigma^2}\, d\tau = -\frac{1}{2\pi}\int_0^{2\pi} v_{\tau\tau}\, d\tau = 0, \quad \sigma = \log s,$$

using the fact that v is locally the real part of an analytic function and so the real part of an analytic function of $\log z = \sigma + i\tau$. Thus on $[r, t]$ we have

$$I(s, u_n) = p_n \log s + q_n$$

for some constants p_n, q_n.

We now have, for $r < s < t$,

$$I(s,u) \le I(s,u_n) = \frac{\log t/s}{\log t/r} I(r,u_n) + \frac{\log r/s}{\log r/t} I(t,u_n) \to \frac{\log t/s}{\log t/r} I(r,u) + \frac{\log r/s}{\log r/t} I(t,u),$$

by the monotone convergence theorem (use the fact that $0 \le u_1 - u_n \uparrow u_1 - u$). This proves (9.3).

Now assume that $a = 0$ and $I(r,u)$ is bounded above as $r \to 0+$. Letting $r \to 0+$, we note that the $\frac{\log t/s}{\log t/r}$ term is positive but tends to 0. Since $\frac{\log r/s}{\log r/t} \to 1$, we get $I(s,u) \le I(t,u)$.

9.2.2 Theorem

Let u be subharmonic and bounded above in $0 < |z| < R$. Then setting $u(0) = K = \lim_{r\to 0+} I(r,u)$ makes u subharmonic in $D(0,R)$.

Proof. Take $M > 0$ such that $u \le M$ on $0 < |z| < R$. Since $I(r,u)$ is a non-decreasing function of r, and tends to K as $r \to 0+$, we automatically get $u(0) \le I(r,u)$ and the sub-mean value property. Thus we only need to show that u is upper semi-continuous at 0. Let $0 < s < R$.

Claim: We have $\limsup_{z\to 0} u(z) \le I(s,u)$.

Let f_n be continuous on $|z| = s$, with $f_{n+1} \le f_n$ and $f_n \to u$ pointwise (again these exist by 1.5.1). Using Poisson's formula let u_n be harmonic on $|z| < s$, continuous on $|z| \le s$ and equal to f_n on $|z| = s$. Then $u_n(0) = I(s, f_n)$ by Poisson's formula.

For a given n take $N > 0$ such that $u_n(0) + N > 0$ and $r_n > 0$ such that $u_n(z) + N > 0$ for $|z| \le r_n$. Let $0 < r \le r_n$ and $D = \{z : r < |z| < s\}$ and set

$$v_n(z) = (M+N)\frac{\log s/|z|}{\log s/r} + u_n(z).$$

Then $v_n(z) = u_n(z)$ on $|z| = s$, while $v_n(z) > M$ on $|z| = r$. Hence $u \le v_n$ on D, since $\limsup_{z\to\zeta, z\in D}(u(z) - v_n(z)) \le 0$ for every $\zeta \in \partial D$, using the fact that u is upper semi-continuous with $u \le u_n$.

Keeping z fixed and letting $r \to 0+$ we get $u(z) \le u_n(z)$ for $0 < |z| < s$ and so

$$\limsup_{z\to 0} u(z) \le \limsup_{z\to 0} u_n(z) = u_n(0) = I(s, f_n) \to I(s,u)$$

as $n \to \infty$.

This proves the Claim. Now letting $s \to 0+$ we get $\limsup_{z\to 0} u(z) \le K$.

9.2.3 Example

The following construction gives a subharmonic function on $\mathbb{C}$ which is not continuous at 0. Let

$$u(z) = \sum_{n=1}^{\infty} \frac{1}{n^2} \log \left| z - \frac{1}{n} \right|.$$

Then u is subharmonic on $\mathbb{C} \setminus \{0\}$; to see this, note that if $0 \neq z_0 \in \mathbb{C}$ then there exists a neighbourhood U_0 of z_0 containing at most one of the singularities $1/n$, say $1/n_0$, which implies that

$$u(z) = \frac{1}{n_0^2} \log \left| z - \frac{1}{n_0} \right| + \mathrm{Re} \left(\sum_{n \neq n_0} \frac{1}{n^2} \log \left(z - \frac{1}{n} \right) \right)$$

on U_0. Since $u(z) \leq \displaystyle\sum_{n=1}^{\infty} \frac{\log 2}{n^2}$ for $0 < |z| < 1$, Theorem 9.2.2 shows that u extends to be subharmonic on $\mathbb{C}$. For $x < 0$ and $m \in \mathbb{N}$ we have $|x - 1/n| \geq 1/n$ and so

$$u(x) \geq -\sum_{n=1}^{\infty} \frac{\log n}{n^2} > -\infty = u(1/m).$$

Thus the extension of u to $\mathbb{C}$ is not continuous at 0.

9.2.4 Theorem

Let D, G be domains in $\mathbb{C}$ and let u be subharmonic on G, and $f : D \to G$ analytic. Then $v = u \circ f$ is subharmonic on D.

Proof. We assume that f is non-constant and that $u \not\equiv -\infty$, since otherwise the result is obvious. We show first that v is upper semi-continuous: if $v(z_0) < L$ then $u(f(z_0)) < L$ so $u(w) < L$ near $f(z_0)$ and so $v(z) < L$ near z_0.

Assume that r is small and positive and that f is one-one near z_0. Take continuous functions v_n, decreasing pointwise to v on $S(z_0, r)$, and form the Poisson integrals V_n. Then V_n is harmonic on $D(z_0, r)$ and $V_n(z) \to v_n(u)$ as $z \to u \in S(z_0, r)$ from inside the circle. Define h_n on the closure of $W = f(D(z_0, r))$ by $h_n(f(z)) = V_n(z)$. Then h_n is harmonic on W and continuous on the closure of W. As $w_m \to w \in \partial W, w_m \in W$, we have, since u is upper semi-continuous,

$$\limsup u(w_m) \leq u(w) = v(f^{-1}(w)) \leq v_n(f^{-1}(w)) = h_n(w)$$

and we get $u \leq h_n$ on W. Hence $v \leq v_n$ on $D(z_0, r)$ and

$$v(z_0) \leq v_n(z_0) = \frac{1}{2\pi} \int_0^{2\pi} v_n(z_0 + re^{i\theta}) d\theta \to \frac{1}{2\pi} \int_0^{2\pi} v(z_0 + re^{i\theta}) d\theta,$$

by the monotone convergence theorem applied to $v_1 - v_n$. Hence v has the sub-mean value property at z_0.

Finally, the multiple points z^* of f are isolated, since they are zeros of f'. By Theorem 9.2.2 these z^* are removable singularities of v, since $v(z) = u(f(z))$ is bounded above as $z \to z^*$.

9.2.5 Theorem

Let u be subharmonic and bounded above in $\mathbb{C}$. Then u is constant.

Proof. Assume u non-constant. Then without loss of generality $u(0) = 0$. Let $v = \max\{u, 0\}$ so that v is subharmonic. Let $0 < r < s < t < \infty$. Then convexity gives

$$I(s) = I(s, v) \leq \frac{\log t/s}{\log t/r} I(r) + \frac{\log r/s}{\log r/t} I(t).$$

Let $t \to \infty$. Since $\frac{\log r/s}{\log r/t} \to 0$ and since $0 \leq I(t) < M$ for some fixed M we get $I(s) \leq I(r)$. Since I is non-decreasing we have I constant on $(0, \infty)$. Since v is upper semi-continuous and $v(0) = 0$ we get $I(s) \equiv 0$. By Theorem 8.3.5 we have, for $|z| \leq s$,

$$0 \leq v(z) \leq 3I(2s, v) = 0$$

and so $v \equiv 0$. But then u has a maximum at 0 and so is constant.

We give another proof of this result (Beardon). Assume that u is non- constant. We can also assume WLOG that $\sup\{u(z) : z \in \mathbb{C}\} = 0$. Let $m = \max\{u(z) : |z| = 1\}$, which exists because u is upper semi-continuous. Then $m < 0$, since otherwise u has a maximum in $\mathbb{C}$ and so is constant. Now fix z_0 with $|z_0| > 1$. Let $\varepsilon > 0$ and let $R > 1$ be large. The function

$$v(z) = u(z) - \varepsilon \log |z|$$

is subharmonic in $0 < |z| < +\infty$, and we have $v(z) \leq m$ for $|z| = 1$ and for $|z| = R$, since R is large. Hence we get

$$u(z_0) \leq v(z_0) + \varepsilon \log |z_0| \leq m + \varepsilon \log |z_0|,$$

and so $u(z_0) \leq m < 0$ since ε may be chosen arbitrarily small. Thus $u(z) \leq m < 0$ for $|z| > 1$ and so on $\mathbb{C}$ by the maximum principle, contradicting the assumption that the supremum of u is 0.

9.2.6 Exercises

(a) Prove *Iversen's theorem*: if f is a non-constant entire function then there exists a path γ tending to infinity such that $f(z) \to \infty$ as z tends to infinity on γ. (Hint: consider a component C_n of the set $E_n = \{z : |f(z)| > n\}$. Prove that f is unbounded on C_n and take a component of E_{n+1}).

(b) Let u be subharmonic in $\mathbb{C}$ such that $u = 0$ on the imaginary axis but $u(z) > 0$ for at least one z in the right half-plane. Let $0 < s < 1/2$. Prove that there exists a path tending to infinity in the right half-plane on which $u(z) > |z|^s$. (Hint: take $s < t < 1/2$ and δ small and positive and consider the function $u(z) - \delta \operatorname{Re}(z^t)$).

9.2.7 Lemma

Let u be subharmonic and bounded above on the domain D in $\mathbb{C}$. Suppose that

$$\limsup_{z \to \zeta, z \in D} u(z) \leq 0 \tag{9.4}$$

for at least one, and for all but finitely many, $\zeta \in X = \partial_\infty D$. Then $u \leq 0$ on D.

Proof. Let $\zeta_1, \dots, \zeta_n$ be the points in X at which (9.4) fails. Let $G = \mathbb{C} \setminus \{\zeta_1, \dots, \zeta_n\}$ and define v on G as follows. On $G \setminus D$ we set $v = 0$, while on D we set $v = \max\{u, 0\}$.

Then v is subharmonic and bounded above in G (it clearly has the sub-mean value property and the fact that v is upper semi-continuous follows from (9.4)), and hence subharmonic and bounded above in $\mathbb{C}$. So v is constant.

Now let $\zeta^* \in X$ be such that (9.4) does hold. Then there exists a sequence $z_n \in D$ such that $z_n \to \zeta^*$ and $u(z_n) \to 0$, so that $v(z_n) \to 0$. Since v is constant this gives $v \equiv 0$ and $u \leq 0$.

9.2.8 Example

This example shows that in the last lemma we cannot delete the hypothesis that u is bounded above. Let

$$u(z) = \mathrm{Re}\left(\frac{1+z}{1-z}\right), \quad |z| < 1.$$

Then for $|w| = 1, w \neq 1$ we have $u(z) \to 0$ as $z \to w$, but u is unbounded in $D(0,1)$.

Chapter 10

Harmonic measure

10.1 Definition of the harmonic measure

For suitable domains D and subsets E of $X = \partial_\infty D$ the harmonic measure $\omega(z, E, D)$ will be defined for $z \in D$. It will then turn out that the harmonic measure is for fixed E a harmonic function of z, while for fixed z it is a measure on a suitable collection of subsets of X.

One of the main applications of harmonic measure is the two-constants theorem 10.2.10 which gives a powerful improvement of the maximum principle for subharmonic functions.

10.1.1 Semi-regular domains

Let D be a domain in $\mathbb{C}$. We say that D is semi-regular if $X = \partial_\infty D$ is infinite and all but finitely many $x \in X = \partial_\infty D$ are regular. Here we use the spherical metric on $\mathbb{C}^*$, which makes $\mathbb{C}^*$ compact. Note that if U is an open subset of X then $U = V \cap X$ for some open subset V of $\mathbb{C}^*$, by definition of the relative topology. Also any closed subset of X is compact, as is X, because a closed subset of a compact set is compact.

For a set Y a collection S of subsets of Y is called a σ-algebra if it is non-empty and has the following two properties: (i) $A \in S$ implies that $Y \setminus A$ is in S; (ii) if $A_1, A_2, \ldots$ are countably many elements of S then their union is in S. Obviously the power set of Y is a σ-algebra. It is easy to prove that if S_t is a σ-algebra of subsets of Y for every $t \in T$ then $\bigcap_{t \in T} S_t$ is also a σ-algebra of subsets of Y. So for any collection U of subsets of Y, taking the intersection of all σ-algebras S of subsets of Y with $U \subseteq S$ gives a σ-algebra, which is said to be generated by U.

If Y is also a topological space then we can form the σ-algebra generated by the open subsets of Y, which is the smallest σ-algebra of subsets of Y containing all the open subsets of Y. Its elements are called Borel sets.

We now identify the Borel sets of $X = \partial_\infty D$. We claim that the Borel subsets of X are precisely the sets $B \cap X$ where B is a Borel subset of $\mathbb{C}^*$. To see this let B_1 be the collection of Borel subsets of X and let F be the collection of Borel subsets of $\mathbb{C}^*$. Then $B_2 = \{B \cap X : B \in F\}$ is a σ-algebra and every open subset of X is an element of B_2 since it is $U \cap X$ for some open $U \in F$. So $B_1 \subseteq B_2$.

But $B_3 = \{W \subseteq \mathbb{C}^* : W \cap X \in B_1\}$ is a σ-algebra and it contains all open subsets of $\mathbb{C}^*$, so $F \subseteq B_3$. Hence $V \in B_2$ gives $V = B \cap X$ with $B \in F$ and hence $B \in B_3$, so that $V = B \cap X \in B_1$. Thus $B_1 = B_2$.

10.1.2 An example of a semi-regular domain

For $n \in \mathbb{N}$ let C_n be the circle $|z-n| = \frac{1}{4}$. Let D be the unbounded component of $\mathbb{C} \setminus \bigcup_{n=1}^{\infty} C_n$ i.e.

$$D = \left\{ w \in \mathbb{C} : |w-n| > \frac{1}{4} \quad \text{for all } n \in \mathbb{N} \right\}.$$

For $x \in C_n$ the component of $\partial_\infty D$ containing x is not $\{x\}$, since C_n is itself connected, and so x is a regular point of X by 9.1.8. On the other hand if $\infty \in E \subseteq X$ and E is connected then $E = \{\infty\}$, because if $y \neq \infty$ is in E we can take a large $n \in \mathbb{N}$ and partition E as

$$\left\{ x \in E : |x| < n + \frac{1}{2} \right\} \cup \left\{ x \in E : |x| > n + \frac{1}{2} \right\}$$

with both sets relatively open and non-empty. Thus the component of $\partial_\infty D$ containing ∞ is just $\{\infty\}$. So ∞ fails to satisfy the sufficient condition 9.1.8 for a barrier, but our definition of harmonic measure will still make sense for D.

10.1.3 A linear functional

Let D be a semi-regular domain in $\mathbb{C}$, with boundary $X = \partial_\infty D$. Let Y be the vector space of functions $f : X \to \mathbb{R}$ which are bounded on X and continuous at all but finitely many points of X. Applying Perron's method gives a harmonic Perron function v_f on D, and Theorem 9.1.7 shows that

$$\lim_{z \to x, z \in D} v_f(z) = f(x)$$

for all but finitely many $x \in X$.

If $f, g \in Y$ and $f = g$ except on a finite set then $v_f - v_g$ is harmonic and bounded and has boundary values 0 except on a finite set, so that $v_f = v_g$ by Lemma 9.2.7. Similarly if $f, g \in Y$ then $v_{f+g} - v_f - v_g$ is harmonic and bounded and again has boundary values 0 except on a finite set, so we get $v_{f+g} = v_f + v_g$. Also if $f \in Y$ with $f \leq 0$ on X then $v_f \leq 0$, again by Lemma 9.2.7. Finally if f is a constant (say M) on X then v_f has boundary values M except on a finite set and so is M by Lemma 9.2.7.

Fix z in D. Then

$$f \to L(z, f) = v_f(z)$$

is a non-negative linear functional on Y (this just means that $f \geq 0$ on X implies that $v_f(z) \geq 0$ on D).

The rest of this section will be devoted to proving the following theorem from first principles: it can, however, be deduced rather quickly from the Riesz representation theorem (2.14 of W. Rudin, *Real and Complex Analysis*).

10.1.4 Theorem

Let D be as above and fix z_0 in D. Then there exist a σ-algebra Π of subsets of X and a probability measure μ on Π (this means a measure $\mu : \Pi \to [0,1]$ with $\mu(X) = 1$) such that:

(i) every Borel subset E of X is in Π;

(ii) for E in Π, the measure $\mu(E)$ is the infimum of $\mu(V)$ over all open V containing E;

(iii) $\mu(E)$ is the supremum of $\mu(K)$ over all compact $K \subseteq E$;

(iv) if $A \subseteq B$ and $\mu(B) = 0$ then $\mu(A)$ exists and is 0;

(v) if V is open then $\mu(V)$ is the supremum of $L(z_0, g) = v_g(z_0)$ taken over all continuous functions $g : X \to [0, 1]$ such that $g \leq \chi_V$ on X.

In order to be a measure, μ must satisfy $\mu(\bigcup E_j) = \sum \mu(E_j)$ whenever the E_j are countably many pairwise disjoint elements of Π. Now V is an open subset of X if and only if $V = U \cap X$ where U is open in $\mathbb{C}^*$, and so if and only if $K = X \setminus V$ is of form $K = F \cap X$ where F is closed in $\mathbb{C}^*$. But then K is compact, since X is compact, and $\mu(V) = 1 - \mu(K)$ and $\mu(E) = 1 - \mu(X \setminus E)$. Hence properties (ii) and (iii) are equivalent provided μ is a measure.

Note that (ii) and (v) imply that this measure μ is unique (because there is a unique definition for open V and hence for every $E \in \Pi$).

10.1.5 Example

Let $D = D(0, 1)$. For a Borel subset A of $X = \partial D$, let $\chi_A(t)$ be 1 on A and 0 elsewhere, and set

$$\mu(A) = \frac{1}{2\pi} \int_0^{2\pi} K(z, e^{it}) \chi_A(e^{it})\, dt,$$

in which $K(z, e^{it})$ is the Poisson kernel. Note that if we keep A fixed then what we get is a harmonic function of z on D. Also (i) is satisfied because this integral exists for every Borel set A.

To check (ii) let $r = |z| < 1$. Obviously if $A \subseteq V$ then $\mu(A) \leq \mu(V)$. But given $A \subseteq X$ and $\delta > 0$ we can look at $B = \{t \in [0, 2\pi] : e^{it} \in A\}$ and the Lebesgue measure $\lambda(B)$ of B is the infimum of the Lebesgue measure of U over all open sets U with $B \subseteq U \subseteq \mathbb{R}$. So we can find an open $U \subseteq \mathbb{R}$ such that $B \subseteq U$ and $U \setminus B$ has Lebesgue measure less than δ. Now let $V = \{e^{it} : t \in U\}$. Then V is an open subset of X with $A \subseteq V$ and $\chi_V(e^{it}) = 1$ implies that $t \in U$. Hence $\chi_V(e^{it}) = \chi_A(e^{it})$ for all $t \in [0, 2\pi]$ apart from a set of Lebesgue measure at most δ. This gives, since

$$0 \leq K(z, e^{it}) \leq \frac{1 + r}{1 - r}, \tag{10.1}$$

the inequality

$$\mu(A) \leq \mu(V) \leq \mu(A) + \delta \left(\frac{1 + r}{1 - r} \right),$$

which proves (ii).

To check (iv) let B be a Borel subset of X with $\mu(B) = 0$ and let A be any subset of B. Let $|z| = r < 1$. Since

$$K(z, t) \geq \frac{1 - r}{1 + r}$$

we have

$$\int_0^{2\pi} \chi_B(e^{it})\, dt = 0.$$

So the set $C = \{t \in [0, 2\pi] : e^{it} \in B\}$ has Lebesgue measure 0. Then every subset of C is Lebesgue measurable with Lebesgue measure 0, and $\chi_A(e^{it}) = 0$ for every $t \in [0, 2\pi]$ apart from a set of Lebesgue measure 0. Hence (10.1) implies that

$$\frac{1}{2\pi} \int_0^{2\pi} K(z, e^{it}) \chi_A(e^{it})\, dt$$

exists and is 0.

Now we check (v). Let $A \subseteq X$ be open. First if $g : X \to [0,1]$ is continuous and $g \leq \chi_A$ then

$$L(z,g) = \frac{1}{2\pi}\int_0^{2\pi} K(z,e^{it})g(e^{it})\,dt \leq \frac{1}{2\pi}\int_0^{2\pi} K(z,e^{it})\chi_A(e^{it})\,dt = \mu(A).$$

Next, $1-\chi_A$ is upper semi-continuous on X and by Theorem 1.5.1 we can take continuous $f_n \uparrow \chi_A$ on X. Thus

$$\mu(A) \geq \frac{1}{2\pi}\int_0^{2\pi} K(z,e^{it})f_n(e^{it})\,dt = L(z,f_n) \uparrow \frac{1}{2\pi}\int_0^{2\pi} K(z,e^{it})\chi_A(e^{it})\,dt = \mu(A)$$

by the monotone convergence theorem. Thus μ satisfies conditions (i) to (v).

The fact that μ is a measure in this example is easily seen from the fact that if $E_1, E_2, \ldots$ are pairwise disjoint subsets of X with union E then $\chi_E = \sum \chi_{E_j}$ and so the integrals add up. Also $\mu(X) = 1$ because the Poisson extension to D of a constant function on X is constant.

For general semi-regular domains D the situation is more complicated and the first step in the proof of Theorem 10.1.4 involves looking at upper semi-continuous functions. Throughout this section D will be a semi-regular domain and X will be $\partial_\infty D$.

10.1.6 Theorem

Let $f : X \to \mathbb{R}$ be upper semi-continuous and let H_f be the family of all continuous real-valued g with $g \geq f$ on X. Choose continuous functions f_n such that $f_n \downarrow f$ on X, and let $u_n(z) = L(z,f_n) = v_{f_n}(z)$. Then $u = \lim u_n$ is either identically $-\infty$ on D or harmonic in D, and $u(z) = \inf\{L(z,g) : g \in H_f\}$. In particular, u is independent of the particular choice of the sequence (f_n).

Proof. The functions f_n exist by Theorem 1.5.1, and the first assertion follows from Harnack's theorem 8.3.10, since $u_n \leq u_{n-1}$ on D. Also the set H_f makes sense, since f is bounded above on X.

To prove that $u(z) = \inf\{L(z,g) : g \in H_f\}$, take any $g \in H_f$, and $\varepsilon > 0$, and any w_0 in X. Set $G(x) = g(x) + \varepsilon$. Then for some large N we have $G(w_0) > f_N(w_0)$ and so $G(w) > f_N(w)$ for all w in a relatively open neighbourhood V of w_0. Hence $G(w) > f_n(w)$ for all w in V and all $n \geq N$. Now the compact set X can be covered by finitely many such V, and so there exists M such that $G(w) > f_M(w)$ for all w in X. Thus $u(z) \leq u_M(z) \leq L(z,G) = L(z,g) + \varepsilon$ for all z in D. This gives $u(z) \leq L(z,g)$ on D.

Next, take any z in D, and $K > u(z)$, and $n \in \mathbb{N}$ with $u_n(z) < K$. Now $f_n \geq f$ on X and $L(z,f_n) = u_n(z) < K$, so that K is not a lower bound for $\{L(z,g) : g \in H_f\}$. Thus $u(z)$ is the greatest lower bound as asserted.

10.1.7 Definition

For an upper semi-continuous function f on X we define $u(z)$ on D as follows. For each z in D, we set $u(z) = u_f(z)$ to be the infimum of $L(z,g) = v_g(z)$ over all continuous g with $g \geq f$ on X. We have just seen that u is harmonic or identically $-\infty$ on D, and we call u the harmonic extension of f to D. Note that if f is itself continuous, then $u(z) = v_f(z) = L(z,f)$; in particular this is true if f is constant. It is clear that if f_1 and f_2 are upper semi-continuous on X with $f_1 \leq f_2$ then the harmonic extension of f_1 is bounded above by that of f_2. Hence if A, B are real numbers and $A \leq f \leq B$ on X then $A \leq u(z) \leq B$ on D.

For a closed subset E of X, the characteristic function χ_E is upper semi-continuous, and we write

$$\omega(z, E, D) = u_{\chi_E}(z).$$

This is a harmonic function on D, bounded above by 1 and below by 0. We will refer to $\omega(z, E, D)$ as the harmonic measure of E with respect to D. Note that if g is continuous on X with $g \geq \chi_E$ then $g \geq 0$ and $h(x) = \min\{g(x), 1\}$ is also continuous with $h \geq \chi_E$. So in fact

$$\omega(z, E, D) = \inf\{L(z, h) : \, h : X \to [0, 1], \, h \geq \chi_E, \, h \text{ continuous } \}.$$

To see this, observe that the set of h as above is a subset of the set of g, so any lower bound for the $L(z, g)$ is a lower bound for the $L(z, h)$, and hence the infimum of the $L(z, h)$ is not less than that of the $L(z, g)$; on the other hand, given g there exists an h with $\chi_E \leq h \leq g$ and so any lower bound for the $L(z, h)$ is a lower bound for the $L(z, g)$ also.

Obviously,

$$\omega(z, X, D) = 1, \quad \omega(z, \emptyset, D) = 0,$$

since in both cases the characteristic function is constant.

Note that if A and B are closed subsets of X and $A \subseteq B$ then any $g : X \to [0, 1]$ which satisfies $g \geq \chi_B$ also satisfies $g \geq \chi_A$, so $\omega(z, B, X) \geq \omega(z, A, X)$.

10.1.8 Urysohn's lemma

Let Y be a compact metric space and let $K \subseteq V \subseteq Y$, with K compact and V open (in both cases relative to Y). Then there exists a continuous function $g : Y \to [0, 1]$, with $g = 1$ on K and $g = 0$ off V.

We just define g by

$$1 - g(y) = \frac{d(y, K)}{d(y, K) + d(y, V^c)}$$

in which $V^c = Y \setminus V$ and d denotes the metric. Here the distance $d(y, A)$ is defined for any closed (and hence compact) $A \subseteq Y$ and is the minimum of the continuous function $d(y, a)$ over $a \in A$. This distance is continuous for a given closed A, and we cannot have $d(y, K) + d(y, V^c) = 0$, because if $d(y, K) = 0$ then $y \in K \subseteq V$ and so $d(y, V^c) > 0$.

10.1.9 Boundary behaviour of the harmonic measure of a closed set

Let E be a closed subset of X.
(a) If $x \in X \setminus E$ and x is a regular boundary point of D then $\omega(z, E, D) \to 0$ as $z \to x$ in D.

To see this, just take $K = E$ and $V = X \setminus \{x\}$ in Urysohn's lemma. This gives a continuous $g : X \to [0, 1]$ with $g \geq \chi_E$ on X and $g(x) = 0$. On D we have

$$0 \leq \omega(z, E, D) \leq v_g(z) \to g(x) = 0, \quad z \to x.$$

In particular this is always the case if D is simply connected, by §9.1.4.

(b) If x is a regular boundary point of D, and an interior point of E (with respect to X), then $\omega(z, E, D) \to 1$ as $z \to x$ in D.

To see this, take $r > 0$ with $V = D(x,r) \cap X \subseteq E$. Then apply Urysohn's lemma to get a continuous $g : X \to [0,1]$ with $g(x) = 1$ and $g = 0$ off V (and so $g = 0$ off E). If $h : X \to [0,1]$ is continuous with $h \geq \chi_E$ on X, then we have $h \geq g$ on X and so $L(z,h) \geq L(z,g)$ on D. Thus, on D,

$$1 \geq \omega(z,E,D) \geq L(z,g) \to g(x) = 1, \quad z \to x.$$

(c) If D is regular then, with $f = \chi_E$, the harmonic measure $\omega(z,E,D)$ agrees with the Perron function $v_f(z)$ defined in §9.1.1.

To prove (c), first let $g : X \to [0,1]$ be continuous with $g \geq f$, and let y be a member of the Perron family $U(f)$ as defined in §9.1.1. Then for every $\zeta \in X$ we have

$$\limsup_{z \to \zeta, z \in D} y(z) \leq f(x) \leq g(x) = \lim_{z \to \zeta, z \in D} v_g(z), \quad \limsup_{z \to \zeta, z \in D} (y(z) - v_g(z)) \leq 0$$

and so the maximum principle shows that $y(z) \leq v_g(z) = L(z,g)$ on D. Thus $y(z) \leq \omega(z,E,D)$ on D, and taking the supremum over these y we get $v_f(z) \leq \omega(z,E,D)$ on D. But we also know that if $x \in X \setminus E$ then $\omega(z,E,D) \to 0$ as $z \to x, z \in D$. Since $\omega(z,E,D) \leq 1$ we get $\omega(z,E,D) \in U(f)$ and so $\omega(z,E,D) \leq v_f(z)$.

10.1.10 Example

In §10.1.9(a), we cannot delete the hypothesis that x is a regular boundary point. Let D_1 be the domain $0 < |z| < 1$ as in Example 9.1.2, let A be the circle $|x| = 1$, and let $g = \chi_A$. Then g is continuous on X and $g = 1$ except at one point, and so $v_g = v_1 = 1$ by the argument in §10.1.3. Since $L(z,h) \geq L(z,g)$ for continuous h with $h \geq g$ on X, this now shows that $\omega(z,A,D_1) = L(z,g) \equiv 1$ on D_1, and so $\omega(z,A,D_1)$ fails to tend to 0 as $z \to 0$.

We return to this theme in §15.1.8.

10.1.11 The harmonic measure of an open set

For a closed subset E of X, we have defined $\omega(z,E,D)$ to be the infimum of $L(z,g) = v_g(z)$ over all continuous $g : X \to [0,1]$ with $g \geq \chi_E$ on X.

If U is an open subset of X, we define $\omega(z,U,D)$ to be the supremum of $L(z,g)$ over all continuous $g : X \to [0,1]$ with $g \leq \chi_U$ on X.

Note that if A and B are relatively open subsets of X and $A \subseteq B$ then any $g : X \to [0,1]$ which satisfies $g \leq \chi_A$ also satisfies $g \leq \chi_B$, so $\omega(z,B,X) \geq \omega(z,A,X)$.

Note also that if E is a clopen (closed and open) subset of X then $h = \chi_E$ is continuous, and so the two definitions both give $\omega(z,E,D) = v_h(z) = L(z,h)$ and in particular they agree.

The next lemma shows that this definition gives the "expected" result that

$$\omega(z,U,D) = 1 - \omega(z,X \setminus U,D).$$

10.1.12 Lemma

Let U be an open subset of X and let $E = X \setminus U$. Then for every z in D we have $\omega(z,U,D) = 1 - \omega(z,E,D)$.

Proof. Obviously if $g : X \to [0,1]$ then $g \le \chi_U$ if and only if $h = 1 - g \ge 1 - \chi_U = \chi_E$ and g is continuous if and only if h is. Thus

$$\begin{aligned}
\omega(z,U,D) &= \sup\{L(z,g) : g : X \to [0,1],\, g \le \chi_U,\, g \text{ continuous }\} \\
. &= \sup\{L(z,g) : g : X \to [0,1],\, h = 1-g \ge \chi_E,\, g \text{ continuous }\} \\
&= \sup\{1 - L(z,h) : h : X \to [0,1],\, h \ge \chi_E,\, h \text{ continuous }\} \\
&= 1 - \inf\{L(z,h) : h : X \to [0,1],\, h \ge \chi_E,\, h \text{ continuous }\} \\
&= 1 - \omega(z,E,D).
\end{aligned}$$

10.1.13 μ-measurable sets

Fix z_0 in D. For *any* subset C of X we define:

$\mu^+(C)$ to be the infimum of $\omega(z_0,U,D)$ over all open U with $C \subseteq U \subseteq X$;

$\mu^-(C)$ to be the supremum of $\omega(z_0,E,D)$ over all closed E with $E \subseteq C \subseteq X$.

Obviously $\mu^+(C)$ and $\mu^-(C)$ both exist, and they are in $[0,1]$.

We say that C is μ-measurable if $\mu^+(C) = \mu^-(C)$, in which case we denote the common value by $\mu(C)$.

10.1.14 Lemma

Let $A \subseteq X$ and let $B = X \setminus A$. Then:
(a) $\mu^-(A) \le \mu^+(A)$;
(b) $\mu^+(A) = 1 - \mu^-(B)$;
(c) if A is μ-measurable then so is B, and $\mu(A) = 1 - \mu(B)$.

Proof. (a) Take closed E and open U with $E \subseteq A \subseteq U$, and using Urysohn's lemma let $g : X \to [0,1]$ be continuous, with $g = 1$ on E and $g = 0$ off U. Then $\chi_E \le g \le \chi_U$ and so

$$\omega(z_0,E,D) \le L(z_0,g) \le \omega(z_0,U,D).$$

This proves (a), and also establishes (iv) of Theorem 10.1.4, because if $A \subseteq C$ and $\mu(C) = 0$ then $0 \le \mu^-(A) \le \mu^+(A) \le \mu^+(C) = 0$.
(b) Here

$$\begin{aligned}
\mu^+(A) &= \inf\{\omega(z_0,U,D) : A \subseteq U \quad (U \text{ open})\} \\
&= \inf\{1 - \omega(z_0,E,D) : E \subseteq B \quad (E \text{ closed})\} \\
&= 1 - \sup\{\omega(z_0,E,D) : E \subseteq B \quad (E \text{ closed})\} \\
&= 1 - \mu^-(B).
\end{aligned}$$

Similarly we get $\mu^+(B) = 1 - \mu^-(A)$ and (c) follows.

10.1.15 Lemma

Let U be an open subset of X. Then U is μ-measurable and $\mu(U) = \omega(z_0,U,D)$.

Proof. Obviously $\mu^+(U) = \omega(z_0, U, D)$. Let $\delta > 0$. Then by the definition of ω for open U there exists a continuous $g : X \to [0, 1]$ with $g \leq \chi_U$ on X and

$$L(z_0, g) > \omega(z_0, U, D) - \delta.$$

Let E be the closed set $E = \{x \in X : g(x) \geq \delta\}$. Then $E \subseteq U$, since $g \leq \chi_U$. This time using the definition of the harmonic measure for closed sets, choose a continuous $h : X \to [0, 1]$, with $h \geq \chi_E$ on X, and with

$$L(z_0, h) < \omega(z_0, E, D) + \delta.$$

Now, on E we have $g \leq 1 \leq \chi_E \leq h$ and on $X \setminus E$ we have $g < \delta$ and $h \geq 0$. Thus

$$g \leq h + \delta$$

on X. So

$$\omega(z_0, E, D) > L(z_0, h) - \delta \geq L(z_0, g) - 2\delta > \omega(z_0, U, D) - 3\delta.$$

Since E is closed and is a subset of U, we have

$$\mu^-(U) \geq \omega(z_0, E, D) \geq \omega(z_0, U, D) - 3\delta = \mu^+(U) - 3\delta,$$

and the lemma follows since δ is arbitrary.

In particular, for open U we have $\mu(U) = \mu^+(U) = \omega(z_0, U, D)$ and this is the supremum of $L(z, g)$ over all continuous g with $g \leq \chi_U$ on X, which immediately establishes assertion (v) of Theorem 10.1.4.

It follows from the last three lemmas that closed sets E are also μ-measurable with $\mu(E) = \omega(z_0, E, D)$. Using the definitions of μ^- and μ^+, we now have (ii) and (iii) of Theorem 10.1.4 for μ-measurable subsets of X.

The next few results deal with the effect of taking unions. For general sets this is rather involved, so it is convenient first to look at disjoint unions of open (and then closed) sets.

10.1.16 Lemma

Let U_1, U_2 be disjoint open subsets of X. Then

$$\mu(U_1) + \mu(U_2) = \mu(W), \quad W = U_1 \cup U_2.$$

Once we have this result for two disjoint open subsets it extends by induction to finitely many pairwise disjoint open subsets.

Proof. Let $E \subseteq W$ be closed. Then so are $E_1 = E \cap U_1 = E \setminus U_2$ and $E_2 = E \cap U_2$. Using Urysohn's lemma form continuous $g_1, g_2 : X \to [0, 1]$ with $g_j = 1$ on E_j and $g_j = 0$ off U_j. Then $g = g_1 + g_2 : X \to [0, 1]$ is continuous and $\chi_E \leq g$ on X. Since $g_j \leq \chi_{U_j}$ this gives

$$\begin{aligned}\omega(z_0, E, D) &\leq L(z_0, g) = L(z_0, g_1) + L(z_0, g_2)\\ &\leq \omega(z_0, U_1, D) + \omega(z_0, U_2, D) = \mu(U_1) + \mu(U_2).\end{aligned}$$

Taking the sup over closed $E \subseteq W$ we get

$$\mu(W) = \mu^-(W) \leq \mu(U_1) + \mu(U_2).$$

Now let $h_j :\to [0,1]$ be continuous with $h_j \leq \chi_{U_j}$. Then $h = h_1 + h_2 : X \to [0,1]$ satisfies $h \leq \chi_W$. So

$$L(z_0, h_1) + L(z_0, h_2) = L(z_0, h) \leq \omega(z_0, W, D) = \mu(W).$$

Taking the sup over h_1 and h_2 gives

$$\mu(U_1) + \mu(U_2) = \omega(z_0, U_1, D) + \omega(z_0, U_2, D) \leq \mu(W).$$

10.1.17 Lemma

Let E_1, E_2 be disjoint closed subsets of X, with union E. Then

$$\mu(E_1) + \mu(E_2) = \mu(E).$$

Again this result extends to finitely many pairwise disjoint closed sets.

Proof. Take an open U with $E \subseteq U \subseteq X$. Since the distance between E_1 and E_2 is positive, as is the distance from E to $X \setminus U$, we can form disjoint open sets U_j with $E_j \subseteq U_j \subseteq U$. To do this we can take a small positive ρ and the union of the open discs of centre $x \in E_j$ and radius ρ. Now

$$\begin{aligned} \mu(E_1) + \mu(E_2) &= \mu^+(E_1) + \mu^+(E_2) \\ &\leq \omega(z_0, U_1, D) + \omega(z_0, U_2, D) \\ &= \mu(U_1) + \mu(U_2) = \mu(U_1 \cup U_2) \leq \mu(U) \end{aligned}$$

by the result for disjoint open sets. Now taking the infimum over $U \supseteq E$ gives $\mu(E_1) + \mu(E_2) \leq \mu(E)$.

Next, take $\delta > 0$ and open sets V_j with $E_j \subseteq V_j$ and $\mu(V_j) = \omega(z_0, V_j, D) < \mu(E_j) + \delta$. By intersecting with open sets U_j as in the previous part it may be assumed that V_1, V_2 are disjoint. Now

$$\mu(E) = \mu^+(E) \leq \mu(V_1 \cup V_2) = \mu(V_1) + \mu(V_2) < \mu(E_1) + \mu(E_2) + 2\delta,$$

and letting $\delta \to 0$ completes the proof.

10.1.18 Theorem

Let V_j be open subsets of X and let $W = \bigcup_{j=1}^{\infty} V_j$. Then $\mu(W) \leq \sum_{j=1}^{\infty} \mu(V_j)$ (Thus μ is countably sub-additive on open sets).

Proof. We first take the case where $W = V_1 \cup V_2$. Let H be a closed (and so compact) subset of W. We claim that there exists $\rho > 0$ such that for each x in H the (spherical) disc $D(x, \rho)$ is contained in one of the V_k. To see this note that each $x \in H$ has $\rho(x) > 0$ such that $X \cap D(x, 2\rho(x))$ is contained in one of the V_j. So by compactness there exist $x_1, \ldots, x_N$ with

$$H \subseteq \bigcup_{j=1}^{N} D(x_j, \rho(x_j)).$$

Let ρ be the minimum of the $\rho(x_j)$, for $j = 1, \ldots, N$. Then each x in H lies in one of the $D(x_j, \rho(x_j))$, and $D(x, \rho)$ is a subset of $D(x_j, 2\rho(x_j))$, which in turns lies in one of the V_k.

Let

$$H_j = \{x \in V_j : \operatorname{dist}(x, X \setminus V_j) \geq \rho\} = \{x \in X : D(x, \rho) \subseteq V_j\}.$$

Then each H_j is closed by continuity and so compact, and $H \subseteq H_1 \cup H_2$. Use Urysohn's lemma to define continuous $g_j : X \to [0,1]$ such that $g_j = 1$ on H_j and $g_j = 0$ off V_j. Then

$$\chi_H(x) \le g(x) = \min\left\{1, \sum_{j=1}^{2} g_j(x)\right\}, \quad g_j(x) \le \chi_{V_j}(x).$$

So, since H is closed and P_j is open,

$$\omega(z_0, H, D) \le L(z_0, g) \le \sum_{j=1}^{2} L(z_0, g_j) \le \sum_{j=1}^{2} \omega(z_0, V_j, D) = \sum_{j=1}^{2} \mu(V_j).$$

Taking the infimum over closed $H \subseteq W$ is arbitrary completes the proof in this case.

This result now extends by induction to the union of finitely many V_j. To handle the general case just note that if $H \subseteq W$ is closed then H is compact and lies in the union of finitely many V_j, so that

$$\omega(z_0, H, D) = \mu^+(H) \le \mu\left(\bigcup_{j=1}^{N} V_j\right) \le \sum_{j=1}^{N} \mu(V_j).$$

10.1.19 Lemma

μ^+ is countably sub-additive on subsets of X.

Proof. Let A_j be subsets of X, $j = 1, 2, \ldots$, let $B = \bigcup_{j \in \mathbb{N}} A_j$ and $\varepsilon > 0$ and choose open U_j such that

$$A_j \subseteq U_j, \quad \omega(z_0, U_j, D) < \mu^+(A_j) + \varepsilon/2^j.$$

Then B is a subset of the union of the U_j and so, by Lemma 10.1.18,

$$\mu^+(B) \le \mu^+\left(\bigcup U_j\right) = \mu\left(\bigcup U_j\right) \le \sum \mu(U_j) = \sum \omega(z_0, U_j, D) \le \varepsilon + \sum \mu^+(A_j).$$

10.1.20 Lemma

Let A be a μ-measurable subset of X and let $\delta > 0$. Then there exist closed E and open U with $E \subseteq A \subseteq U$ and $\mu(U \setminus E) < \delta$.

Proof. Let $\rho > 0$ and choose closed E and open U with $E \subseteq A \subseteq U$ and

$$\mu(A) = \mu^-(A) < \mu(E) + \rho, \quad \mu(A) = \mu^+(A) > \mu(U) - \rho.$$

Take a closed subset G of the open set $U \setminus E$ such that

$$\mu(U \setminus E) < \mu(G) + \rho.$$

Then

$$\mu(U \setminus E) + \mu(E) < \mu(G) + \mu(E) + \rho = \mu(G \cup E) + \rho \le \mu(U) + \rho < \mu(E) + 3\rho.$$

10.1.21 Theorem

Let A_1, A_2 be μ-measurable subsets of X. Then $H = A_1 \cup A_2$ is μ-measurable.

Proof. Take $\delta > 0$ and closed E_j and open U_j with $E_j \subseteq A_j \subseteq U_j$ and $\mu(U_j \setminus E_j) < \delta$. Let $K = E_1 \cup E_2$. Then $H \setminus K \subseteq (U_1 \setminus E_1) \cup (U_2 \setminus E_2)$ and

$$\mu^+(H) \leq \mu^+(K) + \mu^+(H \setminus K) \leq \mu^+(K) + \mu^+(U_1 \setminus E_1) + \mu^+(U_2 \setminus E_2) <$$

$$< \mu^+(K) + 2\delta = \mu^-(K) + 2\delta \leq \mu^-(H) + 2\delta.$$

It follows at once that the union of finitely many μ-measurable sets is μ-measurable.

10.1.22 Lemma

Let A_1, A_2 be pairwise disjoint μ-measurable subsets of X. Then

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2).$$

Proof. First, we have

$$\mu(A_1 \cup A_2) = \mu^+(A_1 \cup A_2) \leq \mu^+(A_1) + \mu^+(A_2) = \mu(A_1) + \mu(A_2)$$

by Lemma 10.1.19. Next, choose closed sets B_j with $B_j \subseteq A_j$. Then $B_1 \cap B_2 = \emptyset$ and so

$$\mu(B_1) + \mu(B_2) = \mu(B_1 \cup B_2) \leq \mu^-(A_1 \cup A_2) = \mu(A_1 \cup A_2).$$

Since B_1 and B_2 are arbitrary we get

$$\mu(A_1) + \mu(A_2) = \mu^-(A_1) + \mu^-(A_2) \leq \mu(A_1 \cup A_2).$$

This lemma obviously extends to the union of finitely many disjoint μ-measurable sets.

10.1.23 Theorem

Let $A_j, j \in \mathbb{N}$, be pairwise disjoint μ-measurable sets, with union B. Then B is μ-measurable and

$$\mu(B) = \sum \mu(A_j).$$

Proof. We know that

$$\mu^+(B) \leq \sum \mu^+(A_j) = \sum \mu(A_j).$$

But, if N is finite,

$$\sum_{j=1}^{N} \mu(A_j) = \mu^+\left(\bigcup_{j=1}^{N} A_j\right) = \mu^-\left(\bigcup_{j=1}^{N} A_j\right) \leq \mu^-(B)$$

and the result follows on letting $N \to \infty$.

10.1.24 Theorem

The μ-measurable subsets of X form a σ-algebra. All Borel subsets of X are μ-measurable.

Proof. We already know that if A is μ-measurable then so is $X \setminus A$, and that finite unions of μ-measurable sets are μ-measurable. If A_j are μ-measurable for $j \in \mathbb{N}$ we just set

$$E_1 = A_1, \quad E_{n+1} = A_{n+1} \setminus \bigcup_{j=1}^{n} A_j.$$

Now $X \setminus A_{n+1}$ is μ-measurable by Lemma 10.1.14, and so is $\bigcup_{j=1}^{n} A_j$, by Lemma 10.1.21, and hence so is $(X \setminus A_{n+1}) \cup \bigcup_{j=1}^{n} A_j$, the complement of which is E_{n+1}.

Then the union B of the A_j is the union of the pairwise disjoint sets E_j and so is μ-measurable, with

$$\mu(B) = \sum \mu(E_j) \leq \sum \mu(A_j).$$

This gives our σ-algebra of subsets of X, and each open set belongs to this σ-algebra by Lemma 10.1.15.

10.2 Properties of the harmonic measure

We have now established Theorem 10.1.4. The measure μ has been constructed, for a fixed z_0, and the next step is to investigate what happens as z_0 varies.

For each z in D, we construct the measure $\mu = \mu_z$. For every Borel subset A of X we define

$$\omega(z, A, D) = \mu_z(A),$$

and we know that $\omega(z, A, D)$ is the infimum of $\omega(z, U, D)$, and the supremum of $\omega(z, E, D)$, over all open U and closed E with $E \subseteq A \subseteq U$.

If A is closed then we have already seen that $\omega(z, A, D)$ is a harmonic function of z on D. The same is true if A is open, because

$$\omega(z, A, D) = 1 - \omega(z, X \setminus A, D).$$

The next theorem shows that $\omega(z, A, D)$ is harmonic for every Borel subset A of X, thus justifying the term harmonic measure.

10.2.1 Theorem

Let A be a Borel subset of X. Then $\omega(z, A, D)$ is harmonic on D.

Proof. Take z_0 in D and $\delta > 0$. Choose closed E, open U, with

$$E \subseteq A \subseteq U, \quad \omega(z_0, E, D) > \omega(z_0, A, D) - \delta, \quad \omega(z_0, U, D) < \omega(z_0, A, D) + \delta.$$

If z is sufficiently close to z_0 then, since $\omega(z, E, D)$ and $\omega(z, U, D)$ are harmonic and so continuous, we have

$$\omega(z, A, D) \leq \omega(z, U, D) \leq \omega(z_0, A, D) + 2\delta, \quad \omega(z, A, D) \geq \omega(z, E, D) \geq \omega(z_0, A, D) - 2\delta,$$

and this proves that $\omega(z, A, D)$ is continuous on D.

Next, if r is small and positive,

$$\omega(z_0, A, D) > \omega(z_0, U, D) - \delta = \frac{1}{2\pi} \int_0^{2\pi} \omega(z_0 + re^{it}, U, D) dt - \delta \geq \frac{1}{2\pi} \int_0^{2\pi} \omega(z_0 + re^{it}, A, D) dt - \delta$$

and

$$\omega(z_0, A, D) < \omega(z_0, E, D) + \delta = \frac{1}{2\pi} \int_0^{2\pi} \omega(z_0 + re^{it}, E, D) dt + \delta \leq \frac{1}{2\pi} \int_0^{2\pi} \omega(z_0 + re^{it}, A, D) dt + \delta.$$

Since δ is arbitrary, we see that $\omega(z, A, D)$ has the mean value property on D and so is harmonic there. The same proof obviously works for any subset A of X such that A is μ_z-measurable for every z in D.

10.2.2 Corollary

Let E be a Borel subset of X. If $\omega(z, E, D) = 0$ for some $z \in D$ then $\omega(z, E, D) = 0$ for all $z \in D$.

This follows at once since $1 - \omega(z, E, D)$ has a maximum in D. We then say that E has zero harmonic measure.

10.2.3 Lemma

Let $x_0 \in X$. Then $\{x_0\}$ has zero harmonic measure.

This is by Lemma 9.2.7, since $\omega(z, \{x_0\}, D)$ has boundary values 0 except on a finite set. It follows using sub-additivity that every countable subset of X has zero harmonic measure.

10.2.4 Theorem

Let D be regular, let E be a Borel subset of X and let $f = \chi_E$. Then $\omega(z, E, D) = v_f(z)$ on D, where v_f is the Perron function defined in §9.1.

Thus the conclusion of 10.1.9(c) extends to all Borel subsets of X. For a more general result see Theorem 10.3.3.

Proof. First let $E \subseteq V \subseteq X$, with V open (i.e. a relatively open subset of X). Let $u \in U(f)$, where $U(f)$ is the Perron family of f as in §9.1. Then, since D is semi-regular and V is open we have, for all $w \in V$,

$$\limsup_{z \to w, z \in D} (u(z) - \omega(z, V, D)) \leq 0$$

by 10.1.9(a), and the same is true for all $w \in X \setminus V$, since $u \in U(f)$. Thus Lemma 9.2.7 gives, for every $z \in D$,

$$u(z) \leq \omega(z, V, D).$$

Applying Theorem 10.1.4(ii) we obtain $u(z) \leq \omega(z, E, D)$, and taking the supremum over $u \in U(f)$ gives $v_f(z) \leq \omega(z, E, D)$.

Now let $K \subseteq E$ be compact. Then $\omega(z, K, D)$ belongs to $U(f)$, and so

$$\omega(z, K, D) \leq v_f(z).$$

Taking the supremum over K we get $\omega(z, E, D) \leq v_f(z)$ by Theorem 10.1.4(iii).

10.2.5 Definition

For a Borel measurable $f : X \to \mathbb{R}^*$ (this means that the set $\{x \in X : f(x) < y\}$ is a Borel set for every $y \in \mathbb{R}$), define

$$M(z, f) = \int_X f(x) d\mu_z(x).$$

Here μ_z is the measure constructed for z. Clearly if $f(x)$ is a constant c then $M(z, f) = c\mu_z(X) = c$.

10.2.6 Theorem

Let $f : X \to [0, \infty]$ be Borel measurable. Then $M(z, f)$ is either harmonic or identically ∞ on D.

Proof. If s is a real-valued simple Borel function on X then $M(z, s)$ is a linear combination of harmonic measures of Borel sets and so Theorem 10.2.1 shows that $M(z, s)$ is harmonic Now just take non-negative simple Borel functions s_n such that $0 \le s_1 \le s_2 \le \ldots$ and $s_n \to f$ pointwise on X. Then

$$M(z, s_n) = \int_X s_n(x) d\mu_z(x) \to \int_X f(x) d\mu_z(x) = M(z, f)$$

by the monotone convergence theorem. The result now follows from Harnack's theorem. It follows at once that if f is a bounded Borel measurable function on X then $M(z, f)$ is harmonic on D.

10.2.7 Lemma

Let f be a bounded Borel measurable function on X, and let f be continuous at the regular boundary point x_0 of D. Then

$$\lim_{z \to x_0, z \in D} M(z, f) = f(x_0).$$

Proof. Assume without loss of generality that $f(x_0) = 0$, and take $\delta > 0$. Assume $|f| \le M < \infty$ on X. Take an open subset U of X with $x_0 \in U$ and $|f(x)| < \delta$ on U. Then

$$\left| \int_U f(x) d\mu_z(x) \right| \le \delta \mu_z(U) \le \delta$$

while

$$\left| \int_{X \setminus U} f(x) d\mu_z(x) \right| \le M\omega(z, X \setminus U, D) \to 0$$

as $z \to x_0$.

10.2.8 Corollary

Let g be a bounded function on X which is continuous at all but finitely many $x \in X$. Then $v_g(z) = L(z, g) = M(z, g)$. Next, let f be upper semi-continuous on X. Then $M(z, f)$ agrees with the harmonic extension $u(z)$ of f defined in Theorem 10.1.6.

Proof. We have $L(z, g) = v_g(z)$ by definition, and

$$\lim_{z \to x_0, z \in D} (v_g(z) - M(z, g)) = 0$$

for all but finitely many $x_0 \in X$. Now apply Lemma 9.2.7.

Next, take continuous $f_n \downarrow f$ on X and let $u_n(z) = L(z, f_n)$. Then Theorem 10.1.6 gives

$$u(z) = \lim u_n(z) = \lim M(z, f_n) = M(z, f),$$

the last step using the monotone convergence theorem.

10.2.9 Theorem: the principle of harmonic measure

Let u be subharmonic and bounded above on D. Let f be a Borel function, bounded above on X, such that

$$\phi(x) = \limsup_{z \to x, z \in D} u(z) \leq f(x)$$

for all $x \in X \setminus E$, where E has harmonic measure 0. Then

$$u(z) \leq \int_X f(x) d\mu_z(x)$$

on D.

Proof. We know from Lemma 8.3.4 that ϕ is upper semi-continuous on X. Take continuous f_n decreasing pointwise to ϕ on X, and write $u_n(z) = v_{f_n}(z) = L(z, f_n) = M(z, f_n)$. For all but finitely many $x \in X$ we have

$$\limsup_{z \to x, z \in D} (u(z) - u_n(z)) \leq 0,$$

and so, since $u - u_n$ is bounded above, we get $u(z) \leq u_n(z)$ on D. Letting $n \to \infty$ we have

$$\int_X (f_1 - f_n)(x)\, d\mu_z(x) \to \int_X (f_1 - \phi)(x)\, d\mu_z(x)$$

by the monotone convergence theorem and so

$$u(z) \leq \lim u_n(z) = \lim \int_X f_n(x)\, d\mu_z(x) = \int_X \phi(x)\, d\mu_z(x) \leq \int_X f(x)\, d\mu_z(x).$$

The following example shows that the principle may fail for unbounded functions. For $D = D(0, 1)$ and $u(z) = \mathrm{Re}\left(\frac{1+z}{1-z}\right)$ we may take $f = 0$, but u is positive on D.

10.2.10 "Two-constants" theorem

Let E_j be finitely many pairwise disjoint Borel subsets of X, with union X. Let u be a function subharmonic and bounded above on D, and let $M_j \in \mathbb{R}$ be such that

$$\limsup_{z \to x, z \in D} u(z) \leq M_j \quad \text{for} \quad x \in E_j.$$

Then

$$u(z) \leq \sum_j M_j \omega(z, E_j, D), \quad z \in D.$$

Proof. Assume first that all M_j are positive. Take open U_j with $E_j \subseteq U_j$ and set

$$v(z) = \sum_j M_j \omega(z, U_j, D).$$

Then for all but finitely many $x \in U_j$ we have

$$\lim_{z \to x, z \in D} \omega(z, U_j, D) = 1, \quad \limsup_{z \to x, z \in D} (u(z) - v(z)) \leq 0.$$

So $u(z) \leq v(z)$ on D. Using (ii) of Theorem 10.1.4 we get

$$u(z) \leq \sum_j M_j \omega(z, E_j, D).$$

If any M_j is negative, take a large positive M. Then we get

$$u(z) + M \leq \sum_j (M_j + M) \omega(z, E_j, D) = M + \sum_j M_j \omega(z, E_j, D).$$

10.2.11 Comparison principle

Let D_1, D_2 be semi-regular domains in $\mathbb{C}$, with $D_1 \subseteq D_2$, and let $X_j = \partial_\infty D_j$. Let $E \subseteq X_1 \cap X_2$ be a Borel subset of X_1 and of X_2. Then

$$\omega(z, E, D_1) \leq \omega(z, E, D_2).$$

Proof. Let F be a closed subset of X_1 and U an open subset of X_2 such that $F \subseteq E \subseteq U$. Let

$$u(z) = \omega(z, F, D_1) - \omega(z, U, D_2).$$

Let $z \to x \in X_1$ with $z \in D_1$. If $x \notin F$ and x is regular for D_1 then $\omega(z, F, D_1) \to 0$ and so $\limsup u(z) \leq 0$. If $x \in F$ then $x \in U$ and so if x is regular for D_2 we get $\omega(z, U, D_2) \to 1$ and again $\limsup u(z) \leq 0$. This means that $\limsup_{z \to x, z \in D_1} u(z) \leq 0$ for all but finitely many $x \in X_1$ and so $u \leq 0$ on D_1. Now take the supremum over F and infimum over U.

10.2.12 Conformal invariance

Let D_1, D_2 be semi-regular domains in $\mathbb{C}$, and let $X_j = \partial_\infty D_j$. Let B_j be a Borel subset of X_j and let $f : D_1 \cup B_1 \to D_2 \cup B_2$ be continuous, such that $f : D_1 \to D_2$ is analytic and f maps B_1 into B_2. Suppose further that at most finitely many x in B_1 are such that $f(x)$ is an irregular boundary point of D_2. Then

$$\omega(z, B_1, D_1) \leq \omega(f(z), B_2, D_2).$$

Proof. Take a compact subset F of $X_2 \setminus B_2$ and an open subset V of X_1 containing $X_1 \setminus B_1$. Define

$$u(z) = \omega(f(z), F, D_2), \quad z \in D_1.$$

Then u is harmonic on D_1. We assert that

$$\limsup_{z \to x, z \in D_1} (u(z) - \omega(z, V, D_1)) \leq 0 \tag{10.2}$$

for all but finitely many $x \in X_1$. Let $x \in X_1$ be regular. If x is in V, then $\omega(z, V, D_1) \to 1$ as $z \to x$ and so (10.2) holds. If x is not in V then x is in B_1 and so $f(x) \in B_2$. In particular, for all but finitely many $x \in X_1 \setminus V$ we have $f(z) \to f(x) \notin F$ and $u(z) \to 0$ as $z \to x$. This proves (10.2). Thus $u(z) \leq \omega(z, V, D_1)$ on D_1. Now (ii) and (iii) of Theorem 10.1.4 give

$$\omega(f(z), F, D_2) \leq \omega(z, X_1 \setminus B_1, D_1)$$

and

$$\omega(f(z), X_2 \setminus B_2, D_2) \leq \omega(z, X_1 \setminus B_1, D_1).$$

10.3 Comparing the harmonic measure and the Perron function

Let D be a semi-regular domain, let E be a Borel subset of X and let $f = \chi_E$. It is natural to ask whether the harmonic measure $\omega(z, E, D)$ agrees on D with the Perron function $v_f(z)$ defined in §9.1, and the following leads to a more general version of 10.1.9(c) and Theorem 10.2.4.

10.3.1 Lemma

Let E be a finite subset of X, and let $g : X \to \mathbb{R}$ with $g = 0$ on $X \setminus E$. Then $v_g = 0$ on D.

Proof. The function g is continuous on $X \setminus E$, and v_g is bounded and harmonic on D. By Theorem 9.1.7, we have $v(z) \to 0$ as $z \to w$ from within D, for all but finitely many $w \in X$. Since a finite set has zero harmonic measure, we obtain $v_g \leq 0$ on D from the two-constants theorem, and $-v_g \leq 0$ in the same way.

10.3.2 Lemma

Let E be a finite subset of X, and let f and g be bounded functions on X with $f \leq g$ on X and $f = g$ off E. Then $v_f = v_g$ on D.

Proof. Write $f = g + h$, where $h \leq 0$ on X and $h = 0$ off E. If $u_1 \in U(g)$ and $u_2 \in U(h)$ then $u_1 + u_2 \in U(f)$ and so (as in Lemma 9.1.6)

$$v_g + v_h \leq v_f \leq v_g.$$

Since $v_h = 0$ by the previous lemma the result follows.

10.3.3 Theorem

Let D be a semi-regular domain, let E be a Borel subset of X and let $f = \chi_E$. Then $\omega(z, E, D) = v_f(z)$ on D, where v_f is the Perron function defined in §9.1.

Thus the conclusion of 10.1.9(c) extends to all Borel subsets of X, even for semi-regular domains.

Proof. First let $E \subseteq V \subseteq X$, with V open (i.e. a relatively open subset of X). Let $u \in U(f)$, where $U(f)$ is the Perron family of f as in §9.1. Then, since D is semi-regular and V is open we have, for all but finitely many $w \in V$,

$$\limsup_{z \to w, z \in D} (u(z) - \omega(z, V, D)) \leq 0$$

by 10.1.9(a), and the same is true for all $w \in X \setminus V$, since $u \in U(f)$. Thus Lemma 9.2.7 gives, for every $z \in D$,

$$u(z) \leq \omega(z, V, D).$$

Applying Theorem 10.1.4(ii) we obtain $u(z) \leq \omega(z, E, D)$, and taking the supremum over $u \in U(f)$ gives $v_f(z) \leq \omega(z, E, D)$.

Now let $K \subseteq E$ be compact, and let Z be the (finite) set of non-regular boundary points. Set $g = f + \chi_Z$. Since $\omega(z, K, D)$ tends to 0 as z tends in D to any regular point in $X \setminus K$, we see that $\omega(z, K, D)$ belongs to $U(g)$, and so

$$\omega(z, K, D) \leq v_g(z) = v_f(z),$$

using the previous lemma. Taking the supremum over K we get $\omega(z, E, D) \leq v_f(z)$ by Theorem 10.1.4(iii).

Chapter 11

Jordan domains and boundary behaviour

11.1 Introduction

The first part of this chapter describes a fairly simple analytic proof of the Jordan curve theorem, which roughly-speaking states that a simple closed curve in $\mathbb{C}$ divides its complement in $\mathbb{C}^*$ into two components, each a domain without holes. The proof is taken from

Topology in the complex plane

by A. Browder, Amer. Math. Mthly. 107 (2000), 393-401.

A Jordan arc is a continuous one-one function $g : [a, b] \to \mathbb{C}$.

A Jordan curve is a continuous one-one function $g : T \to \mathbb{C}$, in which $T = \{e^{it} : 0 \leq t \leq 2\pi\}$.

In either case, the image H is compact and, since $\mathbb{C}$ is Hausdorff, the inverse function is continuous and g is a homeomorphism.

Falconer, *Geometry of Fractal Sets*, p.115 gives a continuous $f : [0, 1] \to \mathbb{R}$ whose graph (obviously a Jordan arc) has Hausdorff dimension in $(1, 2)$.

11.1.1 Preliminaries

Let U be an open subset of $\mathbb{C}$. For $x \in U$ the component C_x of U containing x is the union of all open subsets of U each containing x. If $y \in C_x$ then $C_y = C_x$. Also each C_x is open, and its boundary is a subset of ∂U. The number of components is countable.

From now on in this chapter, X will always be a compact subset of $\mathbb{C}$. The complement of X in $\mathbb{C}$ consists of countably many components, one of which is unbounded.

As usual $C(X)$ will denote the set of all continuous $f : X \to \mathbb{C}$, with L_∞ distance $\rho(f, g) = \sup\{|f(x) - g(x)| : x \in X\}$.

11.1.2 Some groups

For a compact subset X of $\mathbb{C}$ let $C^*(X)$ be the set of all continuous (non-vanishing) $f : X \to \mathbb{C} \setminus \{0\}$. Obviously $C^*(X)$ is a multiplicative group.

Next,

$$e^{C(X)} = \{e^f : f \in C(X)\}$$

is clearly a subgroup of $C^*(X)$. It is the collection of functions in $C^*(X)$ which have a continuous logarithm.

We can then form the quotient group

$$H_X = C^*(X)/e^{C(X)}.$$

This is the same as defining an equivalence relation on $C^*(X)$ by $f \sim g$ iff $f/g \in e^{C(X)}$, and then a multiplication on the classes given by $[f].[g] = [fg]$. The collection of equivalence classes is H_X.

Note that if ϕ is a homeomorphism from X onto a compact $Y \subseteq \mathbb{C}$, then H_X and H_Y are isomorphic via

$$[g]_{H_Y} \to [g(\phi)]_{H_X}.$$

To see this, obviously

$$[g_1][g_2] = [g_1 g_2] \to [g_1(\phi)g_2(\phi)] = [g_1(\phi)][g_2(\phi)]$$

while if $[g]$ is sent to the identity of H_X then $g(\phi) = e^u$ on X and so $g = e^{u(\phi^{-1})}$ on Y.

11.1.3 Lemma

If $f, g \in C(X)$ with $|g| < |f|$ on X then $f \sim f + g$.

We remark that this result is similar in statement and proof to Rouché's theorem.

Proof. We have $\mathrm{Re}(1 + g/f) > 0$ on X. Hence the principal logarithm of $1 + g/f$ is continuous on X.

11.1.4 Lemma

Let $f \in C^*(X)$. Then there exists $\delta > 0$ such that $g \in C^*(X)$ and $g \sim f$ for all $g \in C(X)$ with $\rho(f, g) < \delta$.

Proof. The function $|f|$ has a positive minimum δ on X. Thus if $\rho(f, g) < \delta$ we have $|g - f| < \delta$ on X and so $g \neq 0$, while writing $g = f + (g - f)$ shows that $g \sim f$.

11.1.5 Lemma

Let $f \in C^*(X)$. Then $[f] = \{g \in C^*(X) : g \sim f\}$ is an open and closed subset of $C^*(X)$.

Proof. We've just seen that $[f]$ is open. Now suppose that $g \not\sim f$. Then there exists $\delta > 0$ such that $\rho(g, h) < \delta$ implies $h \sim g$ and so $h \not\sim f$.

11.1.6 Theorem

Let $f, g \in C^*(X)$. Then $f \sim g$ if and only if there exists a continuous $F : X \times [0,1] \to \mathbb{C} \setminus \{0\}$ such that

$$F(z,0) = f(z), \quad F(z,1) = g(z)$$

for all z in X.

Proof. If $f \sim g$ we can write $f = ge^h$ with $h \in C(X)$. We then just set $F(z,t) = f(z)e^{-th(z)}$.

Now suppose that such a F exists. Define f_t by $f_t(z) = F(z,t)$. Now $J = \{t \in [0,1] : f_t \sim f_0\}$ is open and closed in $[0,1]$, by Lemma 11.1.5, and is non-empty. So $J = [0,1]$.

11.1.7 Corollary

Let X have the property that $z \in X$ implies that $tz \in X$ for all $t \in [0,1]$ (starlike about 0). Then $C^*(X) = e^{C(X)}$.

Obviously this applies if X is the closed unit disc, or is $[0,1]$ (which is henceforth always denoted I).

Proof. Just define $F(z,t) = f(tz)$. This shows that $f \sim 1$ for every $f \in C^*(X)$.

11.1.8 Lemma

$e^{C(X)}$ is the maximal connected subset of $C^*(X)$ containing 1.

Proof. Suppose $A \subseteq C^*(X)$, and A properly contains $e^{C(X)}$. Since $e^{C(X)} = [1]$ is an open and closed subset of $C^*(X)$, we may partition A as $e^{C(X)}, A \setminus e^{C(X)}$ and both are relatively open. So A is not connected.

It remains only to show that $e^{C(X)}$ is connected. But if $f, g \in C(X)$ then $t \to e^{tg+(1-t)f}$ is a continuous function from I to $e^{C(X)}$, sending 0 to e^f and 1 to e^g.

11.2 Janiszewski's theorem

11.2.1 Lemma

Let $n \in \mathbb{Z}$ and (henceforth) let $T = \{e^{i\theta} : 0 \le \theta \le 2\pi\}$. If $z^n \in e^{C(T)}$ then $n = 0$.

Proof. Suppose that $n \neq 0$ and $z^n = e^{h(z)}$ on T for some $h \in C(T)$. The principal argument $a(z) = \mathrm{Arg} z$ is continuous on $T_1 = T \setminus \{-1\}$, taking values in $(-\pi, \pi)$, and $e^{ina(z)} = z^n$ on T_1. Thus $(h(z) - ina(z))/2\pi i$ is continuous and integer-valued and so constant on T_1. This is a contradiction since h is continuous on all of T but $a(z)$ is not.

Obviously the same proof shows that if X is the circle $|z - a| = R > 0$ and n is a non-zero integer then $(z-a)^n \notin e^{C(X)}$.

11.2.2 Corollary

If $a \in \mathbb{C}, R > 0$ and n is a non-zero integer then there is no $f \in C^*(\overline{B}(a, R))$ which equals $(z-a)^n$ on $|z-a| = R$.

Here $\overline{B}(a, R)$ denotes the closed unit disc.

Proof. Assume without loss of generality that $a = 0, R = 1$. Suppose we had such an f. By Corollary 11.1.7 we have $f = e^g$ for some g continuous on the closed unit disc. So $z^n = e^{g(z)}$ on T and this contradicts Lemma 11.2.1.

11.2.3 Theorem

Let U be a bounded open subset of $\mathbb{C}$ and let $a \in U$. Let $H = \partial U$ and let n be a non-zero integer. Then there is no $f \in C^*(\overline{U})$ which equals $(z-a)^n$ on H.

Proof. We assume that $U \subseteq D(a, R)$. If f is a continuous non-zero function on the closure of U, which equals $(z-a)^n$ on H, then we extend f to the closed disc by setting $f(z) = (z-a)^n$ for z not in the closure of U. This extended function is continuous, and this contradicts Corollary 11.2.2.

11.2.4 Tietze's extension theorem

Let A be a closed subset of $\mathbb{C}$ and let $f : A \to \mathbb{R}$ be continuous. Then there is a continuous function $g : \mathbb{C} \to \mathbb{R}$ such that $g = f$ on A.

Proof. Urysohn's lemma gives us the following: if B, C are disjoint closed subsets of $\mathbb{C}$ then there exists a continuous $k : \mathbb{C} \to [0, 1]$ with $k = 0$ on B and $k = 1$ on $\mathbb{C}$. In fact, k is

$$k(z) = \frac{d(z, B)}{d(z, B) + d(z, C)}$$

with $d(z, B)$ the distance from z to B. Clearly there is then a continuous $K_0 : \mathbb{C} \to [-1/3, 1/3]$ with $K_0 = -1/3$ on B and $K_0 = 1/3$ on C.

Now let $f : A \to (-1, 1)$ be continuous. Let $B = \{x \in A : f(x) \leq -1/3\}$ and let $C = \{x \in A : f(x) \geq 1/3\}$. Then B, C are closed and there is a continuous $h_1 : \mathbb{C} \to [-1/3, 1/3]$ such that $h_1 = -1/3$ on B and $h_1 = 1/3$ on C. Thus $|h_1 - f| \leq 2/3$ on A.

We claim that there exist continuous $h_n : \mathbb{C} \to \mathbb{R}$ such that $|h_n(x)| \leq 2^{n-1}3^{-n}$ on $\mathbb{C}$ and

$$|f(x) - \sum_{j=1}^{n} h_j(x)| \leq 2^n 3^{-n}, \quad x \in A.$$

Assuming that $h_1, \ldots, h_n$ exist, we apply the first part to the function

$$(3/2)^n (f(x) - \sum_{j=1}^{n} h_j(x)) = K(x).$$

This gives $H(x)$ with $|H| \leq 1/3$ on $\mathbb{C}$ and $|H - K| \leq 2/3$ on A, and we just set $h_{n+1} = (2/3)^n H$.

The function $h(x) = \sum_{j=1}^{\infty} h_j(x)$ is then continuous on $\mathbb{C}$ and equal to f on A.

For a general f, we apply the above proof to $F = f/(1+|f|), f = F/(1-|F|)$.

11.2.5 Lemma

Let $f \in e^{C(X)}$. Then f can be extended to a continuous non-zero function on $\mathbb{C}$.

Proof. Just write $f = e^g$ on X and extend g to a continuous function on $\mathbb{C}$ using Tietze's extension theorem.

11.2.6 Notation

We continue to use X to denote a compact subset of $\mathbb{C}$. Let the distinct components be $U_0, U_1, \ldots$, with U_0 unbounded. Fix $a_k \in U_k$.

11.2.7 Lemma

Let $a, b \in U_k$. Then $z - a \sim z - b$.

If $a \in U_0$ then $z - a \sim 1$.

Proof. Take a path $\gamma : [0,1] \to U_k$ with $\gamma(0) = a, \gamma(1) = b$. Let $F(z,t) = z - \gamma(t)$. Then F is continuous and non-zero on $X \times I$ and $F(z,0) = z - a, F(z,1) = z - b$. Apply Theorem 11.1.6.

Next, choose positive R so large that $\mathrm{Re}(z+R) > 0$ on X and $R \in U_0$. Then $\log(z+R)$ (the principal log) is continuous on X, so $z + R \sim 1$. Hence $z - a \sim 1$ for all a in U_0.

11.2.8 Lemma

Let $q \in \mathbb{N}$ and let

$$f(z) = \prod_{k=0}^{q} (z - a_k)^{n_k} \tag{11.1}$$

with each $n_k \in \mathbb{Z}$. If $f \in e^{C(X)}$ then $n_k = 0$ for all $k \geq 1$.

Proof. N.B. since $a_j \in U_j$ the hypotheses assume that $\mathbb{C} \setminus X$ has at least two components. Assume that $f = e^h$ on X. By Tietze's extension theorem we can assume that h is continuous on $\mathbb{C}$. Thus there is a continuous non-zero function F on $\mathbb{C}$ (Lemma 11.2.5), which equals f on X. Now set

$$g(z) = (z - a_0)^{n_0} \prod_{k=2}^{q} (z - a_k)^{n_k}$$

(with $g(z) = (z - a_0)^{n_0}$ if $q = 1$). Then g is continuous and non-zero on $X \cup U_1$. So there is a continuous non-zero function F/g on $X \cup U_1$, which equals $(z - a_1)^{n_1} = f/g$ on X and so on ∂U_1. This forces $n_1 = 0$, by Theorem 11.2.3.

Remark: this is one of the two key steps of the method. The other will be to show that each $F \in C(X)$ has $F \sim f$ for some f of the form (11.1).

11.2.9 Corollary

If $C^*(X) = e^{C(X)}$ then $\mathbb{C} \setminus X$ is connected.

For otherwise there is a bounded component U_1, and $z - a_1 \not\sim 1$.

11.2.10 Corollary

If X is a Jordan arc (i.e. X is homeomorphic to $I = [0,1]$) then $\mathbb{C} \setminus X$ is connected.

Proof. We have $C^*(I) = e^{C(I)}$ by Corollary 11.1.7. If $f \in C^*(X)$ we have $f(\phi) = e^g \in e^{C(T)}$, where $\phi : T \to X$ is the homeomorphism. Thus $f = e^{g(\phi^{-1})}$ and $C^*(X) = e^{C(X)}$.

Note that $\partial U_0 = X$ in this case. Otherwise some $x \in X$ is not a limit point of U_0 and so some $D(x,r)$ is a subset of X. But then $D(x,r)$ is the homeomorphic image of a connected relatively open subset of I and so of an interval. This is a contradiction, as deleting a point disconnects an interval but not $D(x,r)$.

11.2.11 Corollary

If $a, b \in \mathbb{C} \setminus X$ and $z - a \sim z - b$ then a and b lie in the same U_j.

Proof. Otherwise we have, with $j \neq k$,

$$\frac{z - a_j}{z - a_k} \sim \frac{z - a}{z - b} \sim 1$$

on X. This contradicts Lemma 11.2.8.

11.2.12 Janiszewski's theorem

Let X, Y be compact subsets of $\mathbb{C}$ and let $a, b \in \mathbb{C}$ have the property that a and b both lie in the same component of $\mathbb{C} \setminus X$, and a and b both lie in the same component of $\mathbb{C} \setminus Y$. If $X \cap Y$ is connected then a and b both lie in the same component of $\mathbb{C} \setminus (X \cup Y)$.

Proof. We can write $(z-a)/(z-b) = e^{g(z)}$ on X, and $(z-a)/(z-b) = e^{h(z)}$ on Y. Thus $e^{g-h} = 1$ on $X \cap Y$. But then $g - h$ is constant on $X \cap Y$ and, adding an integer multiple of $2\pi i$ to h if necessary, we get $(z-a)/(z-b) \sim 1$ on $X \cup Y$.

11.3 Convolutions and Runge's theorem

11.3.1 Convolutions

We describe the following ideas for $\mathbb{C}, \mathbb{R}^2$, but they work equally well in $\mathbb{R}^n$. A function $g : \mathbb{C} \to \mathbb{R}$ is said to have compact support if there exists a positive real R with $g(z) = 0$ for $|z| > R$. We say g is C^n if g has continuous n'th order partial derivatives on all of $\mathbb{C}$. Note that if g is C^1 then writing

$$g(x,y) - g(a,b) = g(x,y) - g(x,b) + g(x,b) - g(a,b)$$

and using the MVT shows that g is continuous. C^∞ is the intersection of the C^n and we say g is C_0^n if g is C^n with compact support.

Let g be a real-valued bounded measurable function, zero off the compact set Y, and let h be measurable and locally integrable i.e. the Lebesgue integral

$$\int_{\overline{B}(0,R)} |h(z)|dxdy$$

is finite for every $R > 0$. Define H by

$$H(w) = (h * g)(w) = \int_{\mathbb{C}} h(z)g(w-z)dxdy = \int_{\mathbb{C}} h(w-z)g(z)dxdy = \int_{Y} h(w-z)g(z)dxdy.$$

Fact 1: *if h has compact support so has H.*

This is because if $|w|$ is large enough then $h(w-z) = 0$ for all z in Y.

Fact 2: *if g is continuous then so is H.*

Fix w and take $\varepsilon_1 > 0$. Take $S > 0$ such that $g(u-z) = 0$ for $|u-w| < 1$ and $|z| > S$. Since g is uniformly continuous we can take $\delta \in (0,1)$ such that $|g(u-z) - g(w-z)| < \varepsilon_1$ for $|u-w| < \delta$ and for all z. Hence for these u we have

$$|H(u) - H(w)| \leq \varepsilon_1 \int_{\overline{B}(0,S)} |h(z)|dxdy.$$

Fact 3: *if g is C^1 then so is H, and $\partial_j H = h * \partial_j g$.*

To see this, fix w and S as before, let t be real, small and non-zero, and write

$$\frac{H(w+t) - H(w)}{t} = \int_{\overline{B}(0,S)} h(z)\frac{g(w+t-z) - g(w-z)}{t}dxdy.$$

But the MVT gives

$$\frac{g(w+t-z) - g(w-z)}{t} = \partial_1 g(c),$$

and this is uniformly bounded, since $\partial_1 g$ is continuous with compact support. Taking any sequence $t_n \to 0$ and using the dominated convergence theorem we get the result.

11.3.2 Lemma

Let $Y \subseteq V \subseteq \mathbb{C}$, with Y compact and V open. Then there exists a C_0^∞ function $F : \mathbb{C} \to [0,1]$ with $F = 1$ on Y and $F = 0$ off V.

Proof. Take a small positive t and a non-negative C^∞ function ϕ, vanishing off $D(0,t)$, and with

$$\int_{\mathbb{C}} \phi(z)dxdy = \int_{D(0,t)} \phi(z)dxdy = 1.$$

For example, we may take

$$\phi(x+iy) = \lambda \exp(-1/(t^2 - x^2 - y^2)), \quad x^2 + y^2 < t^2,$$

with $\phi(z) = 0$ otherwise, and with some suitable positive λ.

Now let $h(z) = 1$ for z in a t-neighbourhood of Y (i.e. the distance from z to Y is less than t), with $h = 0$ otherwise. Set

$$F(w) = \int_{\mathbb{C}} h(z)\phi(w-z)dxdy = \int_{\mathbb{C}} h(w-z)\phi(z)dxdy = \int_{D(0,t)} h(w-z)\phi(z)dxdy.$$

Obviously

$$0 \leq F(w) \leq \int_{D(0,t)} \phi(z)dxdy = 1.$$

Also F is C^∞, since ϕ is. Next, if $w \notin V$ then $h(w-z) = 0$ for $|z| < t$, provided t was chosen small enough. Thus $F(w) = 0$. Finally, if $w \in Y$ then $h(w-z) = 1$ for $|z| < t$ and so $F(w) = 1$.

11.3.3 Lemma

Let Y be a compact subset of $\mathbb{C}$ and let $F : Y \to \mathbb{C}$ be continuous. Let $\delta > 0$. Then there exists a C_0^∞ function G such that $|G(w) - F(w)| \leq \delta$ for all w in Y.

Proof. By Tietze's extension theorem we can assume that F is continuous on all of $\mathbb{C}$. By multiplying by a C_0 function which is 1 on Y we can assume F has compact support. Thus F is uniformly continuous on $\mathbb{C}$. Now take a small positive t, so small that $|F(w) - F(w-z)| < \delta$ for all $w \in Y$ and all $z \in D(0,t)$. Take ϕ as in the previous lemma and set

$$G(w) = \int_{\mathbb{C}} F(z)\phi(w-z)dxdy = \int_{D(0,t)} F(w-z)\phi(z)dxdy.$$

This convolution is $\mathbb{C}_0^\infty$ since ϕ is (and F has compact support) and for $w \in Y$ we have

$$F(w) - G(w) = \int_{D(0,t)} (F(w) - F(w-z))\phi(z)dxdy$$

with modulus at most

$$\delta \int_{D(0,t)} \phi(z)dxdy = \delta.$$

11.3.4 The $\overline{\partial}$ operator

We define

$$\overline{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

If h is complex-valued and has continuous first partials on a domain D in $\mathbb{C}$ and $\overline{\partial}h \equiv 0$ then h is analytic, by Cauchy-Riemann.

Also Green's theorem

$$\int_{\partial A} Pdx + Qdy = \int_A Q_x - P_y dxdy$$

(the integral around ∂A once in the positive sense) can be written in the form

$$\int_{\partial A} g(z)dz = \int_{\partial A} g(z)dx + ig(z)dy = \int_A ig_x - g_y dxdy = 2i\int_A \overline{\partial}g dxdy.$$

11.3.5 Lemma

Let g be a complex-valued C^1 function on $\mathbb{C}$, with compact support. Define Tg by

$$Tg(w) = \frac{1}{\pi}\int_{\mathbb{C}} \frac{g(z)}{w-z}dxdy = \frac{1}{\pi}\int_{\mathbb{C}} \frac{g(w-z)}{z}dxdy. \tag{11.2}$$

Then Tg is C^1 on $\mathbb{C}$, and we have

$$\overline{\partial}(Tg) = T(\overline{\partial}g) = g.$$

Proof. The fact that Tg is C^1 and

$$\overline{\partial}(Tg) = T(\overline{\partial}g)$$

follows from §11.3.1 (note that $1/z$ is locally integrable).

Fix w, and choose R so large that $g(z) = 0$ for $|w-z| > R/2$, and take a small positive δ. Let

$$A = \{z \in \mathbb{C} : \delta < |w-z| < R\},$$

and let B be the boundary of A, described once in the positive sense (keeping the interior to the left). Now

$$(T(\overline{\partial}g))(w) = \lim_{\delta\to 0}\frac{1}{\pi}\int_A \frac{(\overline{\partial}g)(z)}{w-z}dxdy.$$

But

$$\frac{(\overline{\partial}g)(z)}{w-z} = \overline{\partial}\left(\frac{g(z)}{w-z}\right)$$

and so Green's theorem gives

$$(T(\overline{\partial}g))(w) = \lim_{\delta\to 0}\frac{1}{\pi}\int_A \overline{\partial}\left(\frac{g(z)}{w-z}\right)dxdy = \lim_{\delta\to 0}\frac{1}{2\pi i}\int_B \frac{g(z)}{w-z}dz = g(w).$$

11.3.6 Lemma

Let $K \subseteq U \subseteq \mathbb{C}$ with K compact and U open. Let g be analytic on U, and let $\delta > 0$. Then there exists a rational function R, with no poles in K, such that $|g(w) - R(w)| < \delta$ for all $w \in K$.

Proof. Replacing g by ϕg, where ϕ is a $\mathbb{C}_0^\infty$ function as in Lemma 11.3.2 which is 1 on some $\{z : \mathrm{dist}(z,K) < t\}$, with $t > 0$, we can assume that g is in fact C_0^1. The function

$$h = \overline{\partial}g$$

is then continuous with compact support, and there is a compact set H not meeting K such that $h = 0$ off H. Lemma 11.3.5 then gives, for $w \in K$,

$$g(w) = (Th)(w) = \frac{1}{\pi}\int_H \frac{h(z)}{w-z}dxdy.$$

To form $R(w)$ we then just approximate the integral by a Riemann sum.

We outline the details. We can assume H is a union of closed rectangles. Let d be the (positive) distance from K to H, and let $|h| \le M$ on H. Take $\rho > 0$ and partition H into closed rectangles H_k, disjoint apart from boundary, and so small that $|z - z_k| < \rho$ and $|h(z) - h(z_k)| < \rho$ for all $z \in H_k$, with z_k the centre of H_k. Let $m(H_k)$ be the area of H_k, and let

$$R(w) = \frac{1}{\pi} \sum \frac{h(z_k)}{w - z_k} m(H_k).$$

Then, for $w \in K$,

$$|g(w) - R(w)| \leq I_1 + I_2,$$

in which

$$I_1 = \left| \sum \int_{H_k} \frac{h(z) - h(z_k)}{w - z} dxdy \right| \leq \rho d^{-1} m(H)$$

and

$$I_2 = \left| \sum \int_{H_k} h(z_k) \left(\frac{1}{w - z} - \frac{1}{w - z_k} \right) dxdy \right| =$$

$$= \left| \sum \int_{H_k} h(z_k) \left(\frac{z - z_k}{(w - z)(w - z_k)} \right) dxdy \right| \leq M m(H) \rho d^{-2}.$$

Remark: this is a weak version of Runge's theorem, which states that R can be chosen so that all its poles lie in the set $\{a_k\}$, in which as before each $a_k \in U_k$ and $U_0, \dots$ are the components of $\mathbb{C} \setminus K$. To see this, suppose that R has a pole at $b \in U_k$. Join b to a_k by a path σ in U_k, and so by a finite sequence of points z_j such that

$$|z_j - z_{j-1}| < \frac{1}{4} \text{dist}(K, \sigma).$$

Now just use the fact that if $|A - B| < \frac{1}{4} \text{dist}(B, K)$ then for $z \in K$ and $m \in \mathbb{N}$ we can write

$$(z - A)^{-m} = (z - B + B - A)^{-m} = (z - B)^{-m}(1 + (B - A)/(z - B))^{-m}$$

and expand out in negative powers of $z - B$, using the fact that $|z - B| > 2|A - B|$ for all z in X.

11.4 Proof of the Jordan curve theorem

11.4.1 Theorem

Let $f(z) \in C^*(X)$. Let N, with $0 \leq N \leq \infty$, be the number of bounded components of $\mathbb{C} \setminus X$. Then there exist integers n_k, all but finitely many of them 0, such that

$$f(z) \sim \prod_{k=1}^{N} (z - a_k)^{n_k}.$$

The n_k are uniquely determined by f. If $N = 0$ then $f \sim 1$.

Proof. Suppose first that f is a rational function of z with no zeros or poles in X. Write $f = P/Q$ and factorize P and Q. If f has a zero of multiplicity m at $\alpha \in U_k$, Lemma 11.2.7 gives $(z-\alpha)^m \sim (z-a_k)^m$, and $(z - \alpha)^m \sim 1$ if $k = 0$.

Now let f be any function in $C^*(X)$. It suffices to show that there exists a rational function R, with no zeros or poles in X, such that $f \sim R$. By Lemmas 11.1.4 and 11.3.3 there exists a function $p \in C_0^\infty$ such that $f \sim p$ i.e. $f/p \in e^{C(X)}$. Let $t > 0$, and let

$$V_j = \{z : \text{dist}(z, X) < jt\}, \quad j = 1, 2, 3.$$

Since $p \neq 0$ on X, we have $p \neq 0$ on V_3, provided t is small enough. We can then use Lemma 11.3.2 to form a function ϕ which is C^∞, is 1 on V_1, and 0 off V_2. Define g by

$$g = \phi \frac{\overline{\partial} p}{p}$$

on V_3, with $g = 0$ off V_3. Then g is C^∞, and Lemma 11.3.5 gives us a C^∞ function $h = Tg$ with $\overline{\partial} h = g$. Let $F = pe^{-h}$. Then $f \sim F$. Also, F is C^∞ and on V_1 we have

$$\overline{\partial} F = e^{-h}\overline{\partial} p - (\overline{\partial} h)pe^{-h} = e^{-h}\overline{\partial} p - gpe^{-h} = 0.$$

So F is analytic on V_1. Let $s = \min\{|F(z)| : z \in X\} > 0$ and choose a rational R, with $|R(w) - F(w)| < s$ on X. Then Lemma 11.1.4 gives $R \sim F$ and so $R \sim F \sim f$.

Finally, the uniqueness follows from Lemma 11.2.8.

11.4.2 Theorem

Let X, Y be compact subsets of $\mathbb{C}$ and let $\phi : X \to Y$ be a homeomorphism. Then $\mathbb{C} \setminus X, \mathbb{C} \setminus Y$ have the same number of components.

Proof. Obviously it suffices to prove that the number of components of $\mathbb{C} \setminus Y$ is at most that of $\mathbb{C} \setminus X$. If $\mathbb{C} \setminus X$ is connected then $C^*(X) = e^{C(X)}$ by Theorem 11.4.1 and so for any $f \in C^*(Y)$ we have $g = f(\phi) \in C^*(X)$ and $g = e^h, h \in C(X)$, which gives $f = e^{h(\phi^{-1})} \in e^{C(Y)}$.

Now suppose that the bounded components of $\mathbb{C} \setminus X$ are $U_1, \ldots, U_n$, and that $\mathbb{C} \setminus Y$ has distinct bounded components $V_1, \ldots, V_{n+1}$. Choose $a_k \in U_k, b_j \in V_j$, and set

$$f_k(z) = z - a_k, \quad F_k = f_k(\phi^{-1}), \quad g_j(z) = z - b_j.$$

For each j, the function $g_j(\phi)$ is in $C^*(X)$. So we can find integers $q_{j,k}, 1 \leq k \leq n$, such that

$$g_j(\phi) \sim \prod_{k=1}^{n} f_k^{q_{j,k}},$$

by which we mean that

$$g_j(\phi) \prod_{k=1}^{n} f_k^{-q_{j,k}} \in e^{C(X)},$$

and so

$$g_j \sim \prod_{k=1}^{n} F_k^{q_{j,k}}.$$

The matrix with entries $q_{j,k}$ has rank at most n over $\mathbb{Q}$, and so we can find integers $m_1, \ldots, m_{n+1}$, not all zero, such that

$$m_1 q_{1,k} + m_2 q_{2,k} + \ldots + m_{n+1} q_{n+1,k} = 0$$

for all k. But this gives (in $C^*(Y)$)

$$\prod_{j=1}^{n+1} g_j^{m_j} \sim 1,$$

which is a contradiction.

11.4.3 The Jordan curve theorem

Let X be a Jordan curve i.e. homeomorphic to T. Then $\mathbb{C} \setminus X$ has two components U_0, U_1, each with boundary X.

We only need show that each U_j has boundary X. Obviously $\partial U_j \subseteq X$. If $\partial U_j \neq X$ for some j then ∂U_j is a subset of a Jordan arc $Y \subseteq X$. So we can join a_0 to a_1 by a path γ in $\mathbb{C} \setminus Y$ which does not meet ∂U_j. This says that both a_m are in U_j and this is obviously a contradiction.

11.5 Boundary extension for Jordan domains

11.5.1 Theorem

Let D be a bounded domain in $\mathbb{C}$ and let f map $\Delta = D(0,1)$ conformally (i.e. one-one analytically) onto D. Then the following are equivalent:

(i) f has a continuous extension to the closed unit disc $\Delta \cup T$;

(ii) ∂D is a closed curve;

(iii) $J = \partial D$ has the following property: to each $\varepsilon > 0$ corresponds $\delta > 0$ such that if $w_1, w_2 \in J$ with $|w_1 - w_2| < \delta$ then w_1, w_2 lie in a compact connected subset B of J with the diameter of B less than ε.

Proof. This is adapted from Pommerenke's book *Boundary behaviour of conformal maps*.

(i) implies (ii) is easy. If f extends continuously to the closed disc then $f(T)$ is a closed curve, and standard results show that $f(T) \subseteq \partial D$. Further, $f(T) = \partial D$, since $f(\Delta \cup T)$ is closed.

(ii) implies (iii). Let $\lambda : [0,1] \to \mathbb{C}$ be any curve. Let $\varepsilon > 0$. Since λ is uniformly continuous, we may partition $[0,1]$ into closed subintervals I_k such that $J_k = \lambda(I_k)$ has diameter less than $\varepsilon/2$. Let $\delta > 0$ be such that if $J_k \cap J_j = \emptyset$ then the distance from J_k to J_j is at least 2δ. Then if w_1, w_2 lie on λ and $|w_1 - w_2| < \delta$ we may write $w_1 \in J_j, w_2 \in J_k, J_j \cap J_k \neq \emptyset$, and we may take $B = J_j \cup J_k$.

(iii) implies (i). We set out to prove that f is *uniformly* continuous on Δ. The extension to T then follows easily. We may assume that $f(0) = 0$ and hence that $D(0,s) \subseteq D$ for some $s > 0$. Let ε be positive, small compared to s, and choose $\delta \in (0, \varepsilon)$ as in (iii).

Let $z_0 \in \Delta$ and let ρ be small and positive. For $\rho \leq r \leq \rho^{1/2}$ let $\gamma_r = S(z_0, r) \cap \Delta$ (recall that $S(z_0, r)$ is the circle of centre z_0, radius r), and let $L(r)$ be the length (possibly infinite) of $f(\gamma_r)$. Parametrizing γ_r by $z = z_0 + re^{i\theta}$ the Cauchy-Schwarz inequality gives

$$L(r)^2 = \left(\int_{\gamma_r} |f'(z)| r d\theta \right)^2 \leq \left(\int_{\gamma_r} r d\theta \right) \left(\int_{\gamma_r} |f'(z)|^2 r d\theta \right) \leq 2\pi r \left(\int_{\gamma_r} |f'(z)|^2 r d\theta \right).$$

Dividing by r and integrating from ρ to $\rho^{1/2}$ we thus have, with A the (finite) area of D,

$$\int_\rho^{\rho^{1/2}} \frac{L(r)^2}{r} dr \leq 2\pi \int_\rho^{\rho^{1/2}} \int_{\gamma_r} |f'(z)|^2 r d\theta dr \leq 2\pi A.$$

Hence there exists r with $\rho \le r \le \rho^{1/2}$ such that

$$L(r)^2 \le \frac{8\pi A}{\log 1/\rho}.$$

Assume that ρ is chosen so small that $L(r) \le \delta/2$, and let $C = \gamma_r$. Then $f(C)$ is a curve in D of length at most $\delta/2$.

Suppose first that $C = S(z_0, r)$. Thus $f(C)$ is a closed curve in D. Choose $z^* \in C$. Then $|f(z) - f(z^*)| \le \delta/2 \le \varepsilon/2$ for all z on C, and the same holds for $z \in D(z_0, r)$, by the maximum principle. Thus $|f(z) - f(z_0)| \le \varepsilon$ for $z \in D(z_0, r)$.

We assume henceforth that $C \ne S(z_0, r)$. Thus C is a circular arc whose closure, when described counter-clockwise, joins end-points $z_1, z_2 \in T$. Since $f(C)$ has finite length, the limits

$$w_j = \lim_{z \to z_j, z \in C} f(z) \in J$$

exist, and $|w_1 - w_2| \le \delta/2$. Thus there exists a compact connected subset K of J of diameter at most ε, with $w_1, w_2 \in K$. Let $M = K \cup f(C)$. Then M has diameter at most 2ε.

Let $z', z'' \in D(z_0, r) \cap \Delta$. We assert that the distance between $f(z'), f(z'')$ is at most 16ε. If this is not the case then at least one of these points, without loss of generality $w' = f(z')$, lies at distance at least 4ε from M. Hence there is a path joining $0 = f(0)$ to w' and not meeting M. There is also a path joining 0 to w' in D, and so not meeting J. Since $J \cap M = K$ is connected, Janiszewski's theorem gives a path σ from 0 to w' and not meeting $M \cup J$. Thus σ is a path in D, not meeting $f(C)$, and so $f^{-1}(\sigma)$ is a path in Δ from 0 to z', not meeting C. By the definition of C, this is impossible.

We have thus shown, in both cases, that $|f(z) - f(z_0)| \le 16\varepsilon$ for $|z - z_0| < \rho$, with ρ independent of z_0. Thus f is uniformly continuous on Δ, as asserted.

11.5.2 Remark

If D_1 is a simply connected proper subdomain in $\mathbb{C}$ then the following method may be used to map D_1 conformally onto a bounded simply connected domain. Take $a \in D_1, b \in \mathbb{C} \setminus D_1$. The function $u(z) = (z - b)^{1/2}$ is analytic and one-one on D_1. Further, if $z_0 \in D_1$ then $u(z)$ does not take the value $-u(z_0)$ on D_1. So there is some $r > 0$ such that $D(u(a), r) \subseteq u(D_1)$ and $|u(z) + u(a)| \ge r$ on D_1, so that $v(z) = (u(z) + u(a))^{-1}$ is bounded and conformal on D_1.

11.5.3 Theorem

Let D be a Jordan domain in $\mathbb{C}$ i.e. a bounded simply connected domain in $\mathbb{C}$ such that $J = \partial D$ is a Jordan curve. Let f map Δ conformally (i.e. analytically and one-one) onto D. Then f has a continuous extension mapping $\Delta \cup T$ one-one onto $D \cup J$.

Proof. Since J is a curve f has a continuous extension mapping $\Delta \cup T$ onto $D \cup J$, by Theorem 11.5.1, and it remains only to show that the extended function is homeomorphic. Certainly f maps T onto J. Let $a \in J$. The set

$$E = \{z \in T : f(z) \ne a\}$$

is a relatively open subset of T. Further, $F = T \setminus E$ has measure zero, as may be seen by applying the two-constants theorem to the function $u(z) = \log |f(z) - a|$, which is subharmonic and bounded above on Δ.

Assume that $|z_1| = |z_2| = 1, z_1 \neq z_2, f(z_1) = f(z_2) = a$. The set $T \setminus \{z_1, z_2\}$ consists of two open arcs of T, denoted A_1, A_2, and we choose $u_j \in A_j$ such that $v_j = f(u_j) \neq a$. Join z_1 to z_2 by a straight line segment L.

Now $\Gamma = f(L) \cup \{a\}$ is a Jordan curve in $D \cup J$, with $f(L) \subseteq D$. So $\mathbb{C} \setminus \Gamma$ has two components, U_0, U_1, with U_0 unbounded.

Claim 1: *Each v_j lies in U_0.*

By rotating D if necessary we may assume that there exists $b \in J$ with $\mathrm{Re}(b) \leq \mathrm{Re}(w)$ for all w in J and $\mathrm{Re}(b) < \mathrm{Re}(a)$. Thus the line N given by $z = b - t, t \geq 0$, does not meet $D \cup \{a\}$. Hence v_j can be joined to points of arbitrarily large modulus by a path not meeting $D \cup \{a\}$ (follow an arc of J to b and then follow N), and so not meeting Γ. This proves Claim 1.

Since $f(L) \subseteq \Gamma = \partial U_1$, there are points arbitrarily close to L whose images under f lie in U_1. Thus we can choose a simple curve M from u_1 to u_2, consisting of two straight line segments, such that M lies in Δ (apart from its end-points), $f(M)$ meets U_1, and M intersects L at precisely one point, v. If $z^* \in M \setminus \{v\}$ then the line segment M^* from z^* to one of the u_j is a path not meeting $L \cup f^{-1}(\{a\})$ and so $f(M^*)$ is a path from v_j to $f(z^*)$ not meeting Γ. Thus $f(z^*) \in U_0$, for all $z^* \in M \setminus \{v\}$, contradicting the assumption that $f(M)$ meets U_1.

11.5.4 Extending the conformal mapping to the plane

Let X be a Jordan curve in $\mathbb{C}$. Assume that $0 \in U_1$. The Riemann mapping theorem gives an analytic homeomorphism $f : D(0,1) \to U_1$ with $f(0) = 0$, and we have seen that f extends to a homeomorphism of $|z| \leq 1$ onto $U_1 \cup X$. By first using the map $z \to 1/z$, and applying the Riemann mapping theorem again, we obtain a homeomorphism g of $\{z : 1 \leq |z| \leq \infty\}$ onto $X \cup U_0 \cup \{\infty\}$ with $g(\infty) = \infty$. Thus $\phi(z) = g^{-1}(f(z))$ is a homeomorphism of T onto itself, and so $f(z) = g(\phi(z))$ for $|z| = 1$. If we set

$$G(z) = f(z), \quad |z| < 1,$$

with $G(\infty) = \infty$ and

$$G(z) = g(|z|\phi(z/|z|)), \quad 1 < |z| < \infty,$$

then f has been extended to a homeomorphism G of the extended plane onto itself, fixing ∞.

11.6 Totally disconnected sets

A non-empty set H is called totally disconnected if its only connected subsets are singleton sets $\{x\}, x \in H$. Compact totally disconnected sets arise for example as the Julia set of $z^2 + c$, for any c not in the Mandelbrot set (in particular for $|c| > 2$).

11.6.1 Lemma

Let E be a connected subset of $\mathbb{C}^* = \mathbb{C} \cup \{\infty\}$, with $\infty \in E$. Then all components U_j of $\mathbb{C}^* \setminus E$ are simply connected domains in $\mathbb{C}$.

Proof. Let γ be a closed piecewise smooth contour in some U_j. Then the winding number $n(\gamma, z)$, which is analytic and integer-valued off γ, is 0 for $|z|$ large enough. Since E is connected, $\mathbb{C} \cap E$ must be unbounded, and so we have $n(\gamma, z) = 0$ for all $z \in E$. Hence if $U_k \neq U_j$ is bounded, we have $n(\gamma, z) = 0$ for all z in U_k, by the maximum principle. Finally, if $U_k \neq U_j$ is unbounded, we have $n(\gamma, z) = 0$ for large z in U_k and so for all z in U_k.

11.6.2 Lemma

Let U be a domain in $\mathbb{C}$. Then there exists a simply connected domain V such that $U \subseteq V \subseteq \mathbb{C}$ and $\partial V \subseteq \partial U$.

(All boundaries here are with respect to $\mathbb{C}^*$).

Proof. Let E be the component of $\mathbb{C}^* \setminus U$ containing ∞. Let V be the component of $F = \mathbb{C}^* \setminus E$ containing U. Then V is a simply connected domain, and $\partial V \subseteq \partial F = \partial E \subseteq \partial U$.

11.6.3 Theorem

Let X be a totally disconnected compact subset of $\mathbb{C}$. Then $\mathbb{C} \setminus X$ is connected.

Proof. Assume that $\mathbb{C} \setminus X$ has a bounded component U. Then there is a simply connected domain V with $U \subseteq V \subseteq \mathbb{C}$, and $W = \partial V \subseteq \partial U \subseteq X$, so that V is bounded. Since V is simply connected, W is a connected subset of X, and so at most a singleton, which is plainly impossible.

Note that Browder's paper suggests an alternative approach to this, based on partitioning X into a disjoint union of compact sets of small diameter.

11.7 Boundary behaviour of analytic functions

11.7.1 Schwarz reflection principle

Let $r > 0$ and $D^+ = \{z : |z| < r, \mathrm{Im}(z) > 0\}$ and $D^- = \{z : |z| < r, \mathrm{Im}(z) < 0\}$ and let u be harmonic on $D^+ = \{z : |z| < r, \mathrm{Im}(z) > 0\}$ with $\lim_{z \to x} u(z) = 0$ for every $x \in (-r, r)$. Then u extends to a harmonic function on $D(0, r)$ satisfying $u(z) = -u(\overline{z})$.

Proof. Let $0 < s < r$ and let $f(t) = u(se^{it})$ on $(0, \pi)$. Extend f to an odd continuous function on $[-\pi, \pi]$. Let U be the Poisson integral of $f(t)$ in $D(0, s)$. Then $U = 0$ on $(-s, s)$ since the Poisson kernel is even when z is real, while f is odd. Thus $U - u = 0$ on $\{z : |z| < s, \mathrm{Im}(z) > 0\}$ since $U - u \to 0$ as z tends to any point on the boundary.

The reflection principle has a very powerful consequence for the boundary behaviour of analytic functions. Suppose that D is a domain in $\mathbb{C}$ and that $D^+ \subseteq D$ and $D^- \cap D = \emptyset$, so that $I = (-r, r) \subseteq \partial D$. Let f be analytic on D such that $u(z) = \mathrm{Re}(f(z)) \to 0$ as $z \to x \in I$. Thus u extends across I to a harmonic function on $D(0, r)$, which is in turn the real part of an analytic function g on $D(0, r)$. Thus g is an analytic extension of f to $D(0, r)$.

We can also handle "corners" as follows. Suppose that $r > 0$ and $0 < \alpha < 2\pi$ and that D is a domain such that

$$D_1 = \{z : 0 < |z| < r, 0 < \arg z < \alpha\} \subseteq D, \quad \{z : 0 < |z| < r, \alpha < \arg z < 2\pi\} \cap D = \emptyset.$$

Thus the corner

$$J = \{0\} \cup \{z : 0 < |z| < r, \arg z = 0, \alpha\}$$

forms part of the boundary of D. Suppose that f is analytic on D and that $Re(f(z)) \to 0$ as $z \to \zeta \in J$. By setting $w = z^{\pi/\alpha}$ and $g(w) = f(z)$, we obtain an analytic function g on a semi-disc. We extend g to the disc and this extends f continuously to J.

Chapter 12

Homotopy and analytic continuation

12.1 Homotopy

Let S be a path-connected topological space (i.e. any two points a, b in S can be joined by a continuous $f : [0,1] \to S$ with $f(0) = a, f(1) = b$), and let x_0, x_1 be points in S (possibly the same). Let γ, σ be two paths in S, both defined on $[0,1] = I$, and both going from x_0 to x_1 i.e. $\sigma(0) = \gamma(0) = x_0, \sigma(1) = \gamma(1) = x_1$.

Suppose that S is a disc, or is $\mathbb{R}^n$, or is some kind of space that can be thought of as having "no holes". Then it's reasonable to believe that we could continuously deform γ into σ by a family of paths in S. What we mean by this is that there is a family of paths $h_u(t), 0 \leq u \leq 1$, in S such that:

(i) each h_u is defined on $[0,1]$, with $h_u([0,1])$ contained in S, and joins x_0 to x_1;

(ii) we have $h_0 = \gamma$ and $h_1 = \sigma$;

(iii) if u is close to v then h_u is close to h_v. More precisely, if we define the function $H(t,u)$ by $H(t,u) = h_u(t)$ then this H will be continuous on $[0,1] \times [0,1]$ (with the usual metric on $I^2 = [0,1] \times [0,1]$).

H is called a homotopy function and we say that γ is homotopic to σ in S.

Example 1: In $\mathbb{C}$, let $\gamma(t) = \cos t, 0 \leq t \leq \pi$, and let $\sigma(t) = \cos t + i \sin t, 0 \leq t \leq \pi$. If we put $h_u(t) = \cos t + iu \sin t, 0 \leq t \leq 1$, we see that γ can be continuously deformed into σ.

Example 2: Let D be a star domain with star centre w. Let $\gamma : [0,1] \to D$ be a closed path, with $\gamma(0) = w$. Then γ is homotopic to the constant path σ given by $\sigma(t) = w$. For h_u we can just take $uw + (1-u)\gamma(t) = w + (1-u)(\gamma(t) - w)$.

Example 3: if $S = \{z : 1 < |z| < 3\}$ then intuitively it's easy to see that the circle $\gamma(t) = 2e^{2\pi i t}$ is not homotopic to the constant path $\sigma(t) = 2$ in S (although they are homotopic in $\mathbb{C}$).

Where no confusion might arise, we drop the phrase "in S".

Remark: we can define homotopy for paths both defined on $[a,b]$ (so that H is then defined on $[a,b] \times [0,1]$) but the formulation above is most usual.

12.1.1 Fact

Homotopy is an equivalence relation.

Clearly each γ is homotopic to itself, with homotopy function $H(t,u) = \gamma(t)$.

If γ is homotopic to σ with family of paths h_u then σ is homotopic to γ: just put $g_u(t) = h_{1-u}(t)$, and $G(t,u) = g_u(t)$ is continuous on I^2.

Finally, if γ is homotopic to σ with family of paths f_u, and σ is homotopic to τ with family of paths g_u, then we form a family of paths h_u which continuously deform γ into τ just by putting $h_u = f_{2u}$ for $0 \le u \le 1/2$ and $h_u = g_{2u-1}$ for $1/2 \le u \le 1$.

Note that $h_{1/2} = f_1 = \sigma = g_0 = g_{2(1/2)-1}$.

12.1.2 Products of paths

Given two paths $\gamma, \sigma : [0,1] \to S$ with $\sigma(0) = \gamma(1)$ we can define a path which is 'γ followed by σ' by

$$(\gamma\sigma)(t) = \gamma(2t), \quad 0 \le t \le 1/2,$$

$$(\gamma\sigma)(t) = \sigma(2t-1), \quad 1/2 \le t \le 1.$$

(sometimes called a "product').

12.1.3 Fact

If γ_0 is homotopic to γ_1 and σ_0 is homotopic to σ_1, with homotopy functions $F(t,u) = \gamma_u(t), G(t,u) = \sigma_u(t), 0 \le t, u \le 1$ respectively, and if the σ_j start where the γ_j finish, then $\gamma_0\sigma_0$ is homotopic to $\gamma_1\sigma_1$.

Just use the paths $h_u = \gamma_u\sigma_u$. Note that $\gamma_u(1) = \sigma_u(0)$.

12.1.4 The "inverse" path

Let $\gamma : [0,1] \to S$ be a path. We can define γ^{-1} (or γ backwards) by $\gamma^{-1}(t) = \gamma(1-t)$. Then $\gamma\gamma^{-1}(t) = \gamma(2t)$ for $0 \le t \le 1/2$ and $\gamma\gamma^{-1}(t) = \gamma^{-1}(2t-1) = \gamma(2-2t)$ for $1/2 \le t \le 1$ and this is a closed curve. It is homotopic to a constant curve as follows:

Let $w = \gamma(0)$ and define $\eta(t) = w$ for $0 \le t \le 1$. Then η is a constant curve, and is homotopic to $\gamma\gamma^{-1}$. Put $h_u(t) = \gamma(ut)$. Then as t goes from 0 to 1, $h_u(t)$ goes along γ as far as $\gamma(u)$. Now put $g_u = h_u h_u^{-1}$.

Clearly $g_0 = \eta, g_1 = \gamma\gamma^{-1}$ and $g_u(t)$ is continuous on I^2. What g_u does is to go along γ as far as $\gamma(u)$, and then retrace its steps back to w.

Note also that if γ is homotopic to σ, with family of paths $p_u(t)$, then using the paths $p_u^{-1}(t) = p_u(1-t)$ we see that γ^{-1} is homotopic to σ^{-1}.

12.1.5 Re-scaling and homotopy

If $\sigma(t) = \gamma(g(t))$ where $g : [0,1] \to [0,1]$ is continuous, non-decreasing and onto, then σ is homotopic to γ.

We say that σ is a re-scaling of γ, and a re-scaling of a path is homotopic to the original path.

To see this, just put $h_u(t) = \gamma(ut + (1-u)g(t))$ (and note that $ut + (1-u)g(t) \in [0,1]$ for $u, t \in [0,1]$).

(I prefer not to use the term re-parametrization, which is reserved for the case where g above is strictly increasing.)

12.1.6 Corollary

If $\rho(t) \equiv \gamma(0)$ (constant curve) then $\rho\gamma$ (ρ followed by γ) is homotopic to γ. Similarly, if $\tau(t) \equiv \gamma(1)$ then $\gamma\tau$ is homotopic to γ.

Proof: $\rho\gamma(t) = \gamma(h(t))$ and $\gamma\tau(t) = \gamma(k(t))$. Here $h(t) = 0$ for $0 \le t \le 1/2$ while $h(t) = 2t - 1$ for $1/2 \le t \le 1$. Similarly $k(t) = 2t$ for $0 \le t \le 1/2$ while $k(t) = 1$ for $1/2 \le t \le 1$.

This leads to a useful fact.

12.1.7 Fact

Let $\sigma, \tau : [0,1] \to S$ be paths from x_0 to x_1. Then σ is homotopic to τ iff $\sigma\tau^{-1}$ (which is σ followed by τ backwards) is homotopic to a constant path.

Why? If σ is homotopic to τ, then $\sigma\tau^{-1}$ is homotopic to $\tau\tau^{-1}$, and we know that the last path is homotopic to a constant path.

If $\sigma\tau^{-1}$ is homotopic to a constant path ρ, then $(\sigma\tau^{-1})\tau$ is homotopic to $\rho\tau$ and so to τ. But $(\sigma\tau^{-1})\tau$ is a re-scaling of $\sigma(\tau^{-1}\tau)$, and so is homotopic to $\sigma\lambda$, where λ is a constant path, and so to σ.

Now we can make a group.

12.1.8 The fundamental group

Let S be a topological space, and let $x_0 \in S$. Consider the family H of all closed paths $\lambda : [0,1] \to S$ starting and finishing at x_0. For a given γ, let $[\gamma]$ be the equivalence class of members λ of H s.t. λ is homotopic to γ.

We can define a multiplication by $[\gamma][\sigma] = [\gamma\sigma]$ (equivalence class of the product path γ "followed by" σ). This is well defined as, if γ_1 is homotopic to γ_2 and σ_1 is homotopic to σ_2, then $\gamma_1\sigma_1$ is homotopic to $\gamma_2\sigma_2$.

Define $I(t) = x_0, 0 \le t \le 1$. Then $[I\gamma] = [\gamma]$ and $[\gamma I] = [\gamma]$ for every γ in H. Also, $[\gamma^{-1}\gamma] = [I]$ for every γ in H. So we have a *group*, with identity $[I]$, and with $[\gamma]^{-1} = [\gamma^{-1}]$, called the *fundamental group* $\pi(x_0, S)$.

1. Note that this group might not be Abelian. However, the multiplication *is* associative, because $\rho(\sigma\tau)$ is a re-scaling of $(\rho\sigma)\tau$.

2. If $x_1 \in S$ then the groups $\pi(x_0, S), \pi(x_1, S)$ are isomorphic. Choose a *fixed* path Λ from x_0 to x_1. For $[\gamma]$ in $\pi(x_1, S)$ we define $T([\gamma])$ to be $[\Lambda\gamma\Lambda^{-1}]$ (which is in $\pi(x_0, S)$).

This is well defined. This is because, if γ and σ both start and finish at x_1 and are homotopic, then $\Lambda\gamma\Lambda^{-1}$ is homotopic to $\Lambda\sigma\Lambda^{-1}$.

Note that $\Lambda\Lambda^{-1}$ is homotopic to the constant path $I(t) \equiv x_0$. Thus $[\Lambda\Lambda^{-1}]$ is the identity in $\pi(x_0, S)$ and $[\Lambda^{-1}\Lambda]$ is the identity in $\pi(x_1, S)$. Thus $T([\sigma][\gamma]) = T([\sigma\gamma]) = [\Lambda\sigma\gamma\Lambda^{-1}] = [\Lambda\sigma\Lambda^{-1}\Lambda\gamma\Lambda^{-1}]$ and this equals $[\Lambda\sigma\Lambda^{-1}][\Lambda\gamma\Lambda^{-1}] = T([\sigma])T([\gamma])$.

Next, T is one-one, as $T([\gamma]) = T([\sigma])$ implies that $\Lambda\gamma\Lambda^{-1}$ is homotopic to $\Lambda\sigma\Lambda^{-1}$, and so $[\gamma] = [\Lambda^{-1}\Lambda\gamma\Lambda^{-1}\Lambda] = [\Lambda^{-1}\Lambda\sigma\Lambda^{-1}\Lambda] = [\sigma]$.

Also T is onto, as $T([\Lambda^{-1}\sigma\Lambda]) = [\Lambda\Lambda^{-1}\sigma\Lambda\Lambda^{-1}] = [\sigma]$.

So T is a group isomorphism and so we often talk just of the fundamental group $\pi(S)$.

12.1.9 Fact

If σ and τ are homotopic paths in X, with family of paths h_u continuously deforming σ into τ, and $f : X \to Y$ is continuous, then $f(\sigma), f(\tau)$ are homotopic paths in Y .

(Just use the paths $f(h_u(t))$.

If f is a homeomorphism from X to Y (i.e. f is one-one and onto, and both f and the inverse f^{-1} are continuous), then paths μ, ν from x_0 to x_1 are homotopic iff $f(\mu)$ and $f(\nu)$ are.

Also, if f is a homeomorphism from X to Y then the fundamental group of X is isomorphic to the fundamental group of Y just by setting $T([\gamma])$ to be the class $[f(\gamma)]$. This gives $T([\gamma])T([\lambda]) = [f(\gamma)][f(\lambda)] = [f(\gamma)f(\lambda)] = [f(\gamma\lambda)] = T([\gamma\lambda])$.

12.1.10 Simple connectivity in terms of homotopy

As remarked before, if S is a suitable space with "no holes" we would expect that two paths in S starting and finishing at the same points would be homotopic.

Let S be a path-connected topological space. We say that S is HSC (homotopy simply connected) if every closed curve $\gamma : [0, 1] \to S$ (i.e. $\gamma(0) = \gamma(1)$) is homotopic to the constant curve η which satisfies $\eta(t) = \gamma(0)$ for all t.

This is the case if and only if, for any pair of curves σ, τ in S such that $\sigma(0) = \tau(0)$, $\sigma(1) = \tau(1)$, it is the case that σ is homotopic to τ. This is by Corollary 12.1.6.

This is also the same as saying that the fundamental group of S is trivial (identity only).

12.1.11 Lemma

A domain D in $\mathbb{C}$ is called convex if, for every z, w in D, the straight line segment $sz+(1-s)w, 0 \le s \le 1$ is contained in D. Convex domains are HSC.

Proof: given γ, σ such that $\gamma(0) = \sigma(0)$, $\gamma(1) = \sigma(1)$, just set $F(t,u) = (1-u)\gamma(t) + u\sigma(t)$.

Alternatively: convex means that any point in the domain can be used as a star centre. Use the method earlier for star domains.

12.1.12 Lemma

Let γ be a closed curve in $\mathbb{C}$ of diameter $L > 0$. Let $w \in \mathbb{C}$ with $\operatorname{dist}\{w, \gamma\} > 8L$. Then γ is null-homotopic with respect to w.

Here the diameter of a set E means $\sup\{|z - z'| : z, z' \in E\}$, and null-homotopic with respect to w means homotopic to a constant in $\mathbb{C} \setminus \{w\}$.

Proof. The hypotheses imply that the variation of $\arg(z - w)$ is less than π on γ. Hence γ lies in a sector with vertex at w (some domain $a < \arg(z - w) < b < a + \pi$) and this is a convex region.

12.1.13 Lemma

Let γ be a closed curve in $\mathbb{C}$ which is null-homotopic with respect to $z_0 \in \mathbb{C}$, and let $\operatorname{dist}\{z_0, \gamma\} \geq 2\delta > 0$. Let $z_1 \in D(z_0, \delta)$. Then γ is null-homotopic with respect to z_1.

Proof. Assume without loss of generality that $z_0 = 0$. There is some homotopy function $H(t,u) : [0,1]^2 \to \mathbb{C} \setminus \{0\}$, with $H(t,0) = \gamma(t)$ and $H(t,1)$ a constant path. There exists $\rho > 0$ such that $|H(t,u)| \geq \rho$ for all (t,u) in $[0,1]^2$. If $\rho \geq \delta$ then $H(t,u)$ never equals z_1 and γ is automatically null-homotopic with respect to z_1.

Assume now that $\rho < \delta$, and set

$$\phi(z) = z\left|\frac{z}{2\delta}\right|^\lambda, \quad |z| < 2\delta, \quad \phi(z) = z, \quad |z| \geq 2\delta,$$

in which $\lambda > 0$ is chosen so that

$$\left(\frac{\rho}{2\delta}\right)^\lambda = \delta.$$

Then ϕ is a homeomorphism of $\mathbb{C}$ onto itself, and $\phi(\gamma(t)) = \gamma(t)$. Thus the composition $K(t,u) = \phi(H(t,u))$ gives a homotopy from γ to a constant path with, for all (t,u), $|K(t,u)| \geq \delta$ and so $K(t,u) \neq z_1$.

12.1.14 Theorem

Let γ be a closed path in $\mathbb{C}$. For $z \in \mathbb{C} \setminus \gamma$, let $\phi(z) = 0$ if γ is null-homotopic with respect to z, with $\phi(z) = 1$ otherwise. Then ϕ is continuous on $\mathbb{C} \setminus \gamma$.

Proof. We have already seen in Lemma 12.1.13 that if $\phi(z_0) = 0$ then $\phi(z) = 0$ for z near z_0. Suppose now that $\phi(z_0) = 1$ and $\operatorname{dist}\{z_0, \gamma\} = \delta$ (necessarily positive). Then for $z_1 \in D(z_0, \delta/8)$ we have $\operatorname{dist}\{z_1, \gamma\} \geq 7\delta/8$ and, again by Lemma 12.1.13, we must have $\phi(z_1) = 1$.

12.1.15 The Riemann sphere $\mathbb{C} \cup \{\infty\}$ is HSC

This seems intuitively obvious, but complications arise when we note that a closed path may visit *every* point on the sphere. Let $\gamma : [0, 1] \to \mathbb{C} \cup \{\infty\}$ be any closed curve.

Case 1: the curve γ lies entirely in $|z| \leq M < \infty$. Then γ lies in the convex domain $D(0, 2M)$ and is homotopic to a constant curve, with some continuous homotopy function $G(t, u) : [0, 1]^2 \to D(0, 2M)$. This G is continuous if we regard it as a function into $\mathbb{C} \cup \{\infty\}$.

Case 2: the curve γ lies entirely in $|z| \geq c > 0$. Then the curve $1/\gamma$ lies entirely in $|z| \leq 1/c$ and so is homotopic to a constant, with some homotopy function $G(t, u)$. Using $1/G(t, u)$ we see that γ is homotopic to a constant.

Case 3: γ sometimes visits both $|z| > 2$ and $|z| < 1/2$. without loss of generality $|\gamma(0)| \geq 1$ (else look at $1/\gamma$). Suppose we have an interval $[a, b]$ on which $|\gamma(t)| \leq 3/4$, with $|\gamma(a)| = |\gamma(b)| = 3/4$ and $|\gamma(t)| \leq 1/2$ for some t with $a < t < b$. We can form a curve $\sigma : [a, b] \to \mathbb{C}$ such that $\sigma(a) = \gamma(a)$ and $\sigma(b) = \gamma(b)$ and $|\sigma(t)| = 3/4$ for every t. But σ and the part of γ for $a \leq t \leq b$ both lie in $|z| \leq 3/4$ and so using a homotopy function (modified to be defined for $a \leq t \leq b, 0 \leq u \leq 1$) we can continuously deform the restriction $\gamma : [a, b] \to \mathbb{C} \cup \{\infty\}$ into σ. We do this for each such interval $[a, b]$, and we have continuously deformed γ into a closed path in $|z| \geq 3/4$, which is now homotopic to a constant.

12.2 Analytic continuation

12.2.1 Example

The function $L_0(z) = \log z = \log |z| + i \arg z$ is analytic in the domain D_0 obtained by deleting from the complex plane the non-positive real axis. Here the argument is chosen to lie in $(-\pi, \pi)$. Obviously the restriction L_1 of L_0 to the upper half plane $D_1 = \{z : \mathrm{Im}(z) > 0\}$ is also analytic.

In the same way, the function $L_2(z) = \log |z| + i \arg z$, with the argument chosen to lie in $(0, 2\pi)$, is analytic in the domain D_2 obtained by deleting from the plane the non-negative real axis.

If we start at 1 and continue $L_0(z)$ counter-clockwise around the circle $|z| = 1$, the argument increases, and $L_0 = L_1 = L_2$ in the quadrant $\pi/2 < \arg z < \pi$. Following the circle further, the argument continues to increase until, on approaching 1 again, the argument tends to 2π. Thus L_0 has been "continued' around the circle, but has not returned to its original value.

12.2.2 Analytic continuation along a path

By a function element we mean a pair (f, D), in which D is a domain in $\mathbb{C}^*$ and f is meromorphic on D. As usual, meromorphic at ∞ means that $f(1/z)$ is meromorphic at 0.

Let (f, D) be a function element and let $z_0 \in D$. Let $\gamma : [a, b] \to \mathbb{C}^*$ be a path with $\gamma(a) = z_0$ (note that continuity is with respect to the spherical metric). An analytic continuation of (f, D) along γ is a family of function elements $(f_t, D_t), a \leq t \leq b$, with the following properties.

(i) $f_a = f$ on a neighbourhood of $z_0 = \gamma(a)$.

(ii) $\gamma(t) \in D_t$ for every t in $[a, b]$.

(iii) For every t in $I = [a, b]$ there exists $\rho_t > 0$ such that the following holds. For $a \leq s \leq b, |s-t| < \rho_t$ we have $\gamma(s) \in D_t$ and $f_s = f_t$ on a neighbourhood of $\gamma(s)$.

Here a neighbourhood of z means an open set containing z. Note that in (iii) we do not require that $f_s = f_t$ on all of $D_s \cap D_t$, but this will be the case if $D_s \cap D_t$ is connected.

Strictly speaking, this is *meromorphic* continuation but, as this term is not normally used, we shall say that the analytic continuation is *finite-valued* if all the f_t map their D_t into $\mathbb{C}$ rather than $\mathbb{C}^*$.

If G, H are domains with $G \subseteq H$ then we say that a function element (g, G) admits unrestricted analytic continuation (UAC) in H if (g, G) can be analytically continued along every path in H starting in G.

12.2.3 Lemma

Suppose that the function element (f, D) is analytically continued along the path $\gamma : [a, b] \to \mathbb{C}^*$ by the family of function elements (f_t, D_t). Then $h(t) = f_t(\gamma(t)) : [a, b] \to \mathbb{C}^*$ is a path.

Proof. We need of course to show that h is continuous. Let W be a neighbourhood of $f_t(\gamma(t))$ and let $U \subseteq D_t$ be a neighbourhood of $\gamma(t)$ such that $f_t(U) \subseteq W$. If s is close enough to t then we have $\gamma(s) \in U$ and, since $f_s = f_t$ near $\gamma(s)$, we get $f_s(\gamma(s)) \in W$.

12.2.4 Theorem

Let (f, D) and (g, D) both be analytically continued along the path $\gamma : [a, b] \to \mathbb{C}^*$. If there exists $u \in [a, b] = J$ such that $f_u = g_u$ on a neighbourhood of $\gamma(u)$ then for every t in J we have $f_t = g_t$ on a neighbourhood of $\gamma(t)$.

In particular if $f = g$ then $f_1 = g_1$ on a neighbourhood of $\gamma(1)$ and so the continuation along γ is (locally) unique.

Proof of the theorem. Let E be the set of t in J such that $f_t = g_t$ on a neighbourhood $U(t)$ of $\gamma(t)$. Let $t \in E$. Then there exists $\rho > 0$ such that if $s \in J$ and $|s - t| < \rho$ then $\gamma(s) \in U(t)$ and $f_s = f_t = g_t = g_s$ on a neighbourhood of $\gamma(s)$. Thus $s \in E$ and E is relatively open.

Now let $t \in J \setminus E$. Then there exists a neighbourhood U of $\gamma(t)$ such that for all z in $V = U \setminus \{\gamma(t)\}$ we have $f_t(z) \neq g_t(z)$. If s is close enough to t we have $\gamma(s) \in U$ and $f_s = f_t$ and $g_s = g_t$ on a neighbourhood W of $\gamma(s)$. Thus for z in $W \setminus \{\gamma(t)\}$ we have $f_s(z) \neq g_s(z)$ and so $s \notin E$. Hence $J \setminus E$ is also relatively open and so, since $E \neq \emptyset$, we have $J = E$ by connectivity.

12.2.5 Critical and asymptotic values

Let $f : \mathbb{C} \to \mathbb{C}^*$ be meromorphic. A critical point z of f is a multiple point of f i.e. a pole of multiplicity at least 2 or a point where $f'(z) = 0$. A point z is critical if and only if there is no neighbourhood of z on which f is one-one. The critical values of f are the values taken at critical points.

Thus $\cos z$ has critical points $n\pi, n \in \mathbb{Z}$, and critical values ± 1, while the only critical value of $1/(e^z - 1)^2$ is ∞.

An asymptotic value of f is an element w of $\mathbb{C}^*$ such that there exists a path γ tending to infinity with $f(z) \to w$ as $z \to \infty$ on γ.

Note that asymptotic values only have relevance for transcendental functions. If f is a rational function then ∞ is a point like any other (though it may be a critical point and/or a pole), since $f(1/z)$ is meromorphic at 0.

For example, e^z has no critical values, but it has asymptotic values $0, \infty$.

Iversen's theorem (see Exercise 9.2.6) says that if f is a non-constant entire function then ∞ is always an asymptotic value of f.

12.2.6 Theorem

Let $f : \mathbb{C} \to \mathbb{C}^*$ be non-constant and meromorphic. Let $z_0 \in \mathbb{C}$ be a non-critical point of f, and let $\gamma : [0,1] \to \mathbb{C}^*$ be a path in $\mathbb{C}^*$ starting at $w_0 = f(z_0)$. Let g be that branch of the inverse function f^{-1} which is defined on a neighbourhood D of w_0 and maps w_0 to z_0. Let S be the supremum of u in $[0,1]$ such that $g = (g, D)$ admits analytic continuation along the path $\gamma : [0,u] \to \mathbb{C}^*$, the function elements g_t finite-valued. Then either (i) $g_t(\gamma(t)) \to \infty$ as $t \to S-$ and $\gamma(S)$ is an asymptotic value of f, or (ii) $g_t(\gamma(t)) \to z^*$ as $t \to S-$, with z^* a critical point of f, and $\gamma(S) = f(z^*)$ a critical value of S, or (iii) g admits finite-valued analytic continuation along $\gamma : [0,S] \to \mathbb{C}^*$.

If γ contains none of the critical and asymptotic values of f, then g admits finite-valued analytic continuation along γ.

(If f is a non-constant rational function then the same proof as below shows that continuation is possible along any γ avoiding critical values of f, although not necessarily finite-valued).

Proof. We first note that $S > 0$, because for small t we can take $D_t = D$ and $g_t = g$.

We begin by noting a consequence of the uniqueness result Theorem 12.2.4. If $0 < u \le u' \le S$ and g_t is a continuation along $\gamma : [0,u] \to \mathbb{C}^*$, while h_t is a continuation along $\gamma : [0,u'] \to \mathbb{C}^*$, then for $0 \le t \le u$ we have $g_t = h_t$ on a neighbourhood of $\gamma(t)$. In particular, $g_t(\gamma(t))$ is uniquely defined, and is a continous function from $[0,S)$ into $\mathbb{C}$. Also, for *every* u with $0 < u < S$ the continuation is possible along $\gamma : [0,u] \to \mathbb{C}^*$.

Next, for small t we have $f \circ g_t(w) \equiv w$ on a neighbourhood of $\gamma(t)$ and so, by Theorem 12.2.4 again, we have $f(g_t(\gamma(t))) = \gamma(t)$ for $0 \le t < S$. Assume that (i) and (ii) do not hold. Then it cannot be the case that $g_t(\gamma(t)) \to \infty$ as $t \to S-$ and so there exists $M > 0$ such that $|g_{t_n}(\gamma(t_n))| \le M$ for a sequence $t_n \to S-$. Take $N > M$ such that $q(f(z), \gamma(S)) > \rho > 0$ on $|z| = N$. Then for t close to S we have $|g_t(\gamma(t))| \ne N$, and it follows that $|g_t(\gamma(t))| < N$ for t close to S.

As $t \to S-$ the continuation $g_t(\gamma(t))$ stays in the compact region $|z| \le N$, and $f(g_t(\gamma(t))) \to w_1 = \gamma(S)$. Hence there exists z_1 with $f(z_1) = w_1$ and $|z_1| \le N$ such that $g_t(\gamma(t)) \to z_1$ as $t \to S-$. Since we have assumed that (ii) does not hold, it follows that z_1 is not a critical point of f and so we can choose a small neighbourhood U_1 of z_1 on which f is one-one. Let h be the inverse function f^{-1} mapping the neighbourhood $W_1 = f(U_1)$ onto U_1 with, obviously, $h(w_1) = z_1$. Since $g_t(\gamma(t)) \to z_1$ as $t \to S-$, and since S is a supremum, we may take u with $0 < u \le S$ and such that $g_u(\gamma(u)) \in U_1$, while (g, D) admits a finite-valued analytic continuation along $\gamma : [0, u] \to \mathbb{C}^*$.

Let t be close to u, with $t \leq u$. Then we have $g_t(w) \in U_1$ and $f(g_t(w)) = w \in W_1$ for w close to $\gamma(t)$. But $f(h(w)) = w$ on W_1, and f is one-one on U_1, so that $g_t(w) = h(w)$ for w close to $\gamma(t)$ and t close to u. Thus we may use the function element (h, W_1) to analytically continue (g, D) all the way along $\gamma : [0, S] \to \mathbb{C}^*$ and, if $S < 1$, along some $\gamma : [0, S'] \to \mathbb{C}^*$ with $S' > S$. This proves the first assertion of the theorem.

The second assertion is easy: the assumptions rule out (i) and (ii), so that S must be 1.

12.2.7 Corollary

Let f be non-constant and meromorphic in $\mathbb{C}$. Suppose that D is a domain in $\mathbb{C}^*$ not containing any critical or asymptotic value of f. Let $z_0 \in \mathbb{C}$ and $f(z_0) = w_0 \in D$. Then the branch of f^{-1} defined near w_0 and mapping w_0 to z_0 admits finite-valued UAC in D.

12.2.8 Corollary

Any local branch of the logarithm $\log z = \log |z| + i \arg z$ admits UAC in $\mathbb{C} \setminus \{0\}$.

12.3 The monodromy theorem

Suppose that we have a function element (f, D) and analytic continuations $(f_t, D_t), (g_t, G_t)$ along paths $\gamma_j : [0, 1] \to \mathbb{C}^*$ from z_0 to z_1. Under what circumstances will $f_1 = g_1$ on a neighbourhood of z_1? This will certainly be the case if $\gamma_1 = \gamma_2$ (Theorem 12.2.4) but we saw in Example 12.2.1 that is is possible to continue $\log z$ around a closed curve and not return to the original branch.

With the γ_j as above we will say that the continuations $(f_t, D_t), (g_t, G_t)$ *lead to the same local function element* if $f_1 = g_1$ on a neighbourhood of z_1. Note that we're not assuming that $D_1 = G_1$.

Note also that if we continue f along $\gamma : [0, 1] \to \mathbb{C}^*$, with function elements (f_t, D_t), and then back along γ^{-1}, the "final" function element f_2 will always equal f_0 near $\gamma(0)$. This is because $(f_t, D_t), 0 \leq t \leq 1$, with $(f_{2-t}, D_{2-t}), 1 \leq t \leq 2$, gives one continuation along $\gamma\gamma^{-1}$ and so in effect, by Theorem 12.2.4, the only continuation.

12.3.1 Monodromy theorem

Let G be a domain in $\mathbb{C}^*$, and let (f, D) be a function element with $D \subseteq G$, and assume that (f, D) admits UAC in G. Let $z_0 \in D$ and let $z_1 \in D$, and let γ, σ be homotopic paths from z_0 to z_1 in G. Then analytic continuation of (f, D) along γ, σ leads to the same local function element near z_1.

Proof. We have some homotopy function

$$H(t, u) = h_u(t) : I^2 \to G, \quad I = [0, 1], \quad h_0 = \gamma, \quad h_1 = \sigma.$$

Let μ be a path in I^2 from $(0, 0)$ to $(1, 1)$. Then $H(\mu)$ is a path in G from z_0 to z_1, and (f, D) admits analytic continuation along $H(\mu)$. With a slight abuse of notation, we refer to this as continuation of f along μ.

We need the following idea of *quadrisection* of a square J, which we describe only for the square

$J = I^2$, but which carries over to any square, with obvious modifications. We define stepwise paths as follows:

(i) $\mu_1(I^2)$ goes $(0,0),(0,1),(1,1)$;

(ii) $\mu_2(I^2)$ goes $(0,0),(0,1/2),(1/2,1/2),(1/2,1),(1,1)$;

(iii) $\mu_3(I^2)$ goes $(0,0),(0,1/2),(1,1/2),(1,1)$;

(iv) $\mu_4(I^2)$ goes $(0,0),(1/2,0),(1/2,1/2),(1,1/2),(1,1)$;

(iv) $\mu_5(I^2)$ goes $(0,0),(1,0),(1,1)$.

For a general square J, the construction is the same, involving vertices and the midpoint of J. Note that each $\mu_k(J)$ has length $2L$, in which L is the side-length of J, and $\mu_k = \mu_{k+1}$ except on an interval of length L. Also μ_k and μ_{k+1} together bound a square of side-length $L/2$.

Assume that the analytic continuations $(f_t, D_t), (g_t, G_t)$ of (f, D) along γ, σ do not lead to the same local function element. Since $h_u(0) = H(0, u) = z_0$ and $h_u(1) = H(1, u) = z_1$ for all u, we may continue f along $\mu_1(I^2)$ and $\mu_5(I^2)$ (on the vertical segments just use f_0, g_1 respectively). Thus continuation of f along $\mu_1(I^2), \mu_5(I^2)$ does not lead to the same local function element. Set $\gamma_0 = \mu_1(I^2), \sigma_0 = \mu_5(I^2), J_0 = I^2, K_0 = [0,2]$.

Claim 1: *For each non-negative integer n there exist paths $\gamma_n, \sigma_n : [0,2] \to I^2$, square regions J_n, and intervals K_n, with the following properties:*

(a) $J_{n+1} \subseteq J_n \subseteq I^2$ and J_n has side-length 2^{-n};

(b) $K_{n+1} \subseteq K_n \subseteq [0,2]$ and K_n has length 2^{1-n};

(c) as t describes the interval K_n, the paths γ_n, σ_n each describe a simple arc of ∂J_n from the bottom left corner to the top right, γ_n clockwise, σ_n counter-clockwise;

(d) for $t \notin K_n$ we have $\gamma_{n+1} = \gamma_n = \sigma_n = \sigma_{n+1}$;

(e) for $0 \le t \le 2$ we have

$$|\gamma_n(t) - \gamma_{n+1}(t)| \le 2^{-n}\sqrt{2}, \quad |\gamma_n(t) - \sigma_n(t)| \le 2^{-n}\sqrt{2}, \quad 0 \le t \le 2.$$

(f) continuation of f along γ_n, σ_n does not lead to the same local function element.

To prove Claim 1 we show how to determine $\gamma_{n+1}, \sigma_{n+1}$ consistent with (a) to (f). We quadrisect the square J_n by paths $\mu_k(J_n)$ as described above. Combining these with the restriction of γ_n to $[0,2] \setminus K_n$ gives us five stepwise curves ν_k joining $(0,0)$ to $(1,1)$ and travelling through J_n from the bottom left corner to the top right, such that $\nu_1 = \gamma_n, \nu_5 = \sigma_n$ and $\nu_k(t) = \gamma_n(t) = \sigma_n(t)$ off the interval K_n. Further, $\nu_k = \nu_{k+1}$ off an interval of length 2^{-n}, and ν_k, ν_{k+1} together bound a square of side-length 2^{-n-1}. Choosing k such that continuation of f along ν_k, ν_{k+1} does not lead to the same local function element, we set $\gamma_{n+1} = \nu_k, \sigma_{n+1} = \nu_{k+1}$.

Let x_0 be the unique point of $\mathbb{R}^2$ lying in the intersection of the nested square regions J_n. Then $x_0 \in I^2$ and (abusing notation again slightly) $H(x_0) \in G$. By the construction, the paths γ_n, σ_n both converge uniformly on $[0,2]$ to a path

$$\eta : [0,2] \to I^2, \quad \eta(t) = \gamma_0(t) + \sum_{n=1}^{\infty}(\gamma_n(t) - \gamma_{n-1}(t)).$$

Let t_0 be the unique point of $[0,2]$ lying in the intersection of the K_n. Since $\gamma_n(t_0) \in J_n$ for all n we have $\eta(t_0) = x_0$. Since η joins $(0,0)$ to $(1,1)$ in I^2, we may continue f along η (i.e. analytically continue the function element (f,D) along $H(\eta)$), using function elements $(F_t, U_t), 0 \le t \le 2$.

Let n be large. Then we may partition $[0,2]$ into three intervals $[0,a],[a,b],[b,2]$ with $0 \le a < b \le 2$, such that γ_n, σ_n, η agree on $[0,a]$ and $[b,2]$, and $H(\gamma_n([a,b])), H(\sigma_n([a,b])), H(\eta([a,b]))$ all lie in U_{t_0}. The analytic continuations of (f,D) along $H(\gamma_n) : [0,a] \to G, H(\sigma_n) : [0,a] \to G, H(\eta) : [0,a] \to G$ all lead to the same local function element, which equals F_{t_0} on a neighbourhood of $H(\eta(a))$. We may then use the function element (F_{t_0}, U_{t_0}) to extend these analytic continuations to the interval $[0,b]$, and finally the function elements $(F_t, U_t), t \ge t_0$ extend these continuations to all of $[0,2]$. Thus there are continuations of f along γ_n, σ_n leading to the same local function element near z_1. By Theorem 12.2.4, the same is true of any continuations of f along γ_n, σ_n. This contradicts the way the γ_n, σ_n were chosen, and proves the theorem.

12.3.2 Corollary

Suppose that G is HSC in Theorem 12.3.1. Then f extends to a meromorphic function on G.

12.3.3 Theorem

Let D be a domain in $\mathbb{C}$ and let $w \in \mathbb{C} \setminus D$. Let $\gamma : [a,b] \to D$ be a closed piecewise smooth contour which is homotopic in D to the constant path $\sigma(t) = \gamma(a)$. Then the winding number

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

is 0. In particular, if D is HSC then D is simply connected in terms of winding number.

Proof: We can assume that $[a,b] = [0,1]$ and $w = 0$. Thus $D \subseteq G = \mathbb{C} \setminus \{0\}$ and γ is homotopic to a constant in G. We define a branch $L(z)$ of $\log z$ near $\gamma(0)$ and continue L analytically along γ. Then $L_1 = L_0 = L$ on a neighbourhood of $\gamma(0)$, by Theorem 12.3.1. Since $L'(z) = 1/z$ we have $L_t'(z) = 1/z$ near $\gamma(t)$. For s near t we have $L_s = L_t$ near $\gamma(s)$, and so we have

$$\frac{d}{dt}L_t(\gamma(t)) = \lim_{s\to t}\frac{L_s(\gamma(s)) - L_t(\gamma(t))}{s-t} = \lim_{s\to t}\frac{L_t(\gamma(s)) - L_t(\gamma(t))}{s-t} = \frac{\gamma'(t)}{\gamma(t)}.$$

Thus

$$2\pi i n(\gamma, 0) = L_1(\gamma(1)) - L_0(\gamma(0)) = 0.$$

12.3.4 Cycle reduction

Suppose that Γ is a cycle made up of piecewise smooth contours $\gamma_1, \ldots, \gamma_n$, and that L is a line segment described in one direction as part of γ_j, say from A to B, and in the opposite direction as part of γ_k, with j, k possibly equal. Then we may "cancel" L without changing any integral $\int_\Gamma f(z)dz$, as follows. Assume for simplicity that γ_j, γ_k both start on L. Follow γ_j from the first time it passes through B to the last time it comes to A, and then follow γ_k from the first time it hits A to the last time it comes to B. This gives a closed curve λ for which

$$\int_\lambda f(z)dz = \int_{\gamma_j} f(z)dz + \int_{\gamma_k} f(z)dz$$

for every continuous f.

12.3.5 Lemma

Let A, B be disjoint non-empty compact subsets of $\mathbb{C}^*$, with $A \subseteq \mathbb{C}$. Let $a \in A$. Then there exists a cycle Γ in $\mathbb{C}$ such that $\Gamma \cap (A \cup B) = \emptyset$ and $n(\Gamma, a) = 1$ but $n(\Gamma, z) = 0$ for all $z \in B$.

Here we are using the convention that $n(\gamma, \infty) = 0$ for a cycle γ in $\mathbb{C}$.

Proof. Since A and B are compact and disjoint the distance s from A to B, measured in the chordal metric, is positive. Let r be small and positive, and cover the plane with a grid of closed square regions S_n of side length r, pairwise disjoint except for common sides and vertices, and with a at the centre of one S_n. Since r is small and A is bounded, no S_n can meet both A and B.

Let T_n be the boundary curve of S_n, described once counter-clockwise, and let Γ_0 be the cycle made up of those T_n for which $S_n \cap A \neq \emptyset$. For these T_n it is clear that $T_n \cap B = \emptyset$ and $n(T_n, b) = 0$ for all $b \in B$, so that $n(\Gamma_0, b) = 0$.
Further, we have $n(T_n, a) = 1$ for precisely one n, and so $n(\Gamma_0, a) = 1$.

Apply the cycle reduction process 12.3.4 repeatedly to obtain a cycle Γ made up of edges of Γ_0, in which no S_n-edge is described in both directions, and for which $\int_\Gamma f(z)dz = \int_{\Gamma_0} f(z)dz$ for every continuous f. We need only show that Γ does not meet A. Suppose that $w \in \Gamma \cap A$. If w is a vertex, then w lies on four squares S_n, and there will be four edges in Γ_0, each described in both directions, and the cancellation of these shows that $w \notin \Gamma$. Similarly, if w lies on a square edge, then w lies on two squares and again the edges are cancelled.

12.3.6 Theorem

Let D be a domain in $\mathbb{C}$. The following are equivalent:

(i) D is simply connected in terms of winding number i.e. $n(\gamma, w) = 0$ for every cycle γ in D and every w not in D.

(ii) D is homeomorphic to the disc $D(0, 1)$.

(iii) D is HSC.

(iv) the complement of D in $\mathbb{C}^*$ is connected;

(v) $\partial_\infty D$ is connected.

In particular the winding number condition (i) implies the intuitive condition of “no holes”.

Proof: We show first that (i) implies (ii). Assume first that $D \neq \mathbb{C}$. Then the Riemann mapping theorem tells us that D is homeomorphic to $D(0,1)$. If $D = \mathbb{C}$ then $w = z/(1+|z|)$ is a homeomorphism from $\mathbb{C}$ to $D(0,1)$ (although not analytic).

(ii) implies (iii). This follows from 12.1.9.

(iii) implies (i). This is by Theorem 12.3.3.

Next, (iv) implies (i). To see this, note that $n(\gamma, w)$ is continuous off γ and is 0 for large w. If we set $n(\gamma, \infty) = 0$ the resulting function is continuous on $\mathbb{C}^* \setminus D$, and this set must be unbounded since it is connected and contains ∞. So $n(\gamma, w) \equiv 0$ on $\mathbb{C}^* \setminus D$, by the connectivity again.

(i) implies (iv). Suppose that $H = \mathbb{C}^* \setminus D$ is not connected. Then we may write $H = A \cup B$ in which A, B are disjoint, non-empty, relatively open subsets of H. Thus A and B are relatively closed, and so are compact subsets of $\mathbb{C}^*$. Assuming without loss of generality that $\infty \in B$ it follows from Lemma 12.3.5 that there is a cycle Γ in D with $n(\Gamma, a) \neq 0$ for some $a \in A$.

(v) implies (iv). If $\mathbb{C}^* \setminus D$ is not connected we form disjoint A, B as above and $\partial_\infty D = \partial A \cup \partial B$ is disconnected.

(i) implies (v). We first prove this when D is unbounded. Assume that the compact set $K = \partial_\infty D$ is not connected. Then we may partition K into disjoint non-empty compact subsets A, B of $\mathbb{C}^*$ and, assuming without loss of generality that $A \subseteq \mathbb{C}, \infty \in B$, we can find a cycle Γ not meeting K and such that $n(\Gamma, a) = 1$ for some $a \in A$.

Let the closed piecewise smooth contours which together make up Γ be Γ_j. Since

$$n(\Gamma, a) = \sum_j n(\Gamma_j, a)$$

there must be some j with $n(\Gamma_j, a) = p > 0$. We assert that $\Gamma_j \subseteq D$. Assuming this not the case, we have $\Gamma_j \cap D = \emptyset$, since Γ_j does not meet the boundary of D. Since a is a boundary point of D, we have $n(\Gamma_j, c) = p$ for some $c \in D$ and hence $n(\Gamma_j, z) = p$ for all $z \in D$. But D is unbounded and so there are z in D with $n(\Gamma_j, z) = 0$.

This contradiction proves the result when D is unbounded. If D is a bounded domain with disconnected boundary, we choose distinct a, b in ∂D and put $G = \phi(D), \phi(z) = (z-a)/(z-b)$, noting that ϕ is a homeomorphism of $\mathbb{C}^*$. Thus G fails to satisfy (i), and so is not HSC, and nor is D.

Chapter 13

Riemann surfaces and the uniformization theorem

13.0.1 Definitions

A surface R is a non-empty connected Hausdorff space with a family of mappings ϕ_α such that each ϕ_α maps an open subset U_α of R homeomorphically onto an open subset V_α of $\mathbb{C}$. The U_α cover R.

The ϕ_α are called *charts*, the collection of charts is an *atlas*, and the U_α are *parametric regions*. If $x \in U_\alpha$ then, for sufficiently small t, the open set $\phi_\alpha^{-1}(D(\phi_\alpha(x), t))$ is a *parametric (open) disc* about x. A parametric closed disc is defined in the obvious analogous way i.e. as the pre-image under a chart of a closed disc of positive radius.

Suppose that $U_\alpha \cap U_\beta \neq \emptyset$. Then the function $\phi_\alpha \circ \phi_\beta^{-1}$ maps $\phi_\beta(U_\alpha \cap U_\beta)$ one-one onto $\phi_\alpha(U_\alpha \cap U_\beta)$. This map is called a *transition map*, and maps one open subset of $\mathbb{C}$ one-one onto another.

We say that R is a *Riemann surface* if every such transition map is analytic (on $\phi_\beta(U_\alpha \cap U_\beta)$). In this case the ϕ_α and U_α are said to define a *conformal structure* on R.

Obviously every open connected non-empty subset D of R is also a Riemann surface, with the obvious conformal structure (i.e. take $U_\alpha \cap D$ as parametric regions).

13.0.2 Theorem

Every surface is path-connected.

The proof is the same as the proof that connected open subsets of $\mathbb{C}$ are path-connected.

13.0.3 Analytic functions on Riemann surfaces

Let R, S be Riemann surfaces, and let $f : R \to S$ be a continuous function. Suppose that R has open sets and mappings U_α, ϕ_α, and correspondingly S has sets W_λ and maps ψ_λ.
Let $x \in R$. Then x lies in one of the open sets U_α, and $f(x)$ lies in an open set W_λ. For z in a neighbourhood of x, we have $f(z)$ in W_λ, by continuity. We look at the function $h = \psi_\lambda f \phi_\alpha^{-1}$ (we omit $\circ$ for convenience). This h is defined near $\phi_\alpha(x)$ and maps a neighbourhood of $\phi_\alpha(x)$ into a neighbourhood of $\psi_\lambda(f(x))$, both of these sets contained in $\mathbb{C}$. Note that we have to assume in advance that f is continuous in order to ensure that this composition is defined. We say that f is analytic (in the

Riemann surface sense) if whenever we do this the function h we get is analytic at $\phi_\alpha(x)$ in the usual sense that $h'(z)$ exists on a neighbourhood of $\phi_\alpha(x)$.

Note that for each x, we only need check this for one open set U_α with $x \in U_\alpha$ and one open set W_λ with $f(x) \in W_\lambda$. For, suppose that we also have $x \in U_\beta$ and $f(x) \in W_\mu$. Look at $g = \psi_\mu f \phi_\beta^{-1}$. Near to $\phi_\beta(x)$, we have $g = \psi_\mu f \phi_\beta^{-1} = \psi_\mu(\psi_\lambda^{-1}\psi_\lambda f \phi_\alpha^{-1}\phi_\alpha)\phi_\beta^{-1} = (\psi_\mu\psi_\lambda^{-1})h(\phi_\alpha\phi_\beta^{-1})$.

Now $\phi_\alpha\phi_\beta^{-1}$ is analytic near $\phi_\beta(x)$ (being a transition map). Also $h(\phi_\alpha(x)) = \psi_\lambda(f(x))$ and $\psi_\mu\psi_\lambda^{-1}$ is analytic near this point (a transition map again). So g is analytic if h is.

Note that a constant function from R to S is always analytic. It is routine to check that the composition of analytic functions (in the Riemann surface sense) is analytic.

13.0.4 The identity theorem

Let $f : R \to S$ be an analytic mapping between Riemann surfaces, and let $b \in S$. Let $E = \{w \in R : f(w) = b\}$. If E has a limit point in R then $E = R$.

Proof. Let F be the set of limit points of E in R. Obviously $R \setminus F$ is open. Let $w \in F$, and choose charts ϕ, ψ at w and b respectively. Then $h = \psi f \phi^{-1}$ is analytic near $\phi(w)$, and $\phi(w)$ is a limit point of zeros of $h(z) - b$. Looking at the Taylor series of h near $\phi(w)$ we see that $h(z) \equiv \psi(b)$ on a neighbourhood of $\phi(w)$, and so F is open. The result now follows since R is connected.

13.1 Examples

13.1.1 Plane domains

Any plane domain D can be made into a Riemann surface by taking just one parametric region $U_\alpha = D$, with ϕ_α the identity.

13.1.2 The Riemann sphere

The extended plane $\mathbb{C} \cup \{\infty\} = \mathbb{C}^*$ is made into a Riemann surface as follows. Set $U_1 = \mathbb{C}$ and $U_2 = \mathbb{C}^* \setminus \{0\}$ and $\phi_1(z) = z, \phi_2(z) = 1/z$. Both $\phi_1\phi_2^{-1}$ and $\phi_2\phi_1^{-1}$ are defined on $\mathbb{C} \setminus \{0\} = U_1 \cap U_2$ and both are just $z \to 1/z$, which is analytic there.

It follows easily that if D is a plane domain and $f : D \to \mathbb{C}^*$ is meromorphic (i.e. analytic apart from isolated poles) then f is an analytic function from D into the Riemann surface $\mathbb{C}^*$ (with the standard conformal structure above). The converse is also true, except that the function which is identically ∞ is not normally regarded as meromorphic.

Since $R(1/z)$ is a rational function when $R(z)$ is, it's also easy to see that rational functions are analytic functions from $\mathbb{C}^*$ into itself.

13.1.3 Theorem

Let $f : \mathbb{C}^* \to \mathbb{C}^*$ be analytic and non-constant. Then f is a rational function.

Proof. We can assume $f(\infty) \neq \infty$ (because $z \to 1/z$ is an analytic function on $\mathbb{C}^*$). So there is some $R > 0$ such that $f(z) \neq \infty$ for $|z| > R$. The set of z in $\mathbb{C}$ with $f(z) = \infty$ has no limit point w in $\mathbb{C}$ and so $f^{-1}(\{\infty\})$ is finite, since $\{z \in \mathbb{C} : |z| \leq R\}$ is a compact set.

If $f(a) = \infty$ with a finite, then near a we use Laurent's theorem to write $f(z) = (z-a)^{-n}H(z) = S_a(z) + H_1(z)$, where H and H_1 are analytic at a, and S_a is a polynomial in $1/(z-a)$. Note that $S_a(z) \to 0$ as $z \to \infty$.

Now we just set $S(z) = \sum S_a$, in which the sum is over the finitely many a for which $f(a) = \infty$. Then $f(z) - S(z)$ stays bounded as z approaches each such a, and so $f(z) - S(z)$ is an entire function. Since $S(\infty) = 0$ and $f(\infty)$ is finite, $f - S$ is a bounded entire function and so constant.

13.1.4 Lifting a conformal structure to a covering space

Let R be a Riemann surface and let X be a path-connected Hausdorff topological space with a mapping $\psi : X \to R$ which is continuous and locally one-one, and maps open sets to open sets. Then X is called a covering space of $\psi(X)$, and X inherits a conformal structure from R as follows.

Let the open sets and mappings of R be U_α, ϕ_α. Then we know that $\phi_\alpha \phi_\beta^{-1}$ is analytic on $\phi_\beta(U_\alpha \cap U_\beta)$ whenever $U_\alpha \cap U_\beta \neq \emptyset$.

Let x be in X. Then $\psi(x)$ lies in some U_α. Take a neighbourhood V_α of x on which ψ is one-one and such that $\psi(V_\alpha) = U_\alpha^*$ is contained in U_α. The open sets for X are just these V_α, and these cover X. The mappings are just the compositions $\psi_\alpha = \phi_\alpha \psi$. If $V_\alpha \cap V_\beta$ is non-empty, then $\psi_\beta(V_\alpha \cap V_\beta) = \phi_\beta(U_\alpha^* \cap U_\beta^*) \subseteq \phi_\beta(U_\alpha \cap U_\beta)$. On $\psi_\beta(V_\alpha \cap V_\beta)$ we have $\psi_\alpha \psi_\beta^{-1} = \phi_\alpha \phi_\beta^{-1}$, which is analytic.

In particular any space homeomorphic to a Riemann surface inherits a conformal structure. Thus a sphere can be made into a Riemann surface, and so can the exterior of a cube.

13.1.5 Multiply-valued functions

The best known application of Riemann surfaces is as the "natural" domain of definition of certain multiply-valued functions.

The logarithm

The complex logarithm $\log z = \ln|z| + i \arg z$, with any choice of the argument, is analytic on the cut plane $D_0 = \{z = re^{i\theta}, r > 0, -\pi < \theta < \pi\}$, but is discontinuous as z approaches the negative real axis.

To get around this difficulty, take countably many copies G_n of D_0, and glue them together along the interval $(-\infty, 0)$ so that as we leave G_n travelling counter-clockwise, we move up to G_{n+1}. On G_n, define $\log z = \ln|z| + i \arg z$ with the argument chosen to lie in $((2n-1)\pi, (2n+1)\pi)$. On the resulting "spiral" surface R, the function $\log z$ so assembled is continuous, and maps R onto $\mathbb{C}$. The charts on R are just projection onto $\mathbb{C}$ in the obvious way.

A slightly more formal approach is to take the surface $S = \{(r\cos t, r\sin t, t) : r > 0, t \in \mathbb{R}\}$ with local charts $(r\cos t, r\sin t, t) \to (r\cos t, r\sin t)$, and $\ln r + it$ maps S onto $\mathbb{C}$.

13.1.6 The square root

The square root $f_0(z) = z^{\frac{1}{2}} = \sqrt{r}e^{i\theta/2}$, with $z = re^{i\theta}, r > 0, -\pi < \theta < \pi$, is again analytic on D_0 but discontinuous at the negative real axis. Set $f_1(z) = -f_0(z)$. For w on $(-\infty, 0)$ we have

$$\lim_{z \uparrow w} f_j(z) = \lim_{z \downarrow w} f_{1-j}(z).$$

Take two copies G_0, G_1 of the Riemann sphere, slit along the open interval $(-\infty, 0)$, and glue them together so that as we leave G_j across $(-\infty, 0)$ we enter G_{1-j}. Define f to be f_j on G_j, and extend it continuously to the edges. The resulting surface is homeomorphic under f to $\mathbb{C}^*$.

13.1.7 Example

The solutions of algebraic equations may be defined as single valued functions on suitable Riemann surfaces: see Ahlfors' *Complex Analysis* [2] for details. We describe here just one example, which leads to a surface which is not simply connected. Define w by

$$w^2 = z(z-1)(z-2).$$

A basic fact from complex analysis states that if F is analytic and zero-free on a simply connected domain G, then F has an analytic square root on G. Hence we may form analytic solutions $w = f_0(z), w = f_1(z) = -f_0(z)$ on the domain formed by cutting the plane along the interval $[0, \infty)$. Let

$$A = (0, 1), \quad B = (1, 2), \quad C = (2, \infty).$$

Then we have

$$\lim_{z \uparrow w} f_j(z) = \lim_{z \downarrow w} f_{1-j}(z), \quad w \in A.$$

On the other hand, writing

$$w = z(1 - 1/z)^{\frac{1}{2}}(z-2)^{\frac{1}{2}}$$

we see that the f_j extend analytically to $1 < |z| < 2$.

Take two copies G_0, G_1 of the Riemann sphere, each slit along the open intervals A, C, and again join them across the cuts. Let $f(z)$ be $f_j(z)$ on the interior of G_j, and extend f continuously to the resulting surface (which is topologically a torus). Apart from at $0, 1, 2, \infty$, the local charts are just projection onto $\mathbb{C}$, but at the four "branch points" we need to be more careful, since projection is not locally one-one there. However, on a neighbourhood of 0 we may use $z^{\frac{1}{2}}$, as in the previous example, and we do the same at $1, 2, \infty$.

13.1.8 The uniformization theorem

The majority of this chapter is devoted to presenting a proof of this important result, which states that if R is a simply connected Riemann surface then R is conformally equivalent to precisely one of the following (in each case with the standard conformal structure): the open plane $\mathbb{C}$; the extended plane $\mathbb{C}^*$; the unit disc $D(0, 1)$.

It follows from Liouville's theorem and compactness that no two of $\mathbb{C}, \mathbb{C}^*, D(0, 1)$ can be conformally equivalent. The fact that R is conformally equivalent to one of these requires the theory of subharmonic functions on Riemann surfaces. The proof presented here is modified from one given by W. Abikoff [1] in the American Math. Monthly, October 1981.

13.2 Subharmonic functions and Perron families

13.2.1 Lemma

Let u be subharmonic on a plane domain D, and let $f : D \to \mathbb{C}$ be conformal. Then $u \circ f^{-1}$ is subharmonic on $f(D)$.

This follows immediately from Theorem 9.2.4. With this result we can define subharmonic functions on Riemann surfaces as follows.

13.2.2 Definition

Let R be a Riemann surface, and let $u : R \to [-\infty, \infty)$ be continuous. We say (initially) that u is subharmonic on R if $u \circ \phi_\alpha^{-1}$ is subharmonic on $\phi_\alpha(U_\alpha)$ for every chart ϕ_α.

13.2.3 Lemma

Let u be subharmonic on the Riemann surface R, let N be a closed parametric disc (with local chart ϕ), and suppose that $u \equiv -\infty$ on a subset E of ∂N such that $\phi(E)$ has positive angular measure. Then $u \equiv -\infty$ on R.

Proof. By the theory of subharmonic functions in the plane (Poisson's formula) we get $u \equiv -\infty$ on N. Let F be the (obviously open) subset of R defined by the property that $w \in F$ iff $u \equiv -\infty$ on a neighbourhood of w. We claim that F is closed, and this holds since if $w_n \in F$ tend to $w \in R$ then $u \equiv -\infty$ on a parametric disc centred at w. So $F = R$ by connectedness.

Henceforth we consider *only* subharmonic functions which are not $\equiv -\infty$. We recall Harnack's theorem from 8.3.10.

13.2.4 Harnack's theorem

Let D be a domain in $\mathbb{C}$ and let u_n be functions harmonic on D such that $u_1 \leq u_2 \leq \ldots$. Let $v(z) = \lim_{n\to\infty} u_n(z)$ for each $z \in D$. Then either $v \equiv \infty$ on D, or v is harmonic on D.

13.2.5 Definitions

Let u be subharmonic on R and let D be a closed parametric disc in R. Then we can form a subharmonic function u_D which satisfies $u \leq u_D$ on R, is harmonic on the interior of D (using Lemma 13.2.3), and equals u off the interior of D. We call u_D the Poisson modification of u.

By a Perron family P we mean a non-empty collection of functions u subharmonic on R, such that:

(i) if $u \in P$ then $u_D \in P$ for every closed parametric disc D in R;

(ii) if $u, v \in P$ then $\max\{u, v\}$ is in P.

13.2.6 Theorem

Let P be a Perron family, and for each $p \in R$ define $g(p) = \sup\{u(p) : u \in P\}$. Then either $g \equiv +\infty$ on R, or g is harmonic on R.

Proof. Take $p_0 \in R$, and a closed parametric disc D_0 centred at p_0.

Suppose first that $g(p_2) = \infty$ for some p_2 in the interior U_0 of D_0. Then there exist $v_n \in P$ such that $v_n(p_2) \to \infty$, and we may assume that $v_n \leq v_{n+1}$ and each v_n is harmonic on U_0 (if not, first take maximums so that $v_n \leq v_{n+1}$ and then replace each v_n by its Poisson modification). Then Harnack's theorem tells us that $v_n \to \infty$ on U_0 and so $g \equiv +\infty$ on U_0.

Suppose now that $g(p) < \infty$ for every $p \in U_0$. Choose $p_1 \in U_0, p_1 \neq p_0$. We can take $u_n \in P, v_n \in P$ such that $u_n(p_0) \to g(p_0), v_n(p_1) \to g(p_1)$, and we may assume that u_n and v_n satisfy $u_n \leq u_{n+1}, v_n \leq v_{n+1}$ and are harmonic on U_0. We may also take $w_n \in P$, harmonic on U_0, such that $w_n \geq \max\{u_n, v_n\}$ on R.

Harnack's theorem gives us $u_n \to h, w_n \to k$, with h, k harmonic on U_0. Since $g(p_0) \geq w_n(p_0) \geq u_n(p_0) \to g(p_0)$, we get $h(p_0) = k(p_0) = g(p_0)$. On the other hand, since $u_n \leq w_n$ we get $h \leq k$ on U_0. The maximum principle now tells us that $h = k$ on U_0.

It follows in particular that $h(p_1) = k(p_1)$. Since $g(p_1) \geq w_n(p_1) \geq v_n(p_1) \to g(p_1)$ we get $g(p_1) = k(p_1) = h(p_1)$. Hence $g = h$ on U_0, and g is harmonic on U_0.

The result now follows by connectedness.

13.3 Green's function

13.3.1 Definition

Let R be a Riemann surface, and let $p_0 \in R$. Let ϕ be a chart near p_0. Consider the Perron family V_{p_0} of all functions v which are subharmonic on $R \setminus \{p_0\}$, and with the following properties:

(i) there exists a compact K_v such that $v = 0$ off K_v;

(ii) $\limsup_{p \to p_0} v(p) + \log|\phi(p) - \phi(p_0)| < \infty$.

Condition (ii) is independent of the particular chart ϕ, because if ψ is another chart then $h = \psi \circ \phi^{-1}$ is locally conformal and

$$|\psi(p) - \psi(p_0)| = |h(\phi(p)) - h(\phi(p_0))| \leq c|\phi(p) - \phi(p_0)|$$

as $p \to p_0$, for some constant c. Obviously $0 \in V_{p_0}$. Set

$$g(p, p_0) = \sup\{v(p) : v \in V_{p_0}\} \geq 0.$$

Then either $g \equiv \infty$, or g is harmonic on $R \setminus \{p_0\}$. In the second case, we call g the Green's function for R, p_0.

13.3.2 Lemma

Let ϕ be a chart at p_0, mapping p_0 to z_0 and a neighbourhood of p_0 onto $D(z_0, r)$. Let $0 < r_1 < r_2$, and define closed parametric discs K_1, K_2 by

$$K_j = \phi^{-1}(\{z : |z| \leq r_j\}). \tag{13.1}$$

Let $v \geq 0, v \in V_{p_0}$, and let

$$m_j = \max\{v(p) : p \in \partial K_j\}. \tag{13.2}$$

Then

$$m_1 + \log r_1 \leq m_2 + \log r_2. \tag{13.3}$$

If $g(p, p_0)$ exists then we also have

$$\max\{g(p, p_0) : p \in \partial K_1\} + \log r_1 \leq \max\{g(p, p_0) : p \in \partial K_2\} + \log r_2. \tag{13.4}$$

Proof. We may assume that $z_0 = 0$. Let $\varepsilon > 0$. The function

$$h(p) = v(p) + (1 + \varepsilon) \log |\phi(p)|$$

tends to $-\infty$ as $z \to p_0$, and is subharmonic on a neighbourhood of K_2. Hence

$$h(p) \leq \max\{h(q) : q \in \partial K_2\}, \quad p \in K_2.$$

In particular

$$m_1 + (1 + \varepsilon) \log r_1 \leq m_2 + (1 + \varepsilon) \log r_2$$

and (13.3) follows on letting $\varepsilon \to 0$. We now get, for $w \in V_{p_0}, w \geq 0$,

$$\max\{w(p) : p \in \partial K_1\} + \log r_1 \leq \max\{g(p, p_0) : p \in \partial K_2\} + \log r_2$$

and the last assertion of the lemma follows.

13.3.3 Lemma

If $g(p, p_0)$ exists then $g(p, p_0) + \log |\phi(p) - \phi(p_0)|$ has a harmonic extension to a neighbourhood of p_0.

Proof. Assume that $\phi(p_0) = 0$. By (13.4), $g(p, p_0) + \log |\phi(p)|$ is bounded above as $p \to p_0$. Now take a small positive r_0 and define v by

$$v(p) = \log r_0 - \log |\phi(p)|, \quad |\phi(p)| < r_0,$$

with $v(p) = 0$ otherwise. Then v is subharmonic on $R \setminus p_0$ and is in V_{p_0}, and so

$$g(p, p_0) \geq v(p) \geq -\log |\phi(p)| - O(1), \quad p \to p_0.$$

13.3.4 Lemma

If $g(p, p_0)$ exists then g is non-constant and positive and $c = \inf\{g(p, p_0) : p \in R\}$ satisfies $c = 0$.

Proof. Lemma 13.3.3 shows that g is non-constant, and thus g is positive, by the maximum principle on $R \setminus \{p_0\}$. Clearly $c \geq 0$. Take $\varepsilon > 0$ and $v \in V_{p_0}$, and set

$$k(p) = (1 - \varepsilon)v(p) - g(p, p_0) + c.$$

Outside a compact set we have $k(p) \leq 0$. Also, as $p \to p_0$ we have

$$v(p) \leq g(p, p_0), \quad g(p, p_0) \to \infty \quad , \quad k(p) \to -\infty.$$

Hence $k(p) \leq 0$ on $R \setminus \{p_0\}$. Letting $\varepsilon \to 0$ we get

$$v(p) \leq g(p, p_0) - c$$

and taking the supremum over v gives $c = 0$.

13.3.5 Existence of the Green's function

If R is compact, then Green's function cannot exist, because g would have a minimum on R, contradicting the fact that g would be harmonic and non-constant on $R \setminus \{p_0\}$.

We consider next necessary conditions, and sufficient conditions, for Green's function to exist on a non-compact surface.

13.3.6 Definition

Let K be a compact subset of R. We say that the maximum principle fails for K if there exists a function h, subharmonic and bounded above on $R \setminus K$, with $\limsup_{p \to K} h(p) \le 0$ and $h(p) > 0$ for some $p \in R \setminus K$.

Here $\limsup_{p \to K}$ means $\limsup_{p \to \partial K, p \in R \setminus K}$.

13.3.7 Examples

The function $\log |z|$ shows that the maximum principle fails for $\{z : |z| \le 1\}$, with respect to the surface $D(0, 2)$.

On the other hand, the maximum principle holds for $\{z : |z| \le 1\}$, with respect to the surfaces $\mathbb{C}^*$ and $\mathbb{C}$, in the latter case because the singularity at ∞ is removable for subharmonic functions which are bounded above.

Note that Green's function $g(p, 0)$ does not exist for $\mathbb{C}$, since $\log^+ R/|z|$ is in V_0 for every $R > 0$.

13.3.8 Lemma

Suppose that $g(p, p_0)$ exists, and that K is a compact subset of R, properly containing $\{p_0\}$. Then the maximum principle fails for K.

Proof. The function $h(p) = -g(p, p_0) \le 0$ has a maximum m on K, and this is taken at some $p_1 \in K, p_1 \neq p_0$. If we had $h(p) \le m$ on $R \setminus K$ then h would have a maximum on $R \setminus \{p_0\}$, which violates the ordinary maximum principle.

In the converse direction, we have:

13.3.9 Theorem

Let K be a compact subset of R, and let p_0 be an interior point of K. Suppose that the maximum principle fails for K. Then $g(p, p_0)$ exists.

Proof. Choose a neighbourhood of p_0, contained in K, mapped onto $D(z_0, r)$ by a chart ϕ. We may assume that $z_0 = 0$. Let $0 < r_1 < r_2 < r$. Let K_j be as in (13.1). Let V be the family of subharmonic functions $v : R \setminus K_1 \to [0, 1]$, such that if $\varepsilon > 0$ then $v(w) < \varepsilon$ for all w outside some compact $L_{v,\varepsilon}$ (thus $v(w) \to 0$ as w "tends to infinity' in $R \setminus K_1$). Set $u(p) = \sup\{v(p) : v \in V\}$. Then u is harmonic on $R \setminus K_1$.

We claim that $u(p) < 1$ for all $p \in R \setminus K_1$. We know that there exists a function h, subharmonic and bounded above by 1 on $R \setminus K$, with $\limsup_{p \to K} h(p) \leq 0$ and $h(p_1) > 0$ for some $p_1 \in R \setminus K$. Let $v \in V$. Then $v + h$ is subharmonic on $R \setminus K$, and

$$\limsup_{p \to K}(v(p) + h(p)) \leq 1, \quad \limsup_{p \to \infty}(v(p) + h(p)) \leq 1.$$

The ordinary maximum principle now gives $v(p_1) \leq 1 - h(p_1) < 1$, which gives $u(p_1) < 1$. Since we now have $u \not\equiv 1$, applying the maximum principle proves the claim.

Now take any $v \in V_{p_0}$. We may assume that $v \geq 0$, since otherwise we replace v by $\max\{v, 0\}$. Define m_j by (13.2). Since v vanishes off a compact set we have $v(p) \leq m_1$ on $R \setminus K_1$. It follows that $v/m_1 \in V$. Hence we get

$$v(p) \leq m_1 u(p), \quad p \in R \setminus K_1.$$

In particular we have

$$m_2 \leq m_1 M_2, \quad M_2 = \max\{u(p) : p \in \partial K_2\} < 1.$$

Combining this with (13.3) leads to

$$m_1 \leq m_2 + \log r_2/r_1 \leq m_1 M_2 + \log r_2/r_1$$

and so

$$m_1 = \max\{v(p) : p \in \partial K_1\} \leq (1 - M_2)^{-1} \log r_2/r_1.$$

Since v is an arbitrary non-negative element of V_{p_0}, while u is fixed, we deduce that $g(p, p_0)$ is finite for $p \in \partial K_1$, and so the Green's function exists.

13.3.10 Corollary

If $g(p, p_0)$ exists for some p_0 in R then $g(p, p_1)$ exists for every p_1 in R.

To see this, take a compact set, the interior of which contains p_0, p_1.

A non-compact Riemann surface R is called *hyperbolic* if Green's function exists, and *parabolic* otherwise.

13.3.11 Lemma

Let R be a non-compact Riemann surface such that there exists a function u non-constant, subharmonic and bounded above on R. Then R is hyperbolic.

The converse is also true, because we can take $-g(p, p_0)$.

Proof. Take $p_0 \in R$. Since u is non-constant, there exists $p_1 \in R$ with $u(p_1) > u(p_0)$. Take a compact subset K of R, with p_0 an interior point of K, and with $p_1 \notin K$, and such that $u(p) < u(p_1)$ on K. Thus the maximum principle fails for K.

13.3.12 Lemma

Suppose that S is a domain on R, and that the boundary of S, relative to R, contains a path γ joining distinct points of R. Then S has a Green's function.

Proof. Take a parametric disc U, the closure of which is contained in R, such that U contains a subpath σ of γ joining distinct points of U. Solving the Dirichlet problem on the image of U in $\mathbb{C}$ we get a function u harmonic on $U \setminus \sigma$, with

$$\lim_{p \to \sigma} u(p) = 1, \quad \lim_{p \to \partial U} u(p) = 0.$$

Extend u to be 0 on $S \setminus U$. Then u is non-constant and subharmonic, but bounded above, on S.

13.4 The uniformization theorem: the hyperbolic case

13.4.1 Theorem

Let R be a hyperbolic Riemann surface (and so open). Then R is conformally equivalent to $D(0,1)$ (with the standard conformal structure).

Proof. Take some p_0 on R and form Green's function $g(p, p_0)$. Let ϕ be a chart near p_0, without loss of generality mapping p_0 to 0. Now $g(p, p_0) + \log|\phi(p)|$ has a harmonic extension $-u(p)$ to a neighbourhood of p_0, on which we define a harmonic conjugate $v(p)$ of u, and f by

$$f(p) = \phi(p) \exp(u(p) + iv(p)), \quad \log|f| = u + \log|\phi| = -g.$$

Since $-g$ has a harmonic conjugate in a neighbourhood of each point of $R \setminus \{p_0\}$, we may analytically continue f subject to $\log|f| = -g$ througout R, and by the monodromy theorem this defines an analytic function $f : R \to D(0,1)$.

It suffices to show that f is univalent, because $f(R)$ will then be a simply connected subdomain of $D(0,1)$ and so conformally equivalent to $D(0,1)$. Let $p_1 \in R \setminus \{p_0\}$, let $a = f(p_1)$, and let

$$T(w) = \frac{w - a}{1 - \bar{a}w},$$

so that $T(a) = 0$. Note that $T(0) = -a$.

Let $\varepsilon > 0$ and let $v_1 \in V_{p_1}$. Let

$$h(p) = v_1(p) + (1 + \varepsilon) \log|T(f(p))|.$$

Outside a compact set we have $v_1 = 0$ and so $h < 0$. Next, let ψ be a chart at p_1, without loss of generality mapping p_1 to 0. Then $Tf\psi^{-1}$ is analytic at 0 and as $p \to p_1$ we have

$$\log|T(f(p))| \leq \log|\phi(p)| + O(1)$$

and hence $h(p) \to -\infty$. Thus h is subharmonic and negative throughout R. Letting $\varepsilon \to 0$ we get

$$v_1(p) \leq -\log|T(f(p))|, \quad g(p, p_1) \leq -\log|T(f(p))|. \tag{13.5}$$

But

$$|T(f(p_0))| = |T(0)| = |a| = |f(p_1)|$$

and so we get

$$g(p_0, p_1) \leq -\log|f(p_1)| = g(p_1, p_0).$$

By symmetry, we get

$$g(p_1, p_0) = g(p_0, p_1),$$

and so

$$g(p_0, p_1) = -\log|f(p_1)| = -\log|T(f(p_0))|.$$

Since $g(p, p_1) + \log|T(f(p))|$ is subharmonic and, by (13.5), non-positive on $R \setminus \{p_1\}$, we deduce that

$$g(p, p_1) = -\log|T(f(p))|, \quad p \in R,$$

in which both sides are infinite at p_1.

Now suppose that $f(p_2) = f(p_1)$. Then $T(f(p_2)) = 0$, and so $g(p_2, p_1) = \infty$. But this gives $p_2 = p_1$, and f is univalent as required.

13.5 The non-hyperbolic case

13.5.1 Divergent curves

Let R be a non-compact Riemann surface. A divergent curve on R is a simple path $\gamma : [0, \infty) \to R$ such that

$$\lim_{t\to\infty} \gamma(t) = \infty,$$

by which we mean that if K is a compact subset of R then there exists $t_0 \geq 0$ such that $\gamma(t) \notin K$ for all $t \geq t_0$.

Since deleting a point from R gives a set which is still connected, γ cannot pass through every point of R.

Obviously if R is compact and $p_0 \in R$ then $R^0 = R \setminus \{p_0\}$ is not compact (take open sets $U_n = R \setminus K_n$, in which K_n are compact neighbourhoods of p_0 decreasing to $\{p_0\}$), and R^0 has a divergent curve (since R is Hausdorff any compact $K \subseteq R^0$ fails to meet some open neighbourhood of p_0).

13.5.2 Lemma

Let γ be a divergent curve on the simply connected non-compact Riemann surface R. For $t \geq 0$ set $R_t = R \setminus \gamma([t, \infty))$. Then R_t is an open set. Further, if U is an open parametric disc centred at $\gamma(s)$ and s_1, s_2 are sufficiently close to s then there is a homeomorphism f of R_{s_2} onto R_{s_1} which is the identity outside U.

Proof. Assume without loss of generality that $s_1 \leq s_2$. Let $s_3 = \inf\{t \geq s : \gamma(t) \notin U\}$. Then $V = U \setminus \gamma([s_1, s_3))$ is a domain (by the Jordan curve theorem, using the fact that a disc is homeomorphic to $\mathbb{C}$) and is simply connected (via the homeomorphism to $\mathbb{C}$ and a winding number argument). The Riemann mapping theorem gives a homeomorphism g of V onto $W = D(0, 1) \setminus [\frac{1}{2}, 1)$, and we then take a homeomorphism of W which is the identity outside $D(0, \frac{3}{4})$.

13.5.3 Lemma

R_t is a simply connected domain, and a hyperbolic Riemann surface.

Proof. We prove first that R_t is connected. To see this, assume without loss of generality that $t = 0$. Let $a, b \in R_0$ and let $s = \inf A$, where A is the set of $t > 0$ such that there exists a path from a to b in R_t. Then $s < \infty$ since R is connected.

Take an open parametric disc U centred at $\gamma(s)$, and not containing a, b. Take s_1, s_2 close to s, with $s_1 \leq s < s_2$ and $s_2 \in A$. By assumption, there exists a path ρ from a to b in R_{s_2}. Let f be as in Lemma 13.5.2. Then $f(a) = a, f(b) = b$, and $f(\rho)$ is a path from a to b in R_{s_1}. It follows that $s \in A$ and $s = 0$.

We now show that R_t is simply connected. Assume again that $t = 0$, and let Γ be a closed curve in R_0. Since R is simply connected, Γ is homotopic in R to a constant curve, via a homotopy function F. Let $s = \inf B$, with B the set of $t > 0$ such that Γ is homotopic to a constant in R_t. Then $s < \infty$, because for large s we can take F.

Take an open parametric disc U, centred at $\gamma(s)$, such that U does not meet Γ. Take $s_1 \leq s < s_2 < \infty$, with $|s_j - s|$ small, such that $s_2 \in B$. Thus Γ is homotopic to a constant in R_{s_2}, via a homotopy function F_2. Take f as in Lemma 13.5.2 again, so that $f(\Gamma(t)) = \Gamma(t)$. It follows that $f \circ F_2$ is a homotopy in R_{s_1}, deforming Γ to a constant path.

Since the boundary of R_t in R contains a simple path, each R_t is hyperbolic.

13.5.4 Theorem

Let R be a simply connected Riemann surface, with a divergent curve, and having no Green's function. Then R is conformally equivalent to $\mathbb{C}$.

Note that $\mathbb{C}$ is homeomorphic to $D(0,1)$, and thereby inherits a conformal structure with a Green's function. So the hypothesis that R has no Green's function is not redundant.

Proof. Fix $p_0 \in R_0$, and a chart ϕ at p_0. We may assume that $\phi(p_0) = 0$. Each R_n is hyperbolic, and so there is a conformal map $G_n : R_n \to D(0,1)$ with the standard conformal structure on $D(0,1)$, and with $G_n(p_0) = 0$. Let $g_n = G_n \circ \phi^{-1}$, and let $g_n'(0) = 1/c_n$. Let $F_n = c_n G_n$. Then F_n maps R_n conformally onto $B_n = D(0, |c_n|)$, and $f_n = F_n \circ \phi^{-1}$ has $f_n'(0) = 1$.

For $n \geq m$, we have $R_m \subseteq R_n$, and so $F_n \circ F_m^{-1}$ maps B_m conformally into B_n, with 0 mapped to 0. Since $F_n \circ F_m^{-1} = f_n \circ f_m^{-1}$ near 0, we see that the derivative of $F_n \circ F_m^{-1}$ at 0 is 1.

Now the family of functions f analytic and univalent on a fixed disc $D(0,r)$, with the normalization $f(0) = f'(0) - 1 = 0$, is a normal family, by Koebe's distortion theorem. The limit function of any convergent sequence in this family is analytic and, by Hurwitz' theorem, univalent.

We apply the diagonalization process. Take a subsequence F_{1n} of F_n such that as $n \to \infty$ the sequence $F_{1n} \circ F_1^{-1}$ converges LU on B_1 to a function H_1 analytic and univalent there. Take a subsequence F_{2n} of F_{1n} such that, as $n \to \infty$, the sequence $F_{2n} \circ F_2^{-1}$ converges LU on B_2, to H_2. We repeat this. Note that $H_k \circ F_k$ is a conformal map of R_k into $\mathbb{C}$.

Now let $P_n = F_{nn}$. For each k, this sequence P_n is eventually a subsequence of F_{kn} and so $P_n \circ F_k^{-1}$ converges LU on B_k to H_k.

Let $k \leq m$. Then $R_k \subseteq R_m$ and we have, on R_k,

$$H_k \circ F_k = \lim(P_n \circ F_k^{-1}) \circ F_k = \lim(P_n \circ F_m^{-1} \circ F_m \circ F_k^{-1}) \circ F_k = \lim(P_n \circ F_m^{-1}) \circ F_m = H_m \circ F_m.$$

Since the union of the R_k is R, this defines a conformal map H of R onto a simply connected domain D in $\mathbb{C}$. If $D \neq \mathbb{C}$ then there exists a conformal map ψ of D onto $D(0,1)$, so that $\psi \circ H$ maps p_0 to 0.

But then the function $\log|\psi \circ H|$ is subharmonic, non-constant and bounded on R, contradicting the assumption that R is not hyperbolic.

To handle the remaining cases of the uniformization theorem, we need the following lemma, the proof of which we postpone till the next section.

13.5.5 Lemma

Let R be a simply connected Riemann surface and let $p_0 \in R$. Let $R^0 = R\setminus\{p_0\}$. If R has no divergent curves, then R^0 is simply connected.

In particular R^0 is simply connected if R is simply connected and compact.

13.5.6 Theorem

Let R be a simply connected Riemann surface with no divergent curves. Then R is conformally equivalent to $\mathbb{C}^*$ (and so compact).

Proof. Take $p_0 \in R$ and form the punctured surface R^0, which by Lemma 13.5.5 is simply connected. Obviously R^0 has a divergent curve.

If R^0 is parabolic then R^0 is conformally equivalent to $\mathbb{C}$, via some conformal map f, and a simple argument shows that $f(p) \to \infty$ as $p \to p_0$. Thus R is conformally equivalent to $\mathbb{C}^*$.

It remains only to show that R^0 cannot be hyperbolic. Assuming that R^0 has a Green's function, we obtain a conformal mapping f of R^0 onto $D(0,1)$. The singularity at p_0 is removable, and the maximum principle gives $f(p_0) \in D(0,1)$ and so $f(p_0) = f(p_1)$ for some $p_1 \in R^0$. But then, by the open mapping theorem, all values near $f(p_1)$ are taken by f near p_0 and near p_1, contradicting the univalence of f on R^0.

13.6 The case of no divergent curves

In this section we prove Lemma 13.5.5. We puncture R to form $R^0 = R\setminus\{p_0\}$. Let $N = N_0$ be a closed parametric disc centred at p_0, and let $R^1 = R \setminus N_0$. Since an annulus is homeomorphic to a punctured disc, R^1 is homeomorphic to R^0, and it will therefore suffice to prove that R^1 is simply connected. Fix $p_1 \in R^1$, and form the Green's function $g(p, p_1)$. This exists, by Lemma 13.3.12.

13.6.1 Lemma

We have $g(p, p_1) \to 0$ as $p \to \partial N$ from R^1.

Proof. Take a closed parametric disc N_1 centred at p_0, such that N_0 lies in the interior of N_1. Solve the Dirichlet problem with boundary values $g(p)$ on ∂N_1 and 0 on ∂N_0. Note that to do this we only need the Dirichlet problem for a plane annulus. Let the resulting function be h. If u is a function in the Perron family defining g, then $\limsup_{p\to\partial N_j}(u(p) - h(p)) \leq 0$. Thus $u \leq h$ on $N_1 \setminus N_0$ by the maximum principle, giving $g \leq h$ there. Since $h(p) \to 0$ as $p \to \partial N_0$ we get $\limsup_{p\to\partial N_0} g(p) \leq 0$. Since $g \geq 0$, the result follows.

13.6.2 Lemma

Let $\infty > T > 0$ and let Y be a component of the set $\{p \in R^1 : g(p) > T\}$. Then g is not bounded above on Y and there exists a path $\gamma : [0, \infty) \to Y$ such that $g(\gamma(t)) \to \infty$ as $t \to \infty$.

Proof. Suppose first that g is bounded above on Y. Obviously $g(p) = T$ on ∂Y, which does not meet N. Define $v(p) = g(p)$ for $p \in Y$, with $v(p) = T$ for $p \in R \setminus Y$. Then v is subharmonic and bounded above on R, contradicting Lemma 13.3.11 and the assumption that R is not hyperbolic.

The construction of the path is now standard. Fix $y_0 \in Y$. Let $Y_0 = Y$ and, assuming that y_n, Y_n have been defined, let Y_{n+1} be a component of the set $\{p : g(p) > T + n + 1\}$ such that $Y_{n+1} \subseteq Y_n$, and choose $y_{n+1} \in Y_{n+1}$. Join y_n to y_{n+1} by a path in Y_n.

Since R has no divergent curves, $\gamma(t)$ must tend to p_1 and we get immediately:

13.6.3 Lemma

Let $\infty > T \geq 0$. Then the set $\{p \in R^1 : g(p) > T\}$ has a unique component.

13.6.4 Lemma

Let $\infty > T > 0$, and suppose that $p_3 \in R^1$ with $g(p_3) \leq T$. Then p_3 can be joined to ∂N by a path $\sigma(t), 0 \leq t \leq 1$, such that $g(\sigma(t)) < T$ for $0 < t < 1$.

Proof. Assume that $g(p_3) < T$ (if not, first join p_3 to p_4 with $g(p_4) < T$). Let Y be that component of the set $\{p : g(p) < T\}$ which contains p_3. We assert that ∂Y (the boundary of Y with respect to R) meets ∂N, from which the existence of the required path is immediate. Suppose that ∂Y does not meet ∂N. Then $\partial Y \subseteq R^1$ and $g(p) = T$ on ∂Y. Let $v(p) = -g(p)$ for $p \in Y$, with $v(p) = -T$ for $p \in R \setminus Y$. Then v is subharmonic, non-constant and bounded above on R, contradicting Lemma 13.3.11.

13.6.5 Critical points

If ϕ is a local parameter near $p_2 \in R^1, p_2 \neq p_1$, then p_2 is a critical point of g if $G = g \circ \phi^{-1}$ has a critical point (i.e. $G_x = G_y = 0$) at $\phi(p_2)$. This property is independent of the choice of ϕ, by the chain rule. The critical points of g are isolated, and if f is analytic on a neighbourhood U of p_2 with $\log|f| = -g$, then critical points of g coincide with critical points of f.

13.6.6 The local behaviour of g

Let $p_2 \in R^1, p_2 \neq p_1$, and assume that p_2 is not a critical point of g. Then there is a unique curve C through p_2 with the following property. If f is analytic near p_2 with $\log|f| = -g$ then $\arg f$ is constant on C. Further, g is strictly monotone on C. To see this note that if f and f^* are both analytic with $\log|f| = -g$ then f^*/f is constant. The curve C is a level curve of the function $\mathrm{Im}(\log f)$, this function a harmonic conjugate of $-g$.

Suppose next that p_2 is a critical point of g. Then there are $n \geq 2$ curves through p_2 on each of which $\arg f$, for f as above, is constant. To see this, choose n so that $(f(p) - f(p_2))^{1/n}$ has a simple zero at p_2. This allows us to write $f(p) = G(F(p)^n)$ with G and F locally one-one. Choose a curve for G through $F(p_2)^n$: then this curve has n pre-images through $F(p_2)$.

Indeed, let $q_2 = \log f(p_2)$ and draw a horizontal and a vertical straight line through q_2. This forms four quadrants which we label 1, 2, 3, 4 counter-clockwise, starting at the top right in the usual way. There are $4n$ curves emanating from p_2, in which n is the multiplicity of the zero of $f(p) - f(p_2)$ at p_2, and $4n$ "sector-like" regions with vertex at p_2. On $2n$ of these curves we have $\arg f$ constant, while $\log|f|$ is constant on the other $2n$. There are n curves emanating from p_2 on which $\arg f$ is constant and $\log|f|$ decreases, and as we cross one of these in the counter-clockwise sense, $\arg f$ decreases (here counter-clockwise is interpreted with respect to the image under a local chart).

13.6.7 The main step

Define a function f_1, analytic near p_1, with $\log|f_1| = -g$. Then f_1 has a simple zero at p_1. Let h be the inverse function of f_1, defined on a neighbourhood of 0. Let r be the supremum of positive s such that h extends to be analytic on $D(0,s)$. Then h is analytic on $D(0,r)$. Since f_1 may be analytically continued along any path in R^1, starting at p_1, and since $f_1 \circ h$ is the identity near 0, it follows that

$$g(h(w)) = \log \frac{1}{|w|}, \quad w \in D(0,r). \tag{13.6}$$

We see now that $H = h(D(0,r))$ is a connected set on which $g(p) > \log 1/r$.

Let $p_n \to p^*$, with $p_n \in H$ and $p^* \in \partial H$. Then we may write $p_n = h(w_n)$ and without loss of generality $w_n \to w^*$ with $|w^*| \leq r$. Suppose that $|w^*| < r$. Then $p_n \to h(w^*)$, and $p^* = h(w^*)$ is an interior point of H, by the open mapping theorem. This is a contradiction, and hence $|w^*| = r$ and $g(p_n) = \log 1/|w_n| \to \log 1/r$. We deduce that H is a component, and so by Lemma 13.6.3 the unique component, of the set $\{p \in R^1 : g(p) > \log 1/r\}$.

13.6.8 Lemma

h is locally univalent on $D(0,r)$, and $H = h(D(0,r))$ contains no critical point of g.

Proof. Take $w_1 \in D(0,r)$, and analytically continue f_1 along the image under h of the line segment from 0 to w_1. This gives a function f_2 analytic near w_1, with $\log|f_2| = -g$, and $f_2(h(w)) = w$ near w_1. So h must be one-one near w, and f_2 must be one-one near $h(w)$.

13.6.9 Lemma

h is univalent on $D(0,r)$.

Proof If $|w_j| < r$ and $h(w_1) = h(w_2), w_1 \neq w_2$, then by (13.6) we have $|w_1| = |w_2|$. Suppose now that $0 \leq \theta_1 < \theta_2 < 2\pi$ and set

$$E = \{t \in (0,1) : h(te^{i\theta_1}) = h(te^{i\theta_2})\}.$$

Obviously $(0,1) \setminus E$ is open, and this set is non-empty since h is one-one near 0. Suppose $t_1 \in E$, and let $p_3 = h(w_1), w_j = t_1 e^{i\theta_j}$. Let T_j be the image under h of the ray $\arg w = \theta_j$, and analytically continue f_1 along the curves T_j. This gives F_1, F_2 analytic on a parametric disc V centred at p_3, and $F_j \circ h(w) = w$ near w_j. Thus $\arg F_j$ is constant on an arc of T_j passing through p_3.

By 13.6.6 there is a unique curve C passing through p_3 on which $\arg F_j$ is constant, and g is strictly monotone on C. Since $g(h(te^{i\theta_j})) = \log 1/t$, we have $h(te^{i\theta_1}) = h(te^{i\theta_2})$ for t close to t_1. Hence E is open and so empty, by connectedness.

The next lemma is now obvious.

13.6.10 Lemma

f_1 extends to be analytic and univalent, with $f_1 = h^{-1}$, on the simply connected domain $H = h(D(0,r))$.

If $r = 1$ then we have finished, and we assume henceforth that $r < 1$. Since $H = h(D(0,r))$ is the unique component of the set $\{p \in R^1 : g(p) > \log 1/r\}$, the closure of H does not meet ∂N, by Lemma 13.6.1.

13.6.11 Lemma

Suppose that $\zeta \in \partial H$, and that ζ is not a critical point of g. Then f_1 extends analytically and univalently to a neighbourhood V of ζ.

Proof. Let G be analytic near ζ, with $\mathrm{Re}(G) = g$. Then G is one-one near ζ, and we let V be the pre-image under G of a disc centred at $G(\zeta)$. Then $V \cap H$ is connected, since its image under G is a half-disc. Choose f_2 analytic near ζ, with $\log|f_2| = -g$. Thus f_2 is univalent near ζ, and f_2/f_1 is constant on $V \cap H$. Multiplying f_2 by a constant gives the required extension. Since $|f_2(p)| \geq r$ on $V \setminus H$ the extended function remains univalent.

13.6.12 Lemma

Let $\theta \in [0, 2\pi)$, and define $\gamma_\theta(t) = h(te^{i\theta})$ for $0 \leq t < r$. Then there exists $\zeta \in \partial H$ such that $\gamma_\theta(t) \to \zeta$ as $t \to r$.

Proof. Since R has no divergent curves, the curve $\gamma(t) = \gamma_\theta(t)$ visits some compact set through a sequence tending to r. Thus there exists $\zeta \in R^1$ such that $\gamma(t_n) \to \zeta$ through a sequence $t_n \to r$, and ζ is a boundary point of $H = h(D(0,r))$, since $g(\gamma(t_n)) = \log 1/t_n \to \log 1/r$. Finally, $\gamma(t) \to \zeta$ by 13.6.6, since $\arg f_1$ is constant on γ.

13.6.13 Lemma

g has a critical point on ∂H.

Proof. Assume not. Then for each $\theta \in [0, 2\pi)$, the curve $\gamma_\theta(t)$ tends to $\zeta = \zeta_\theta \in \partial H$, and f_1 extends analytically and univalently to a neighbourhood V of ζ. We may assume that $W = f_1(V)$ is a disc.

We have $f_1(\zeta) = re^{i\theta}$, since

$$f_1(\zeta) = \lim f_1(\gamma(t_n)) = \lim t_n e^{i\theta},$$

in which t_n increases with limit r. Let h^* be the inverse function of f_1, mapping $W = f_1(V)$ onto V. On

$$W \cap D(0,r) = f_1(V \cap H)$$

we have $h = f_1^{-1}$ and $h^* = f_1^{-1}$, and so h^* extends h to $D(0,r) \cup W$.

We do this for each θ, and obtain an extension of h to a disc W_θ centred at $re^{i\theta}$. Since the intersection of any two W_θ is connected, and meets $D(0,r)$ unless the intersection is void, it follows by compactness that this permits us to extend h analytically to a larger disc $D(0,r')$, contradicting the choice of r.

13.6.14 A closed curve

Choose a critical point ζ of g on ∂H. There exist (at least) two curves $\eta_j, j = 1, 2$ emanating from ζ, on which g increases and $\arg f$ is constant, for any any f analytic on a neighbourhood V of ζ with $\log|f| = -g$. Note that f_1/f is constant on every connected subset of $V \cap H$. These curves lie, apart from their starting point, in H, and so by the constancy of $\arg f_1$ on η_j we see that each η_j is the image T_j under h of a ray $\arg w = \theta_j$. Mark "lower" and "upper" sides of T_j as T_j^l, T_j^u, so that $\arg f_1(p)$ increases as p crosses from T_j^l to T_j^u. By 13.6.6, we go from T_j^u to T_j^l as we cross T_j moving counter-clockwise around ζ.

Thus the union of T_1, T_2, p_1 and ζ gives a closed curve σ on R with two well defined "edges', which we will label "positive" and "negative".

13.6.15 Lemma

$R \setminus \sigma$ is path-connected.

Proof. We shall show that every $p \in R \setminus \sigma$ can be joined to p_0 by a path avoiding σ. This is true if $p \in N$ or if $g(p) \leq \log 1/r$, by Lemma 13.6.4. Suppose now that $g(p) > \log 1/r$, so that $p \in H$. Now only finitely many curves $\gamma_\theta(t)$ can land at ζ, and so we first move from p to a point p' not lying on any of these. Following a curve $\arg f_1 = c$ from p' we land at $\zeta' \in \partial H, \zeta' \neq \zeta$, and since $g(\zeta') \leq \log 1/r$ we can continue on to p_0.

13.6.16 Lemma

There exists a continuous function from R to $\mathbb{C} \setminus \{0\}$, not having a continuous logarithm.

Proof. To the cut surface $R \setminus \sigma$ we adjoin two copies of σ, labelled σ^+, σ^-, corresponding to the positive and negative edges of σ. Let the resulting space be X. We construct a continuous function $q : X \to [0, 1]$, with $q = 0$ on σ^- and $q = 1$ on σ^+. The function $Q = \exp(2\pi i q)$ will then be well-defined and continuous on R, but does not have a continuous logarithm q^* on R, because for any such q^* the function $q^* - q$ would be constant on $R \setminus \sigma$.

To construct q, cover σ by finitely many closed parametric discs P_j, each contained in a small open parametric disc D_j. Let $F_j : R \to [0, 1]$ be continuous, with $F_j = 0$ on P_j, and $F_j = 1$ off D_j. We then define G_j on X to be the same as F_j, except that $G_j = 1$ on σ^+ and on all points of D_j on the "positive" side of σ. Thus $G_j : X \to [0, 1]$ is continuous, with $G_j = 0$ on $\sigma^- \cap P_j$, and $G_j = 1$ on σ^+. Finally set $q(x) = \min\{F_j(x)\}$.

This result contradicts the following standard lemma, and the proof of Lemma 13.5.5 is complete.

13.6.17 Lemma

Let S be any simply connected Riemann surface, and let $Q : S \to \mathbb{C} \setminus \{0\}$ be continuous. Then Q has a continuous logarithm on S.

Proof. Fix $a \in S$ and assume without loss of generality that $Q(a) = 1$. Let σ_1, σ_2 be paths in S joining a to b. Then $Q(\sigma_1), Q(\sigma_2)$ are homotopic paths in $\mathbb{C} \setminus \{0\}$ starting at 1, and by the ordinary monodromy theorem the continuations of $\log w$ along these paths agree near $Q(b)$, so that $\log Q(b)$ is well defined.

Chapter 14

The Phragmén-Lindelöf principle

14.1 Introduction

This represents a refinement of the maximum principle for subharmonic and analytic functions. The classical proofs have largely been supplanted by use of harmonic measure. We begin with:

14.1.1 Lemma

Let D be a domain in $\mathbb{C}$ and let u be subharmonic and bounded above on D, with

$$\limsup_{z\to\zeta, z\in D} u(z) \le 0$$

for all finite boundary points of D. Then $u(z) \le 0$ on D.

This follows at once from Lemma 9.2.7. The next lemma is a refinement of Lemma 14.1.1 for functions having slow growth as z tends to infinity in D.

14.1.2 Lemma: the classical Phragmén-Lindelöf principle

Let D be a domain in $\mathbb{C}$ and let u be subharmonic on D, such that

$$\limsup_{z\to\zeta, z\in D} u(z) \le 0$$

for all finite boundary points z of D. Suppose further that there exists $v(z)$ harmonic on D, with

$$\liminf_{z\to\zeta, z\in D} v(z) \ge 0$$

for every finite boundary point z of D, and such that for every $\delta > 0$ we have

$$\limsup_{z\to\infty, z\in D} (u(z) - \delta v(z)) \le 0.$$

Then $u(z) \le 0$ on D.

Proof. Fix w in D. For $\delta > 0$ the maximum principle gives

$$u(w) - \delta v(w) \le 0$$

and we just let $\delta \to 0$.

14.1.3 Corollary

Let $R > 0$ and $M > 0$ and let $-\pi \leq a < b \leq \pi$. Let f be analytic on the domain

$$D = \{z : \quad |z| > R, \quad a < \arg z < b\},$$

with

$$\limsup_{z \to \zeta, z \in D} |f(z)| \leq M < \infty$$

for all finite boundary points ζ of D. Assume that

$$\log |f(z)| < |z|^s \tag{14.1}$$

for all large z in D, in which $s < S < \pi/(b-a)$. Then $|f(z)| \leq M$ in D.

Proof. We may clearly assume that $M = 1$ (otherwise replace f by $f(z)/M$, which does not affect the existence of an s as in (14.1)). By considering $f(ze^{it})$ in place of f, for some fixed t, we may assume that $b > 0, a = -b$. Thus $s < S < \pi/2b$.

Take $u(z) = \log |f(z)|$ and

$$v(z) = |z|^S \cos(S \arg z) = \mathrm{Re}(z^S).$$

For z in D we have

$$|S \arg z| \leq Sb < \pi/2, \quad \cos(S \arg z) \geq \cos Sb = \mu > 0$$

and so

$$v(z) \geq |z|^S \mu.$$

Thus, for every $\delta > 0$ we have

$$u(z) - \delta v(z) \to -\infty$$

as $z \to \infty$ in D. By Lemma 14.1.2, we get $u(z) \leq 0$ on D.

This result is sharp: to see this, take $0 < b \leq \pi$ and $a = -b$ and $f(z) = \exp(z^{\pi/(b-a)})$. Then f is bounded on the finite boundary of D but unbounded in D.

Thus the narrower the sectorial region D is, the faster f has to grow in D in order to not be bounded. We will see a far-reaching generalization of this idea in the section on the Carleman-Tsuji estimate for harmonic measure.

14.2 Applications

The next two results are among the most useful applications of this strand of ideas.

14.2.1 Theorem

Let D be an unbounded simply connected domain in $\mathbb{C}$, not the whole plane. Let f be analytic and bounded on D, and continuous on $D \cup \partial D$. Assume that $f(z) \to 0$ as z tends to infinity on ∂D. Then $f(z) \to 0$ as z tends to infinity in D.

Proof. Assume without loss of generality that $|f(z)| \leq 1$ on D. Let $0 < \delta < 1$ and let E be a closed subset of ∂D such that $|f(z)| < \delta$ on $\partial D \setminus E$. Since ∞ is a regular point of $X = \partial_\infty D$, by §9.1.4, we get

$$\omega(z, E, D) \to 0, \quad z \to \infty,$$

using §10.1.9. But now the two-constants theorem gives

$$|f(z)| \leq \delta + \omega(z, E, D)$$

and the result follows.

14.2.2 Theorem

Let D be a simply connected domain as in Theorem 14.2.1, such that the boundary of D consists of two simple curves C_1, C_2 both tending to infinity, and disjoint apart from their common starting point $a \in \partial D$. Let f be analytic and bounded in the domain D, and continuous in $D \cup C_1 \cup C_2$. Assume that $f(z) \to a_j$ as z tends to infinity on C_j. Then $a_1 = a_2$.

Proof. It is clear that the a_j are finite, since f is bounded. Assume $a_1 \neq a_2$ and apply Theorem 14.2.1 to $g(z) = (f(z) - a_1)(f(z) - a_2)$. Thus $g(z) \to 0$ as z tends to infinity in the closure of D. Let $\varepsilon > 0$, and take $M > 0$ such that $|g(z)| < \varepsilon$ for $z \in D, |z| > M$.

We now use the fact that $J = C_1 \cup C_2 \cup \{\infty\}$ is a Jordan curve on the Riemann sphere (in particular J cannot contain a disc), and so a rotation of D is a Jordan domain in $\mathbb{C}$. Take a curve I which lies in the closure of D and joins C_1 to C_2, with $|z| > M$ for all z on I. Such a curve exists by Theorem 11.5.3: take the Riemann mapping h from $D(0,1)$ to D and extend it to a homeomorphism on $|z| \leq 1$. The curve I is then the image of an arc of a circle centred at $h^{-1}(\infty)$. We have $|g(z)| < \varepsilon$ on I and so, by connectedness, either $f(z) - a_1$ is small on all of I or $f(z) - a_2$ is small on all of I. This contradicts the fact that $f(z) - a_j$ is small for large z on C_j.

Chapter 15

The Carleman-Tsuji estimate for harmonic measure

15.1 The Carleman-Tsuji estimate

15.1.1 Parseval's formula for a continuous function

Let f be a continuous real-valued function on $[-\pi, \pi]$. Define the Fourier coefficients

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx.$$

These are uniformly bounded. As shown in the section on Poisson's formula (8.2.2),

$$F(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\frac{e^{it}+z}{e^{it}-z}dt = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt + \frac{1}{2\pi i}\int_{|w|=1}\frac{2f(\arg w)}{w-z}dw$$

is analytic in $D(0,1)$ with $c_0 = F(0) = \frac{1}{2}a_0$. Also $u = \mathrm{Re}(F)$ is bounded and, as $z \to e^{is}$, $-\pi < s < \pi$, we have $u(z) \to f(s)$.

Differentiation gives, for $n > 0$,

$$F^{(n)}(z) = \frac{n!}{2\pi i}\int_{|w|=1}\frac{2f(\arg w)}{(w-z)^{n+1}}dw$$

and

$$c_n = \frac{F^{(n)}(0)}{n!} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt = a_n - ib_n.$$

Thus Taylor's theorem applied to F gives

$$u(re^{it}) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} r^n(a_n\cos nt + b_n\sin nt),$$

the series uniformly convergent on each closed disc $|z| \le r < 1$. The orthogonality of the trigonometric functions gives

$$I(r) = \frac{1}{\pi}\int_{-\pi}^{\pi} u(re^{it})^2dt = \frac{1}{2}a_0^2 + \sum_{n=1}^{\infty} r^{2n}(a_n^2 + b_n^2).$$

Since $u(re^{it})$ is uniformly bounded and tends pointwise to $f(t)$ on $(-\pi,\pi)$ the dominated convergence theorem gives Parseval's formula

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)^2 dt = \frac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2+b_n^2).$$

15.1.2 Wirtinger's inequality

Suppose that f is a real-valued function such that f' is continuous on $[a,b]$ and $f(a)=f(b)=0$. Then

$$\int_a^b f'(x)^2 dx \geq \frac{\pi^2}{(b-a)^2}\int_a^b f(x)^2 dx.$$

Proof. It suffices to prove this when $a=0, b=\pi$. Extend f to an odd function on $[-\pi,\pi]$. In the Fourier expansion of f we have $a_n=0$ and

$$b_n = \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\, dx.$$

Further, f' can be extended to an even function h on $[-\pi,\pi]$ with $\int_{-\pi}^{\pi} h(x)dx = 2\int_0^{\pi} f'(x)dx = f(\pi)-f(0)=0$. The Fourier expansion of h has no $\sin nx$ terms and has

$$A_n = \frac{2}{\pi}\int_0^{\pi} f'(x)\cos nx\, dx = nb_n,$$

using integration by parts. Parseval's formula gives

$$\frac{2}{\pi}\int_0^{\pi} f(t)^2 dt = \sum_{n=1}^{\infty} b_n^2 \leq \sum_{n=1}^{\infty} n^2 b_n^2 = \sum_{n=1}^{\infty} A_n^2 = \frac{2}{\pi}\int_0^{\pi} f'(t)^2 dt.$$

15.1.3 Definition

For $0<t<\infty$ and a domain D in $\mathbb{C}$ we define $\theta_D^*(t)$ as follows. If D contains the whole circle $|z|=t$ then $\theta_D^*(t)=\infty$. If $D\cap\{z:|z|=t\}$ is not the whole circle $|z|=t$ then it consists of countably many open arcs, and we define $\theta_D^*(t)$ to be the angular measure of the longest of these (if one has angular measure $s>0$ then at most finitely many can have angular measure $>s$). Note that if $\theta_D^*(t)>y$, then $D\cap\{z:|z|=t\}$ contains a closed arc A of angular measure y, and D contains a neighbourhood of A. Thus $\theta_D^*(t')>y$ for t' close to t and so $\theta_D^*(t)$ is measurable (we've shown that $-\theta_D^*(t)$ is upper semi-continuous i.e. $\theta_D^*(t)$ is LSC).

Obviously if D,U are domains with $D\subseteq U$ then $\theta_D^*(t)\leq\theta_U^*(t)$.

15.1.4 The Carleman-Tsuji estimate: a special case

Let $0<r<\infty$ and let D be a domain in $\mathbb{C}$ with $0\in D$ such that D meets the circle $|z|=r$. Assume that there exist a positive increasing sequence $\rho_n\to\infty$ and a finite subset $\{\theta_j\}$ of $[0,2\pi]$ such that ∂D consists of:

(i) arcs of circles $|z|=\rho_n$, each such circle contributing at most finitely many arcs;

(ii) radial line segments $z = se^{i\theta_j}, \rho_n \leq s \leq \rho_{n+k}$;

Let D_r be the component of $D \cap D(0,r)$ containing 0. Let $\theta_r = \partial D_r \setminus \partial D$.

Then θ_r is a subset of the circle $|z| = r$, since $w \in \theta_r$ implies that $|w| \leq r$ and that w is a limit point of D and so in D.

Let

$$u(z) = \omega(z, \theta_r, D_r)$$

and extend u to a function v subharmonic in $D(0,r)$, by setting $v = 0$ in $D(0,r) \setminus D_r$. To do this, note that if $w \in D(0,r) \cap \partial D_r$ then $u(z) \to 0$ as $z \to w$ with z in D_r.

Since $v \geq 0$ we see that v^2 is upper semi-continuous. Also for $|z_0| < r$ and small $s > 0$, Cauchy-Schwarz gives

$$v(z_0)^2 \leq \left(\frac{1}{2\pi} \int_0^{2\pi} v(z_0 + se^{i\theta}) d\theta \right)^2 \leq \frac{1}{2\pi} \int_0^{2\pi} v(z_0 + se^{i\theta})^2 d\theta$$

and so v^2 is subharmonic. For $0 < \rho < r$ let

$$m(\rho) = \frac{1}{2\pi} \int_0^{2\pi} v(\rho e^{i\theta})^2 d\theta = \frac{1}{2\pi} \int_{\theta_\rho} u(\rho e^{i\theta})^2 d\theta \tag{15.1}$$

in which $\theta_\rho = D_r \cap \{z : |z| = \rho\}$ for $0 < \rho < r$. Then, by Theorem 9.2.1, $m(\rho)$ is a convex non-decreasing function of $\log \rho$ on $(0,r)$ and in particular m is continuous. Also, since u is harmonic and so continuous at 0, we have $\lim_{\rho \to 0+} m(\rho) = u(0)^2$.

By 1.1 the derivative $\mu = \frac{\partial m}{\partial \log \rho}$ exists on $J = (0,r) \setminus E_0$, where E_0 is a countable set, and μ is non-decreasing on J.

Claim 1: *μ is positive on J.*

To prove the claim we note that u is harmonic and non-constant near the origin, using the identity theorem for harmonic functions. So near the origin u is the real part of a non-constant analytic function and there are constants a_n, b_n such that we can write, for small ρ,

$$u(\rho e^{i\theta}) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \rho^n (a_n \cos n\theta + b_n \sin n\theta)$$

and

$$m(\rho) = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} \rho^{2n} (a_n^2 + b_n^2)$$

so that $m'(\rho) > 0$. This proves the claim.

Let $\rho_n < \rho < \rho_{n+1}$. Then θ_ρ consists of finitely many open arcs of $|z| = \rho$. On $P_n = \{z \in \partial D_r : \rho_n < |z| < \rho_{n+1}\}$ we have $u = 0$, and P_n consists of finitely many open radial segments, across which u can be extended by the Schwarz reflection principle 11.7.1. So all partial derivatives of u extend continuously up to P_n.

Let $t = \log \rho$. For $\rho_n < \rho < \rho_{n+1}$,

$$m_t = \frac{1}{\pi} \int_{\theta_\rho} u u_t d\theta. \tag{15.2}$$

Also, writing u locally as a harmonic function of $\log z = t + i\theta$,

$$m_{tt} = \frac{1}{\pi}\int_{\theta_\rho} (u_t)^2 + uu_{tt}d\theta = \frac{1}{\pi}\int_{\theta_\rho} (u_t)^2 - uu_{\theta\theta}d\theta \tag{15.3}$$

and so integration by parts gives

$$m_{tt} = \frac{1}{\pi}\int_{\theta_\rho} (u_t)^2 + (u_\theta)^2 d\theta \geq 0. \tag{15.4}$$

By (15.2) and Cauchy-Schwarz,

$$(m_t)^2 \leq \frac{1}{\pi^2}\int_{\theta_\rho} u^2 d\theta \int_{\theta_\rho} (u_t)^2 d\theta \tag{15.5}$$

and so

$$\frac{(m_t)^2}{2m} \leq \frac{1}{\pi}\int_{\theta_\rho} (u_t)^2 d\theta. \tag{15.6}$$

Define $\theta^*(\rho)$ for $\rho_n < \rho < \rho_{n+1}$ as follows. If θ_ρ consists of the whole circle $|z| = \rho$ then put $\theta^*(\rho) = \infty$. If θ_ρ is not the whole circle $|z| = \rho$ then it consists of finitely many open arcs θ_ρ^j. Then $\theta^*(\rho) = \theta_D^*(\rho)$ is the angular length of the longest of these.

In the second case we get by Wirtinger's inequality 15.1.2

$$\int_{\theta_\rho^j} (u_\theta)^2 d\theta \geq \frac{\pi^2}{|\theta_\rho^j|^2}\int_{\theta_\rho^j} u^2 d\theta \geq \frac{\pi^2}{\theta^*(\rho)^2}\int_{\theta_\rho^j} u^2 d\theta,$$

since u vanishes at the end-points, and summing gives

$$\frac{1}{\pi}\int_{\theta_\rho} (u_\theta)^2 d\theta \geq \frac{2\pi^2}{\theta^*(\rho)^2} m(\rho), \quad \rho \neq \rho_n. \tag{15.7}$$

Thus (15.4), (15.6) and (15.7) give

$$m_{tt} \geq \frac{(m_t)^2}{2m} + \frac{1}{2}\left(\frac{2\pi}{\theta^*(\rho)}\right)^2 m(\rho), \quad \rho \neq \rho_n. \tag{15.8}$$

Put

$$t = \log \rho, \quad t_n = \log \rho_n, \qquad M(t) = m(\rho), \quad h(t) = \frac{2\pi}{\theta^*(\rho)}. \tag{15.9}$$

Then (15.8) becomes

$$M'' \geq \frac{(M')^2}{2M} + \frac{1}{2}h^2 M, \quad t_n < t < t_{n+1}. \tag{15.10}$$

In particular, using Claim 1,

$$M'(t) > 0, \quad M''(t) > 0, \quad t \neq t_n. \tag{15.11}$$

Thus

$$L'' + (L')^2 \geq \frac{1}{2}(L')^2 + \frac{1}{2}h^2,$$

and

$$2L'' + (L')^2 \geq h^2, \quad L = \log M. \tag{15.12}$$

This gives

$$(M''/M')^2 = (L''/L' + L')^2 \geq (L')^2 + 2L'' \geq h^2$$

and so, using (15.11),

$$M''/M' \geq h, \quad t_n < t < t_{n+1}. \tag{15.13}$$

So for $t_n < s < s' < t_{n+1}$ we have

$$M'(s') \geq M'(s)\exp(\int_s^{s'} h(t)dt).$$

Iterating this and using the fact that $M'(s) \leq M'(s')$ for $s, s' \notin \{t_n\}$ with $s < s'$, since M is convex, we get

$$M'(\tau) \geq M'(t)\exp(\int_t^{\tau} h(s)ds), \quad -\infty < t < \tau < \log r, \quad t, \tau \notin \{t_n\}. \tag{15.14}$$

Now put

$$t = \log\rho, \quad \sigma = e^{\tau}, \quad t^* = \log r \tag{15.15}$$

and assume that $k \in (0,1)$. Suppose that $0 < \rho < kr$. Then, since M is continuous, non-negative and non-decreasing,

$$1 \geq \lim_{\tau \to t^*} M(\tau) \geq \int_t^{t^*} M'(\tau)d\tau,$$

we get, using (15.14),

$$\begin{aligned}
1 &\geq M'(t)\int_t^{t^*} \exp\left(\int_t^{\tau} h(s)ds\right)d\tau \\
&= M'(t)\int_\rho^r \exp\left(\int_\rho^{\sigma} \frac{2\pi dx}{x\theta^*(x)}\right)\frac{d\sigma}{\sigma} \\
&\geq M'(t)\int_{kr}^r \exp\left(\int_\rho^{\sigma} \frac{2\pi dx}{x\theta^*(x)}\right)\frac{d\sigma}{\sigma} \\
&\geq M'(t)\int_{kr}^r \exp\left(\int_\rho^{kr} \frac{2\pi dx}{x\theta^*(x)}\right)\frac{d\sigma}{\sigma} \\
&\geq (1-k)M'(t)\exp\left(\int_\rho^{kr} \frac{2\pi dx}{x\theta^*(x)}\right)
\end{aligned}$$

since

$$\int_{kr}^r \frac{d\sigma}{\sigma} = \int_k^1 \frac{d\sigma}{\sigma} \geq \int_k^1 d\sigma = 1-k.$$

This gives

$$M'(t) \leq (1-k)^{-1}\exp\left(-\int_\rho^{kr} \frac{2\pi dx}{x\theta^*(x)}\right), \quad 0 < \rho = e^t < kr, \quad 0 < k < 1. \tag{15.16}$$

Now (15.10) gives

$$M''(t) \geq \frac{1}{2}h(t)^2 M(t) \geq \frac{1}{2}h(t)M(t) = \frac{\pi m(\rho)}{\theta^*(\rho)}. \tag{15.17}$$

Let $t < \tau$ with $t, \tau \notin \{t_j\}$ and let $t_\nu < \ldots < t_n$ be those t_j lying in (t, τ). Since M' is non-decreasing,

$$M'(\tau) \geq M'(\tau) - M'(t_n^+) + M'(t_n^+) - M'(t_n^-) + \ldots + M'(t_\nu^+) - M'(t_\nu^-) + M'(t_\nu^-) - M'(t) \geq \int_t^{\tau} M''(s)ds$$

and so (15.17) gives

$$M'(\tau) \geq \pi \int_\rho^\sigma \frac{m(x)dx}{x\theta^*(x)} \geq \pi m(\rho) \int_\rho^\sigma \frac{dx}{x\theta^*(x)}. \tag{15.18}$$

Using (15.16), with t replaced by $\tau = \log \sigma$, and (15.18), we now obtain, for $\rho < \sigma < kr, 0 < k < 1$,

$$\pi m(\rho) \int_\rho^\sigma \frac{dx}{x\theta^*(x)} \leq M'(\tau) \leq (1-k)^{-1} \exp\left(-\int_\sigma^{kr} \frac{2\pi dx}{x\theta^*(x)}\right)$$

and in particular

$$\pi m(\rho) \int_\rho^\sigma \frac{dx}{x\theta^*(x)} \leq (1-k)^{-1} \exp\left(-\int_\rho^{kr} \frac{2\pi dx}{x\theta^*(x)} + \int_\rho^\sigma \frac{2\pi dx}{x\theta^*(x)}\right). \tag{15.19}$$

If $0 < \rho < kr$ and

$$\int_\rho^{kr} \frac{2\pi dx}{x\theta^*(x)} > 1 \tag{15.20}$$

then we choose $\sigma \in (\rho, kr)$ with

$$\int_\rho^\sigma \frac{2\pi dx}{x\theta^*(x)} = 1$$

and (15.19) gives

$$m(\rho) \leq \frac{2e}{(1-k)} \exp\left(-2\pi \int_\rho^{kr} \frac{dx}{x\theta^*(x)}\right). \tag{15.21}$$

On the other hand if (15.20) fails then the RHS of (15.21) is at least $2/(1-k) > 2 > m(\rho)$. Thus (15.21) always holds.

Letting $\rho \to 0+$ we get $m(\rho) \to u(0)^2$ and so

$$u(0) = \omega(0, \theta_r, D_r) \leq \frac{(2e)^{1/2}}{(1-k)^{1/2}} \exp\left(-\pi \int_0^{kr} \frac{dt}{t\theta^*(t)}\right), \quad 0 < k < 1. \tag{15.22}$$

15.1.5 The Carleman-Tsuji estimate: the main step

Let $0 < r < \infty$ and let D be a semi-regular domain containing 0 and meeting the circle $|z| = r$. Let D_r be the component of $D \cap D(0, r)$ containing 0, and let $H_r = \partial D_r \setminus \partial D$, so that $H_r \subseteq D \cap \{z : |z| = r\}$. Note that H_r is a relatively open subset of ∂D_r. Note also that D_r is semi-regular (if x is a boundary point of D and D_r then a barrier for x, D will serve for x, D_r, while if x is in H_r then x satisfies the condition 9.1.8). Let E_0 be a compact subset of ∂D_r such that $E_0 \subseteq H_r$ and let $u(z) = \omega(z, E_0, D_r)$.

Let n be a positive integer, and define building blocks of the n'th stage to be the sets

$$\{z : |z| \leq 2^{-n} r\}, \quad \{z : p2^{-n} r \leq |z| \leq (p+1)2^{-n} r, \quad \pi q 2^{-n} \leq \arg z \leq \pi(q+1)2^{-n}\}, \quad p, q \in \mathbb{N}.$$

Let D_n^* be the union of all blocks of the n'th stage which are contained in D, and let D_n be that component of the interior D_n^{**} of D_n^* which contains 0. Obviously $D_n^* \subseteq D_{n+1}^*$ and so $D_n^{**} \subseteq D_{n+1}^{**}$ and $D_n \subseteq D_{n+1}$.

Claim 1: $D = \bigcup_{n=1}^\infty D_n$.

To see this, take $z \in D$ and join 0 to z by a path γ in D. If n is large then 2^{-n} is small compared to $\mathrm{dist}(\gamma, \partial D)$ and so a neighbourhood of γ lies in D_n^*. This proves Claim 1.

Note that it follows that D_n, for large n, meets the circle $|z| = r$.

For each n, let $D_n(r)$ be the component of $D_n \cap D(0,r)$ containing 0, and define $\theta_n(r) = \partial D_n(r) \setminus \partial D_n$. Since $D_n \subseteq D_{n+1}$ we clearly have $D_n(r) \subseteq D_{n+1}(r)$.

Claim 2: $\bigcup_{n=1}^{\infty} D_n(r) = D_r$.

Since $D_n \subseteq D$ we have $D_n(r) \subseteq D_r$. Now join 0 to z in D_r by a path γ in D_r. Then for large n we have $\gamma \subseteq D_n^{**}$ and so $\gamma \subseteq D_n$.

Claim 3: let $x \in E_0$; then there exists $\delta > 0$ such that, for all sufficiently large n, we have $D(x,\delta) \cap D(0,r) \subseteq D_n(r)$ and $\{z : |z| = r\} \cap D(x,\delta) \subseteq \theta_n(r)$.

To see this, note that since $x \in H_r$ we have $x \in D$. Thus by Claim 1 there exists $\delta > 0$ such that $D(x,\delta)$ is in D_p for all sufficiently large p. Thus $V = D(x,\delta) \cap D(0,r) \subseteq D_p \cap D(0,r)$ for all large p.

Since x is a boundary point of D_r, there exists $y \in V \cap D_r$. By Claim 2, we have $y \in D_n(r)$ for all large n, and so $V \subseteq D_n(r)$ for large n, since $V \subseteq D_n \cap D(0,r)$ and V is connected. Since $D(x,\delta) \cap \{z : |z| = r\} \subseteq Cl(D_n(r)) \cap D_n$, the second assertion of Claim 3 follows.

It follows from compactness that the same n will serve for all $x \in E_0$, if sufficiently large.

Let $u_n(z) = \omega(z, \theta_n(r), D_n(r))$ for large n. Since $D_n(r) \subseteq D_{n+1}(r) \subseteq D$, §15.1.4 gives

$$u_n(0) = \omega(0, \theta_n(r), D_n(r)) \leq \frac{(2e)^{1/2}}{(1-k)^{1/2}} \exp\left(-\pi \int_0^{kr} \frac{dt}{t\theta_D^*(t)}\right). \tag{15.23}$$

Further, $u_n \leq u_{n+1}$, by the comparison principle, and

$$v(z) = \lim_{n\to\infty} u_n(z)$$

is harmonic in D_r, by Harnack's theorem, using Claim 2 and the fact that $u_n \leq 1$.

We compare u to v. By Claim 3, if $x \in E_0$ then x is an interior point of $\theta_n(r)$ for large n. Again by Claim 3, there is some $\delta > 0$ such that $D(0,r) \cap D(x,\delta) \subseteq D_n(r)$ and as $z \to x$ with $z \in D(0,r)$ we have $1 \geq v(z) \geq u_n(z) \to 1$ (note here that $D_n(r)$ is regular).

If $x \in \partial D_r \setminus E_0$ then $x \in \partial D$ or $|x| = r$. Since D is semi-regular, it follows that, with finitely many exceptions, $u(z) \to 0$ as $z \to x$ from within D_r. Thus we get $u \leq v$ on D_r. Since E_0 is an arbitrary compact subset of H_r we have, using (15.23),

$$\omega(0, H_r, D_r) \leq \frac{(2e)^{1/2}}{(1-k)^{1/2}} \exp\left(-\pi \int_0^{kr} \frac{dt}{t\theta_D^*(t)}\right). \tag{15.24}$$

15.1.6 The Carleman-Tsuji estimate

Let D be a semi-regular domain in $\mathbb{C}$ and let $z \in D$. Let $0 < r < \infty, 0 < k < 1$ and $2|z| \leq kr$. Let D_r be the component of $D \cap D(0,r)$ containing z. Then with $S(0,r)$ the circle $|z| = r$,

$$u(z) = \omega(z, S(0,r), D_r) \leq \frac{3(2e)^{1/2}}{(1-k)^{1/2}} \exp\left(-\pi \int_{2|z|}^{kr} \frac{dt}{t\theta_D^*(t)}\right). \tag{15.25}$$

Proof. We should more precisely write $S(0,r)\cap\partial D_r$ in place of $S(0,r)$ on the LHS of (15.25). However, the statement here is slightly stronger than that in Tsuji's book, in which only $\omega(z, S(0,r)\cap D, D_r)$ is considered.

Assume first that D meets $S(0,r)$ and let $H_r = \partial D_r \setminus \partial D$. Let $U = D \cup D(0, 2|z|)$ and let U_r be the component of $U \cap D(0,r)$ containing z. Since $D \subseteq U$ we have $D_r \subseteq U_r$ and $H_r \subseteq L_r = \partial U_r \setminus \partial U$. Now

$$\theta_U^*(t) = \theta_D^*(t), \quad (2|z| < t < r),$$

and $\theta_U^*(t) = \infty$ for $0 < t < 2|z|$. Thus, using the comparison principle and Harnack's inequality,

$$\omega(z, H_r, D_r) \le \omega(z, L_r, U_r) \le 3\omega(0, L_r, U_r)$$

and (15.24) gives

$$\omega(z, H_r, D_r) \le \frac{3(2e)^{1/2}}{(1-k)^{1/2}} \exp\left(-\pi \int_{2|z|}^{kr} \frac{dt}{t\theta_D^*(t)}\right). \tag{15.26}$$

Now take s with $r - s$ small and positive. Let G be the component of $D \cap D(0,s)$ containing z, and let $L = \partial G \setminus \partial D$. Obviously $G \subseteq D_r$. Also $\theta_G^*(t) \le \theta_D^*(t)$ and (15.26) gives

$$v(z) = \omega(z, L, G) \le \frac{3(2e)^{1/2}}{(1-k)^{1/2}} \exp\left(-\pi \int_{2|z|}^{ks} \frac{dt}{t\theta_D^*(t)}\right). \tag{15.27}$$

We compare $u(w)$ to $v(w)$ on G. If $w \to x \in L$, then $v(w) \to 1$, since L is a relatively open subset of $S(0,s)$. On the other hand, if $w \to x \in \partial G \setminus L = \partial G \cap \partial D$ then $|x| \le s$. Thus x is in the closure of D_r but not in D, and so x is in the relatively open set $\partial D_r \setminus S(0,r)$. Provided x is a regular boundary point of D_r it follows that $u(w) \to 0$ as $w \to x$, and we have already seen in §15.1.5 that D_r is semi-regular since D is. Hence, with finitely many exceptions, $\limsup_{w\to x\in\partial G}(u(w) - v(w)) \le 0$ and so $u(z) \le v(z)$. Since s is arbitrary in (15.27), we get (15.25).

Remark. With the above notation let Y_r be the part of ∂D lying in $|z| \ge r$. Using the comparison theorem we get

$$\omega(z, D(0,r)\cap\partial D_r, D_r) \le \omega(z, D(0,r)\cap\partial D_r, D) \le \omega(z, D(0,r)\cap\partial D, D),$$

since evidently $D(0,r)\cap\partial D_r \subseteq D(0,r)\cap\partial D$. Taking complements we get

$$\omega(z, Y_r, D) \le \omega(z, S(0,r)\cap\partial D_r, D_r)$$

and (15.25) can be applied again.

The next theorem is a typical application of this estimate 15.1.6 and is a powerful refinement of Lemma 14.1.1 and Corollary 14.1.3.

15.1.7 Theorem

Let v be subharmonic on the semi-regular domain D in $\mathbb{C}$, and assume that

$$\limsup_{z\to\zeta, z\in D} v(z) \le 0$$

for every finite boundary point ζ of D. Assume further that $r_n \to \infty$ and

$$B(r_n, v) \exp\left(-\pi \int_1^{r_n/2} \frac{dt}{t\theta_D^*(t)}\right) \to 0$$

as $n \to \infty$, in which

$$B(r, v) = \sup\{v(z) : z \in D, |z| = r\}.$$

Then $v(z) \leq 0$ on D.

Proof. Assume that $v(z) > 0$ for some z in D. The two-constants theorem gives

$$v(z) \leq B(r_n, v)\omega(z, S(0, r_n), D_{r_n})$$

and applying the Carleman-Tsuji estimate 15.1.6 the RHS tends to 0.

Note that if D is a sectorial region $\{z : |z| > R, a < \arg z < b\}$ then

$$\theta_D^*(t) = (b - a), \quad \exp\left(-\pi \int_1^{r/2} \frac{dt}{t\theta_D^*(t)}\right) = (r/2)^{-\pi/(b-a)},$$

and so Corollary 14.1.3 is a special case of this result.

15.1.8 Boundary behaviour of harmonic measure: revisited

Let D be a semi-regular domain in $\mathbb{C}$ with $0 \in X = \delta_\infty D$. Let $E \subseteq X$ be closed, with $0 \notin E$. If 0 is regular for D then by §10.1.9 we know that $\omega(z, E, D) \to 0$ as $z \to 0, z \in D$, whereas if 0 is not regular then Example 10.1.10 shows that this may fail.

On the other hand if $S(0, t)$ meets the complement of D for every $t > 0$ then $\theta_D^*(t) \leq 2\pi$ for $t > 0$ and the Carleman-Tsuji estimate shows that for a given component D_r of $D \cap D(0, r)$ we have $\omega(z, S(0, r), D_r) \to 0$ as $z \to 0$. Indeed, for this to hold it is only necessary that $S(0, t)$ meet $\mathbb{C} \setminus D$ for a sufficiently "thick" set of t tending to 0.

Chapter 16

Two fundamental results on asymptotic values

16.1 Transcendental singularities of the inverse function

Let f be non-constant and meromorphic in the plane, let $a \in \mathbb{C}$ and $t > 0$, and let $C(t)$ be a (non-empty) component of the set $\{z \in \mathbb{C} : |f(z) - a| < t\}$.

Then $1/(f - a)$ must be unbounded on $C(t)$. To see this, assume without loss of generality that $t = 1$, and suppose that $1/(f - a)$ is bounded on $C(1) = D$. Then $1/(f - a)$ is analytic on D, and setting

$$u(z) = \log \frac{1}{|f(z) - a|} \quad (z \in D), \quad u(z) = 0 \quad (z \notin D),$$

defines a non-constant bounded subharmonic function in $\mathbb{C}$. This is a contradiction.

Assume now that we have a family of such components $C(t), 0 < t < t_0$, with the property that $C(t) \subseteq C(s)$ for $t < s$. Then there are two cases to consider.

Case 1: *there exists* $z_0 \in \bigcap_{0<t<t_0} C(t)$.
Then evidently $f(z_0) = a$. Let $s > 0$ and pick r with $0 < r \leq s$ such that $f - a$ has no zeros in $0 < |z - z_0| \leq r$. Let $T = \min\{|f(z) - a| : |z - z_0| = r\}$. Then $C(t) \subseteq D(z_0, r) \subseteq D(z_0, s)$ for $t < T$. In particular, $\bigcap_{0<t<t_0} C(t) = \{z_0\}$ and $C(t)$ is bounded for small positive t.

Case 2: $\bigcap_{0<t<t_0} C(t) = \emptyset$.
For each large positive integer n, choose $z_n \in C(n)$ and a path γ_n from z_n to z_{n+1} in $C(n)$. The union of these gives a path $\gamma(t)$ such that $f(\gamma(t)) \to a$ as $t \to \infty$.

We assert that $\gamma(t) \to \infty$ as $t \to \infty$. Assume not. Then there exist $t_m \to \infty$ with $|\gamma(t_m)| \leq M < \infty$ for all $m \in \mathbb{N}$, and without loss of generality $\gamma(t_m) \to w \in \mathbb{C}$ as $m \to \infty$. Since $f(\gamma(t_m)) \to a$ we must have $f(w) = a$. Let L be a positive real number. Then $|f(z) - a| < L$ on some open neighbourhood U_L of w. Since $\gamma(t_m) \to w$ and $\gamma(t_m) \in C(L)$ for large m, we have $U_L \cap C(L) \neq \emptyset$ and hence $w \in U_L \subseteq C(L)$. This contradicts the assumption that the $C(t)$ have empty intersection.

We have thus shown in Case 2 that $f(z)$ tends to a along a path tending to infinity. In particular, if f is transcendental then a is an asymptotic value of f and the $C(t)$ are said to determine a transcendental singularity of f^{-1} over a.

Note that transcendental singularities do not arise for rational functions, as Case 2 for rational functions simply corresponds to $f(\infty)=a$. Further, critical values of a meromorphic function f are sometimes referred to as algebraic singularities of f^{-1}.

Conversely, suppose that the transcendental meromorphic function $f(z)$ tends to a as z tends to infinity along a path γ in $\mathbb{C}$. Then for each positive real number t there exists a unique component $C(t)$ of the set $C'(t)=\{z\in\mathbb{C}:|f(z)-a|<t\}$, such that $C(t)$ contains an unbounded subpath of γ. It is clear that $C(t)\subseteq C(s)$ if $0<t<s$. Since the $C(t)$ are all unbounded, they must satisfy Case 2, and their intersection must be empty.

A transcendental singularity of f^{-1} over a is said to be direct if $C(t)$, for some $t>0$, contains finitely many zeros of $f-a$. Since the intersection of all the $C(t)$ is empty there then exists $t_1>0$ such that none of these zeros lies in $C(t)$ for $t>t_1$. In particular $C(t)$, for small positive t, contains no zeros of $f-a$. The contrary case is that of an indirect singularity, in which $C(t)$ contains infinitely many zeros of $f-a$, for every $t>0$. Transcendental singularities of f^{-1} over ∞, direct or otherwise, are defined by considering $1/f$.

For example, the function $z/\sin z$ tends to infinity along the positive real axis, and this singularity is indirect, while ze^z has direct singularities over 0 and ∞.

16.2 The Denjoy-Carleman-Ahlfors theorem

16.2.1 Lemma

Let $n\geq 2$ be an integer. Let $D_j, j=1,\ldots,n$ be pairwise disjoint domains, and let u_j be non-constant subharmonic functions such that u_j vanishes outside D_j. Assume that $h(r)$ is a positive function such that that, for each j, we have $B(r,u_j)\leq O(h(r))$ as $r\to\infty$. Then we have

$$\liminf_{r\to\infty}\frac{h(r)}{r^{n/2}}>0.$$

Proof. Since each u_j is non-constant and vanishes outside D_j, each domain D_j must be unbounded. Let $\theta_j(t)=\theta^*_{D_j}(t)$ be defined as in 15.1.3. Note that if t is large then $\theta_j(t)<2\pi$, because $n\geq 2$ and the circle $|z|=t$ meets D_k for $k\neq j$. Theorem 15.1.7 implies that for each j, as $r\to\infty$,

$$\pi\int_1^r\frac{dt}{t\theta_j(t)}\leq\log B(2r,u_j)+O(1)\leq\log h(2r)+O(1). \tag{16.1}$$

But, since the D_j are pairwise disjoint, the Cauchy-Schwarz inequality gives

$$n^2=\left(\sum_{j=1}^n\theta_j(t)^{1/2}\theta_j(t)^{-1/2}\right)^2\leq\sum_{j=1}^n\theta_j(t)\sum_{j=1}^n\theta_j(t)^{-1}\leq 2\pi\sum_{j=1}^n\theta_j(t)^{-1}$$

if t is large. Thus for large r we have, using (16.1),

$$n^2\log r-O(1)\leq 2\sum_{j=1}^n\pi\int_1^r\frac{dt}{t\theta_j(t)}\leq 2n\log h(2r)+O(1),$$

and this proves the lemma.

16.2.2 Theorem (Denjoy-Carleman-Ahlfors)

Suppose that f is transcendental and meromorphic in the plane, and that the inverse function f^{-1} has $n \geq 2$ direct transcendental singularities, lying over $a_1, \ldots, a_n$ (not necessarily distinct). Then

$$\liminf_{r \to \infty} \frac{T(r,f)}{r^{n/2}} > 0. \tag{16.2}$$

In particular, the lower order of f is at least $n/2$.

Moreover, if F is transcendental and meromorphic in the plane and F^{-1} has a direct transcendental singularity over ∞ and $F(z)$ is bounded on a path tending to infinity then

$$\liminf_{r \to \infty} \frac{T(r,F)}{r^{1/2}} > 0.$$

Proof. To prove the first part assume that all the a_j are finite. Thus there exists $\delta > 0$ such that for each $j = 1, \ldots n$ we can find a non-empty component D_j of the set $\{z \in \mathbb{C} : |f(z) - a_j| < \delta\}$, such that $f(z) \neq a_j$ on D_j and such that the D_j are pairwise disjoint (if $a_j = a_k$ then D_j, D_k are distinct and so disjoint components).

For each j we define a non-constant subharmonic function u_j by

$$u_j(z) = \log \left| \frac{\delta}{(f(z) - a_j)} \right| \quad (z \in D_j), \quad u_j(z) = 0 \quad (z \notin D_j).$$

By Theorem 8.3.5,

$$B(r, u_j) = \sup\{u_j(z) : |z| = r\}$$

satisfies

$$B(r, u_j) \leq 3 \int_0^{2\pi} u_j(2re^{it})dt \leq 3m(2r, 1/(f - a_j)) + O(1) \leq 3T(2r, f) + O(1).$$

The result then follows by applying Lemma 16.2.1 with $h(r) = T(2r, f)$.

To prove the second part, apply the first part to $f(z) = F(z^2)$, which has two direct singularities over ∞.

The theorem is sharp, since e^z has order 1 and two direct transcendental singularities (over $0, \infty$).

16.2.3 Corollary

Let f be a transcendental entire function with $n \geq 1$ finite asymptotic values. Then f satisfies (16.2).

Proof. We can assume that there are simple paths $\gamma_j \to \infty, j = 1, \ldots, n$, pairwise disjoint except that each starts at 0, and such that $f(z) \to a_j \in \mathbb{C}$ as $z \to \infty$ on γ_j. (To ensure that each path is simple we can first approximate by a stepwise curve and then delete any repeated segments of the curve).

Let $\gamma_{n+1} = \gamma_n$. In the region D_j between γ_j and γ_{j+1} the function f must be unbounded, by Theorem 14.2.2. This gives us n direct transcendental singularities over ∞, with f bounded on the intermediate paths, and proves the result.

This theorem is also sharp, as

$$f(z) = \int_0^z \frac{\sin t}{t} dt$$

has order 1, and two finite asymptotic values.

16.3 Two lemmas needed for the Bergweiler-Eremenko theorem

The Bergweiler-Eremenko theorem is a striking result from [20] connecting the critical and asymptotic values of a meromorphic function, which has subsequently found widespread application in value distribution theory. The result shows that if f is a transcendental meromorphic function of finite order then any direct transcendental singularity of f^{-1} must be a limit point of critical values. We will present the subsequent modification by Hinchliffe [46], which shows that the result remains true for functions of finite lower order.

The proof will require the following lemma [59, p.287] on isolated singularities of the inverse function.

16.3.1 Lemma

Let f be transcendental and meromorphic in the plane, and let $0 < S < \infty$, and let C be a component of the set $\{z : S < |f(z)| \leq \infty\}$. Let $z_0 \in C$ with $w_0 = f(z_0)$ finite and $f'(z_0) \neq 0$, and let g be that branch of the inverse function f^{-1} which maps w_0 to z_0. Suppose that g admits unrestricted analytic continuation in the annulus $S < |w| < \infty$, starting at w_0. Then C is simply connected, and contains either one pole (possibly multiple) of f, or no pole of f but instead a path σ tending to infinity on which $f(z) \to \infty$.

Proof. We may assume that $S = 1$. Choose v_0 such that $e^{v_0} = w_0 = f(z_0)$. Then, starting at v_0,

$$h(v) = g(e^v) = f^{-1}(e^v)$$

admits unrestricted analytic continuation in the half-plane U given by $\mathrm{Re}(v) > 0$. By the monodromy theorem, h then extends to an analytic function on U, with $f(h(v)) = e^v$.

Next, h maps U into C. Indeed, $h(U) = C_0 = \{z \in C : f(z) \neq \infty\}$, for if $z_1 \in C_0$ we can choose a simply connected domain C_1 with $\{z_0, z_1\} \subseteq C_1 \subseteq C_0$. Since f maps C_0 into $1 < |w| < \infty$, we may define an analytic branch of $F = \log f$ on C_1, mapping C_1 into U. Further, $e^F = f$ maps z_0 to w_0 and $h(F)$ is the identity near z_0, and this remains the case throughout C_1 by the identity theorem. Thus $z_1 = h(F(z_1)) \in h(F(C_1)) \subseteq h(U)$.
There are now two possibilities to consider.

Case 1: suppose that h is univalent on U. In this case the image under $z = h(v)$ of $\mathrm{Re}(v) = 1$ is a simple curve L, on which $|f(z)| = e$. We assert that L must tend to infinity in both directions. Since $f(h(1+k2\pi i)) = e$ for every integer k, it is clear that $h(1+iy)$ must be unbounded as $y \to +\infty$, and as $y \to -\infty$, in both cases with y real. If we have $|h(1+iy)| \leq M < \infty$ for arbitrarily large $|y|$, with y real, then there must be infinitely many points on the circle $|z| = 2M$ with $|f(z)| = e$. This is impossible, since f is transcendental, and so $h(1+iy) \to \infty$, as asserted.

Next, the function $H(v) = 1/(h(v) - h(1/2))$ is bounded on $\mathrm{Re}(v) \geq 1$, by the open mapping theorem and the assumption that h is univalent. Thus, by the Phragmén-Lindelöf principle, we have $H(v) \to 0$, and $h(v) \to \infty$, as $v \to \infty$ with $\mathrm{Re}(v) \geq 1$. It follows that C_0 is an unbounded simply connected domain and for the path σ we may take $h : [1, \infty) \to C_0$.

We deduce that C cannot contain a pole of f. To see this, suppose z_2 is a pole of f in C, and take a sequence u_n in C_0, with $u_n \to z_2$, and $s_n \in U$ such that $h(s_n) = u_n$. Since $e^{s_n} = f(h(s_n)) = f(u_n) \to \infty$, we get $s_n \to \infty, \mathrm{Re}(s_n) > 1$, and so $u_n = h(s_n) \to \infty$, which is a contradiction. Thus $C = C_0$ and C is simply connected; further, $F = \log f$ may be defined on C, mapping z_0 to v_0, and $h(F)$ is the identity near z_0 and so throughout C, while $F(h)$ is the identity near v_0 and so on U. This completes the proof of the lemma in this case.

Case 2: Suppose that we have $v_1, v_2 \in U$ with $v_1 \neq v_2$, $h(v_1) = h(v_2)$. Then $e^{v_1} = f(h(v_1)) = f(h(v_2)) = e^{v_2}$ and so $v_2 = v_1 + m2\pi i$ for some integer m. If v is close to v_1 then the open mapping theorem tells us that h takes the value $h(v)$ at some v' close to v_2, and we must have $(v' - v)/2\pi i \in \mathbb{Z}$ and so $v' = v + m2\pi i$. Thus h has period $m2\pi i$ near v_1 and so throughout U.

Let k be the smallest positive integer such that h has period $k2\pi i$. In this case the function $G(\zeta) = h(k \log \zeta) = g(\zeta^k)$ is analytic and univalent in $W = \{\zeta : 1 < |\zeta| < \infty\}$, and maps W onto C_0. To see this, just note that G can be analytically continued along any path in W, and continuation in W once around 0 leads back to the same function element, by the periodicity of h. Since G is univalent, $z_1 = \lim_{\zeta \to \infty} G(\zeta)$ exists.

Suppose that $z_1 = \infty$. If τ is large, then G takes the value τ at ζ with ζ large, and this gives $f(\tau) = f(G(\zeta)) = f(g(\zeta^k)) = \zeta^k$ so that $f(\tau)$ is large. But this gives $\lim_{\tau \to \infty} f(\tau) = \infty$, contradicting the fact that f is transcendental. Thus z_1 is finite. The same argument shows that $f(z)$ is large for z close to z_1, and so z_1 is a pole of f.

We now see that G is univalent on $W^* = W \cup \{\infty\}$, mapping W^* onto $C_0 \cup \{z_1\}$, which is therefore simply connected, so that z_1 is the only pole of f in C. This may also be seen as follows: if $u_n \in C_0$ and $f(u_n) \to \infty$ take $\zeta_n \in W$ with $G(\zeta_n) = u_n$. Then $\zeta_n^k = f(g(\zeta_n^k)) = f(G(\zeta_n)) = f(u_n) \to \infty$ and so $\zeta_n \to \infty$ and $u_n = G(\zeta_n) \to z_1$.

Finally, we note that since $z = G(\zeta) = g(\zeta^k)$ is univalent on W^*, it follows that $\zeta = G^{-1}(z) = f(z)^{1/k}$ is meromorphic and univalent on C, so that z_1 is a pole of multiplicity k. This completes the proof of Lemma 16.3.1.

The next lemma [55] gives an estimate for the length of level curves of a meromorphic function, and is a slightly more precise version of [56, Lemma 2].

16.3.2 Lemma

Let G be transcendental and meromorphic in the plane, and let $\alpha \in (1, \infty)$. For $w \in \mathbb{C}$ and positive r and R, let $L(r, w, R, G)$ denote the length of the level curves $|G(z)| = R$ lying in $D(w, r)$, and set $L(r, R, G) = L(r, 0, R, G)$. Let $\psi(t)$ be continuous, positive and non-decreasing on $[1, \infty)$ such that

$$\int_1^\infty \frac{1}{t\psi(t)} dt < \frac{\log \alpha}{4}. \tag{16.3}$$

Then if the positive constant S is large enough there exist uncountably many $R \in (S, 2S)$ such that

$$L(r, R, G)^2 \leq cr^2 \psi(\alpha r)(T(\alpha r, G) + \log S), \quad r \geq 1, \tag{16.4}$$

in which c is a positive constant depending only on α.

Proof We use the length-area inequality as in [39, Theorem 2.1, p.29] (see also [73, p.44]). Let Δ be an open disc in $\mathbb{C}$ of area A. Then

$$\int_S^{2S} \frac{L(\Delta, R, G)^2}{p(\Delta, R, G)R} dR \leq 2\pi A, \tag{16.5}$$

in which $L(\Delta, R, G)$ is the length of the curves $|G(z)| = R$ in Δ and

$$p(\Delta, R, G) = \frac{1}{2\pi} \int_0^{2\pi} n(\Delta, Re^{i\phi}, G) d\phi, \tag{16.6}$$

where $n(\Delta, a, G)$ is the number of roots of $G(z) = a$ in Δ, counting multiplicity.

Denote by c_j positive constants depending only on α. Set $\beta = \sqrt{\alpha}$ and $r_q = \beta^q, q = 0, 1, 2, \ldots$. Then (16.3) gives

$$\sum_{q=1}^{\infty} \frac{1}{\psi(r_q)} \leq \frac{1}{\log \beta} \int_1^{\infty} \frac{1}{t\psi(t)} dt < \frac{1}{2}. \tag{16.7}$$

Assume that S is large. Then for ϕ real and for $S \leq R \leq 2S$ we have $\infty \geq |G(0) - Re^{i\phi}| \geq 1$. This gives, for $r \geq 1$,

$$\begin{aligned} n(D(0,r), Re^{i\phi}, G) &\leq n(r, Re^{i\phi}, G) \\ &\leq c_0 N(\beta r, 1/(G - Re^{i\phi})) \\ &\leq c_0 T(\beta r, G - Re^{i\phi}) + C_1^* \\ &\leq c_0 T(\beta r, G) + c_0 \log R + c_0 \log 2 + C_1^* \\ &\leq c_0 T(\beta r, G) + c_1 \log S. \end{aligned}$$

Here the constant C_1^* only arises if $G(0) = \infty$. Substituting this estimate into (16.5) and (16.6) gives, for $r \geq 1$,

$$\int_S^{2S} \frac{L(r, R, G)^2}{R} dR \leq c_2 r^2 \left(T(\beta r, G) + \log S \right).$$

Hence if the positive constant c_3 is chosen large enough then for each $q \in \mathbb{N}$ there exists a subset E_q of $(S, 2S)$ with

$$\int_{E_q} \frac{dR}{R} < \frac{\log 2}{\psi(r_q)} \tag{16.8}$$

such that for all $R \in (S, 2S) \setminus E_q$ and for $r_{q-1} < r \leq r_q$ we have

$$\begin{aligned} L(r, R, G)^2 &\leq L(r_q, R, G)^2 \\ &\leq c_3 r_q^2 \psi(r_q)(T(\beta r_q, G) + \log S) \\ &\leq c_4 r^2 \psi(\alpha r)(T(\alpha r, G) + \log S). \end{aligned}$$

Since (16.7) and (16.8) give

$$\int_{\bigcup_{q=1}^{\infty} E_q} \frac{dR}{R} < \frac{\log 2}{2}, \tag{16.9}$$

(16.4) follows.

16.4 The Bergweiler-Eremenko theorem: preliminaries

Following [20, 46], the key step is to prove the following proposition.

16.4.1 Proposition

Let f be transcendental and meromorphic of finite lower order in the plane, such that f^{-1} has an indirect transcendental singularity over 0. Let the components $C(t)$ be as in §16.1. Then for every $t > 0$ the component $C(t)$ contains infinitely many zeros of f'.

The proof of Proposition 16.4.1 will take up the whole of this section. Assume throughout that f is transcendental and meromorphic of finite lower order in the plane, and that f^{-1} has an indirect transcendental singularity over 0, such that $C(\varepsilon)$, for some $\varepsilon > 0$, contains finitely many zeros of f'. By reducing ε, if necessary, it may be assumed that $C(\varepsilon)$ contains no zeros of f'.

16.4.2 Lemma

Let $0 < \delta < \varepsilon$. Let $z_1 \in C(\delta)$, with $f(z_1) = 0$. Then there exist a with $0 < |a| = r < \delta$ and a simply connected domain $D \subseteq C(\delta)$, such that f maps D univalently onto $D(0, r)$, and D contains a path σ tending to infinity on which $f(z) \to a$ as $z \to \infty$, mapped by f onto the line segment $w = ta, 0 \leq t < 1$.

Proof. Let g be that branch of the inverse function f^{-1} which maps 0 to z_1. Next, let r be the supremum of positive real t such that g extends to be analytic in $D(0, t)$. We have $r > 0$, since f is univalent on a neighbourhood of z_1. Further, g is analytic on $D(0, r)$, and univalent there, since $g(w_1) = g(w_2)$ gives $w_1 = f(g(w_1)) = f(g(w_2)) = w_2$. Moreover, $D = g(D(0, r))$ is a simply connected domain and $|f(z)| \to r$ as z tends to the finite boundary ∂D, and so D is that component of the set $\{z : |f(z)| < r\}$ which contains z_1. It follows that $r < \delta$, for otherwise we would have $C(\delta) \subseteq D$, which contradicts the fact that $C(\delta)$ contains infinitely many zeros of f. In particular, we now have $D \subseteq C(\delta)$.

Now suppose that, for every a with $|a| = r$, the branch g of f^{-1} can be analytically continued along the line segment $w = ta, 0 \leq t \leq 1$. Then each such continuation defines an extension h_a of g to a disc $U_a = D(a, d_a), d_a > 0$. If $U_a \cap U_b \neq \emptyset$, then $h_a = h_b = g$ on the non-empty intersection $U_a \cap U_b \cap D(0, r)$. Since $U_a \cap U_b$ is connected we get $h_a = h_b$ on $U_a \cap U_b$. But the circle $|w| = r$ is compact, and can be covered by finitely many such U_a, from which it follows that g extends analytically to a disc $D(0, r_1), r_1 > r$, and this is a contradiction.

It follows that there is some a with $|a| = r$ such that g does not admit analytic continuation along the path $w = ta, 0 \leq t \leq 1$. Now the path $g(ta), 0 \leq t < 1$, lies in D and so in $C(\delta)$, and so does its closure in the finite plane, since $r < \delta$. It follows that, as $t \to 1-$, $g(ta)$ must tend either to infinity or to a critical point of f, and the latter is ruled out since f has no critical points in $C(\delta)$. Thus we obtain the path σ.

16.4.3 Lemma

There exist points $z_j \to \infty, z_j \in C(\varepsilon)$, and distinct complex numbers a_j with $0 < |a_j| < \varepsilon/2$, and pairwise disjoint simply connected domains $D_j \subseteq C(\varepsilon)$, with $0 \notin D_j$, with the following properties. First, f maps D_j univalently onto $D(0, r_j)$, with $f(z_j) = 0$. Second, each D_j contains a path $\sigma_j \to \infty$ on which $f(z) \to a_j$ as $z \to \infty$, and the path σ_j is mapped by f onto the line segment $w = ta_j, 0 \leq t < 1$.

Proof. The z_j, a_j will be defined inductively. Take $z_1 \in C(\frac{1}{2}\varepsilon)$ with $f(z_1) = 0$, and let $a = a_1, D =$

$D_1, \sigma = \sigma_1$ be as in Lemma 16.4.2. Assuming that z_{n-1}, D_{n-1} have already been determined, we need only take $z_n \in C(\frac{1}{2}r_{n-1})$, with $f(z_n) = 0$ and $z_n \neq z_j, 1 \leq j \leq n-1$, and determine D_n, r_n, a_n, σ_n as in Lemma 16.4.2. We assert that the D_j are pairwise disjoint. If $m < n$ and D_n meets D_m then, since D_n is a component of the set $\{z : |f(z)| < r_n\}$, we have $D_n \subseteq D_m$. But this is a contradiction since $z_n \neq z_m$ and f is univalent on D_m. It now follows that the D_j may be chosen so that $0 \notin D_j$, by deleting one of the D_j if necessary.

16.4.4 Lemma

Let the z_j, a_j and D_j be as in Lemma 16.4.3. For $t > 0$, let $t\theta_j(t)$ be the length of the longest open arc of $|z| = t$ which lies in D_j. As z tends to infinity on σ_j, we have

$$\log \frac{r_j}{|f(z) - a_j|} \geq \int_{|z_j|}^{|z|} \frac{dt}{t\theta_j(t)} - \log 2. \tag{16.10}$$

Proof. The function $h_j(z) = f(z)/r_j$ maps D_j univalently onto Δ, with z_j mapped to 0. By §7.2.3 and Lemma 7.2.5 we then have

$$\log\left(\frac{1 + |h_j(z)|}{1 - |h_j(z)|}\right) = [z_j, z]_{D_j} \geq \int_{|z_j|}^{|z|} \frac{dt}{t\theta_j(t)}. \tag{16.11}$$

But f maps σ_j onto the line segment $w = ta_j, 0 \leq t < 1$, and so $1 - |h_j(z)| = |f(z) - a_j|/r_j$. Since $\log 2 \geq \log(1 + |h_j(z)|)$, (16.10) now follows from (16.11).

16.4.5 Lemma

Let u lie on σ_j. Then there exists v on σ_j, with $|u| \leq |v| \leq |u| + 1$, such that

$$\max\{|f(v) - a_j|, |f'(v)|\} \leq |f(u) - a_j|.$$

Proof. Starting at u, follow σ_j in the direction in which $|f(z) - a_j|$ decreases. Then σ_j describes an arc γ joining the circle $|z| = |u|$ and $|z| = |u| + 1$. Then the inverse function $g = f^{-1}$ maps a sub-segment I of $[f(u), a_j)$ onto γ, and so

$$1 \leq \left| \int_I g'(\zeta) d\zeta \right| \leq |f(u) - a_j| \max\{|g'(\zeta)| : \zeta \in I\}.$$

16.4.6 A sequence on which $T(r, f')$ grows slowly

Since f has finite lower order, so has f', and hence there exist a real number $M > 12$ and a sequence (s_n) tending to infinity such that

$$T(s_n^5, f') + T(s_n^5, 1/f') \leq s_n^M. \tag{16.12}$$

Let N, K and L be integers, with N/M, K/N and L/K large. Set

$$G(z) = z^N f'(z), \tag{16.13}$$

and apply Lemma 16.3.2 to $1/G$, with $\alpha = e^8$ and $\psi(t) = t$. This gives a small positive η such that G has no critical values w with $|w| = \eta$ and such that

$$L(r, \eta, G)^2 = O(r^3 T(\alpha r, G)) = O(r^3 T(\alpha r, f'))$$

as $r \to \infty$. In particular, since $M > 12$,

$$L(s_n^4, \eta, G) = O(s_n^6 T(\alpha s_n^4, f')^{1/2}) = O(s_n^{6+M/2}) \leq s_n^M \tag{16.14}$$

as $n \to \infty$.

16.4.7 Lemma

For each large n there exist t_n, T_n satisfying

$$s_n^{1/2} - 1 \leq t_n \leq s_n^{1/2}, \quad s_n^2 \leq T_n \leq s_n^2 + 1, \tag{16.15}$$

such that

$$\max\{|\log|f'(z)|| : z \in S(0, t_n) \cup S(0, T_n)\} \leq s_n^{M+1}. \tag{16.16}$$

Proof. By (16.12) and standard estimates (see §3.2.7), the number of zeros and poles of f' in $s_n^{1/4} \leq |z| \leq s_n^4$, counting multiplicity, is

$$q_n \leq n(s_n^4, f') + n(s_n^4, 1/f') \leq s_n^M.$$

Label these zeros and poles as $w_1, \ldots, w_{q_n}$ and let U_n be the union of the discs $D(w_j, s_n^{-M-1})$, $j = 1, \ldots, q_m$. Then the discs of U_n have sum of radii at most s_n^{-1} and so for large n it is possible to choose t_n, T_n satisfying (16.15) and such that the circles $S(0, t_n), S(0, T_n)$ do not meet U_n. But then standard estimates based on the Poisson-Jensen formula give, for $z \in S(0, t_n) \cup S(0, T_n)$,

$$|\log|f'(z)|| \leq O(T(s_n^3, f')) + O(q_n \log s_n) \leq s_n^{M+1}$$

if n is large enough.

16.4.8 Lemma

Let $\tau > 0$ and for large n let t_n and T_n be as in Lemma 16.4.7. Provided the positive integer N was chosen large enough, any component C_n of the set

$$\{z \in \mathbb{C} : t_n < |z| < T_n, \ |G(z)| < \eta\}$$

satisfies

$$\textit{diam} f(C) < \tau.$$

Proof. Fix $z^* \in C$ and choose any $z \in C$. Join z^* to z by a path λ in the closure of C consisting of part of the ray $\arg u = \arg z^*, |u| > t_n$, part of the circle $|u| = |z|$, and part of ∂C. Since

$$|f'(u)| \leq \eta |u|^{-N} \leq \eta s_n^{-N/4}$$

on λ this gives, using (16.14),

$$|f(z) - f(z^*)| \leq 2\pi\eta|z|^{1-N} + \eta \int_{t_n}^{\infty} t^{-N} dt + s_n^M \eta s_n^{-N/4} = o(1),$$

which proves the lemma.

16.4.9 Lemma

Let $Q \in \mathbb{N}, Q \geq 4L$. Let $n \in \mathbb{N}$ be large and let $E_1, \ldots, E_Q$ be pairwise disjoint domains such that for each j and each $t > 0$ the circle $S(0,t)$ is not contained in E_j. For $t > 0$ let $\phi_j(t)$ be the angular measure of $S(0,t) \cap E_j$.

Then at least $Q - 2L$ of the domains $E_1, \ldots, E_Q$ are such that

$$\pi \int_{[4s_n^{1/2}, s_n/4]} \frac{dt}{t\phi_j(t)} > K \log s_n \quad \text{and} \quad \pi \int_{[4s_n, s_n^2/4]} \frac{dt}{t\phi_j(t)} > K \log s_n. \tag{16.17}$$

Proof. Assume that at least L of the E_j, without loss of generality $D_1, \ldots, E_L$, are such that the second inequality of (16.17) fails (in particular, the closure of each of these E_j must therefore meet $S(0, 4s_n)$ and $S(0, s_n^2/4)$, because otherwise the given integral would evidently be infinite, and hence each of these E_j must meet $S(0,t)$ for $4s_n < t < s_n^2/4$). By our assumption,

$$\pi \int_{[4s_n, s_n^2/4]} \sum_{j=1}^{L} \frac{dt}{t\phi_j(t)} \leq LK \log s_n. \tag{16.18}$$

But the Cauchy-Schwarz inequality gives, for $t \in (4s_n^{1/2}, s_n^2/4)$,

$$L^2 \leq \left(\sum_{j=1}^{L} \phi_j(t) \right) \left(\sum_{j=1}^{L} \frac{1}{\phi_j(t)} \right) \leq 2\pi \left(\sum_{j=1}^{L} \frac{1}{\phi_j(t)} \right).$$

On integrating from $4s_n$ to $s_n^2/4$ and using (16.18), this leads to

$$L^2 \log(s_n/16) \leq 2LK \log s_n,$$

an obvious contradiction if n is large enough, since L and K were chosen in §16.4.6 with L/K large. The same argument shows that it is not possible for the first inequality of (16.17) to fail for L of the E_j.

16.4.10 Completion of the proof of Proposition 16.4.1

Let $a_1, \ldots, a_{6L}$ be as in Lemma 16.4.3, and choose τ such that

$$0 < \tau < \varepsilon/4, \quad 4\tau < \min\{|a_j - a_{j'}| : 1 \leq j < j' \leq 6L\}. \tag{16.19}$$

Let $n \in \mathbb{N}$ be large, and apply Lemma 16.4.9 to the domains $D_1, \ldots, D_{6L}$ corresponding to $a_1, \ldots, a_{6L}$ as in Lemma 16.4.3. Let $u_j \in \sigma_j$ with $|u_j| = s_n$. Then applying Lemma 16.4.4 gives

$$\log \frac{r_j}{|f(u_j) - a_j|} \geq K \log s_n - O(1),$$

by (16.10) and (16.17) for at least $4L$ of the a_j, which after re-labelling we may assume are $a_1, \ldots, a_{4L}$. Applying Lemma 16.4.5 now gives v_j satisfying

$$v_j \in D_j, \quad s_n \leq |v_j| \leq s_n + 1, \quad \max\{|f(v_j) - a_j|, |f'(v_j)|\} \leq s_n^{1-K}, \quad j = 1, \ldots, 4L. \tag{16.20}$$

Since K/M is large this gives, in particular, $|G(v_j)| < \eta$, where G and η are as in (16.13) and (16.14). Let t_n and T_n be as in Lemma 16.4.7, and let C_j be that component of the set

$$\{z \in \mathbb{C} : t_n < |z| < T_n, \, |G(z)| < \eta\}$$

which contains v_j. Then the diameter of $f(C_j)$ is at most τ by Lemma 16.4.8, and

$$|f(z) - a_j| \leq \tau + s_n^{1-K} \leq \tau + o(1)$$

for all $z \in C_j$. In particular this implies using (16.19) that the following hold:
(a) $C_1, \ldots, C_{4L}$ are pairwise disjoint;
(b) each of $C_1, \ldots, C_{4L}$ lies in $C(\varepsilon)$ and so contains no zeros of f';
(c) for $1 \leq j, j' \leq 4L$, $j \neq j'$ the component C_j does not meet the path $\sigma_{j'}$ of Lemma 16.4.3, since $f(z) = a_{j'} + o(1)$ for $z \in \sigma_{j'}, |z| > t_n$, and in particular C_j cannot contain a circle $S(0,t), t > 0$.

Lemma 16.4.9 may now be applied again, this time with $E_j = C_j$ and $\phi_j(t)$ the angular measure of $C_j \cap S(0,t)$, and it may be assumed without loss of generality that (16.17) holds for $j = 1, \ldots, 2L$. The Carleman-Tsuji estimate for harmonic measure (§15.1.6) and the conformal invariance of harmonic measure now give

$$\begin{aligned} \omega(v_j, C_j, S(0,T_n) \cup S(0,t_n)) &\leq c_1 \exp\left(-\pi \int_{2|v_j|}^{T_n/2} \frac{dt}{t\phi_j(t)}\right) + c_1 \exp\left(-\pi \int_{2t_n}^{|v_j|/2} \frac{dt}{t\phi_j(t)}\right) \\ &\leq c_2 s_n^{-K}, \end{aligned} \tag{16.21}$$

using (16.15), (16.17) and (16.20), in which c_1, c_2 are positive constants independent of j and n.

Since

$$|f'(z)| = \eta |z|^{-N} \geq \eta T_n^{-N} \geq \frac{1}{2}\eta s_n^{-2N} \quad \text{for} \quad z \in \partial C_j \setminus (S(0,t_n) \cup S(0,T_n)),$$

the two constants theorem 10.2.10 may be applied to $\log|1/f'(z)|$, which is subharmonic on C_j by (b). This gives, using (16.16), (16.20) and (16.21),

$$(K-1)\log s_n \leq \log \frac{1}{|f'(v_j)|} \leq c_2 s_n^{M+1-K} + 2N \log s_n + O(1),$$

a contradiction if n is large enough, since K and N were chosen in §16.4.6 with K/N large.

16.5 Statement and proof of the Bergweiler-Eremenko theorem

16.5.1 Theorem

Let f be transcendental and meromorphic of finite lower order in the plane, with an indirect transcendental singularity over $a \in \mathbb{C}$. Then for every $t > 0$, the corresponding component $C(t)$ contains infinitely many critical points z of f with $f(z) \neq a$.

In particular, if f has finite lower order and finitely many critical values then every every asymptotic value of f corresponds to a direct transcendental singularity of the inverse function f^{-1}.

Proof. Assume the contrary. Then there is some $\varepsilon > 0$ such that the only critical points of f in $C(\varepsilon)$ are zeros of $f - a$. Assume without loss of generality that $a = 0$. Since f has finite lower order, f can have only finitely many direct transcendental singularities, by the Denjoy-Carleman-Ahlfors theorem, and we assume that ε is so small that there is no w with $0 < |w| < \varepsilon$ such that f^{-1} has a direct transcendental singularity over w.

Take $z_0 \in C(\varepsilon)$, with $f(z_0) = w_0 \neq 0$, and a path $\gamma : [0,1] \to \{w : \delta \leq |w| \leq \varepsilon - \delta\}$, with δ positive but small compared to $|w_0|$, such that γ starts at w_0. Let g be that branch of f^{-1} mapping $w_0 = f(z_0)$ to z_0, and suppose that analytic continuation of g along γ is not possible. Then there exists

$S \in [0,1]$ such that as $t \to S-$, $z = g(\gamma(t))$ either tends to infinity or to a critical point z_1 of f with $\delta \leq |f(z_1)| \leq \varepsilon - \delta$. But the latter may be excluded since $g(\gamma(t)) \in C(\varepsilon)$ for $0 \leq t < S$, which implies, since $|f(z_1)| \leq \varepsilon - \delta$, that $z_1 \in C(\varepsilon)$, which is impossible by assumption. It follows that the path σ given by $z = g(\gamma(t)), 0 \leq t < S$, is a path tending to ∞, and lying in $C(\varepsilon)$, on which $f(z) \to w_1$ as $z \to \infty$, with $\delta \leq |w_1| \leq \varepsilon - \delta$. But then an unbounded subpath of σ lies in a component C' of the set $\{z : |f(z) - w_1| < \delta/2\}$, and $C' \subseteq C(\varepsilon)$. Hence f' has no zeros on C'. Further, the singularity over w_1 must be indirect, since we have excluded direct singularities with $0 < |w| < \varepsilon$, and this contradicts Proposition 16.4.1.

Since δ may be chosen arbitrarily small, we now see that g admits unrestricted analytic continuation in $0 < |w| < \varepsilon$. But, using Lemma 16.3.1, this implies that $C(\varepsilon)$ is simply connected, and contains at most one zero of f, which contradicts the definition of an indirect singularity.

16.5.2 Theorem

Let f be transcendental and meromorphic in the plane.
(a) Suppose that f' has finitely many zeros. Then

$$\liminf_{r\to\infty} \frac{T(r,f)}{r} > 0. \tag{16.22}$$

(b) Suppose that f'/f has finitely many zeros. Then

$$\liminf_{r\to\infty} \frac{T(r,f)}{r^{1/2}} > 0. \tag{16.23}$$

If, in addition, f has finitely many poles, then (16.22) holds.

Theorem 16.5.2 is Hinchliffe's refinement [46] of results from [25, 26]. The elementary examples $\tan z, \tan^2 \sqrt{z}$, as well as examples of larger order constructed in [25] using Riemann surfaces, show that both parts of the theorem are sharp. The proof here will be based on the unified approach given in [20], and in particular on the following lemma.

16.5.3 Lemma

Let f be transcendental and meromorphic in the plane such that f has infinitely many zeros but no asymptotic values in $A = \{w \in \mathbb{C} : 0 < |w| < \infty\}$. Then f has infinitely many critical points z with $f(z) \in A$.

Proof. Let $a \in \mathbb{C}$ with $f(a) = 0$ and let $m \in \mathbb{N}$ be the order of the zero of f at a. Let $g(z) = f(z)^{1/m}$ near a and let h be the branch of g^{-1} mapping 0 to a. Let r be the supremum of positive t such that h admits unrestricted analytic continuation in $|w| < t$. Then h extends to be analytic on $D(0,r)$ and hence r must be finite, since otherwise h is a univalent entire function and so a linear function, from which it follows that f is a rational function, which is a contradiction.

A compactness argument then gives b with $|b| = r$ such that h cannot be analytically continued along the closed line segment $[0,b]$. As $t \to 1-$ with $t \in (0,1)$ the preimage $z = h(tb)$ cannot tend to infinity, because otherwise we obtain a path tending to infinity on which $g(z)$ tends to b and $f(z)$ tends to $b^m \in A$, a contradiction. Hence there exist a sequence $t_m \in (0,1)$ with $t_m \to 1$ such that $z_m = h(t_m b)$ tends to $z^* \in \mathbb{C}$, and $g(z^*) = b, f(z^*) = b^m \in A$. Thus z^* must be a critical point of f, because otherwise h could be continued along $[0,b]$ all the way to b.

Since f has infinitely many zeros and since a critical point of f can be associated as above to at most finitely many zeros of f, it follows that f has infinitely many such critical points z^* with $f(z^*) \in A$.

16.5.4 Proof of Theorem 16.5.2

Let f be transcendental and meromorphic in the plane such that that f'/f has finitely many zeros (which is obviously the case if f' has finitely many zeros) but f does not satisfy (16.22). Then f has finitely many critical values and by Theorem 16.5.1 every asymptotic value of f corresponds to a direct transcendental singularity of the inverse function f^{-1}. By the Denjoy-Carleman-Ahlfors theorem, f has at most one asymptotic value.

Assume first that f' has finitely many zeros. Choose $a \in \mathbb{C}$ such that $f - a$ has infinitely many zeros. Applying Lemma 16.5.3 shows that f has a finite asymptotic value $b \neq a$. Thus ∞ is not an asymptotic value of f, and so by Iversen's theorem f must have infinitely many poles. Hence the function

$$g(z) = \frac{1}{f(z) - b}$$

has infinitely many zeros and asymptotic value ∞. By Lemma 16.5.3 the function g also has a finite non-zero asymptotic value, contradicting the fact that f has at most one asymptotic value. This proves part (a).

To prove part (b) we may assume without loss of generality that f has infinitely many zeros, since otherwise the result follows from part (a). Hence Lemma 16.5.3 shows that f has a finite asymptotic value $b \neq 0$, and again ∞ is not an asymptotic value of f, and f has infinitely many poles, which proves the last assertion of part (b).

Now choose $\delta > 0$ such that f has no critical or asymptotic values in $0 < |w - b| < 2\delta$. Then there exists a component C of the set $\{z \in \mathbb{C} : |f(z) - b| < \delta\}$ containing a path tending to infinity on which $f(z)$ tends to b, and by Lemma 16.3.1 the component C is simply connected. By the choice of δ the boundary of C is a union of simple curves each tending to infinity in both directions, and so $1/(f(z) - b)$ is bounded on a path tending to infinity. Thus f satisfies (16.23) by the Denjoy-Carleman-Ahlfors theorem. This completes the proof of part (b).

Bibliography

[1] W. Abikoff, The uniformization theorem, Amer. Math. Monthly 88 (1981), 574-592.

[2] L.V. Ahlfors, Complex Analysis, McGraw-Hill, New York, 1953.

[3] T.M. Apostol, Mathematical Analysis, Second Edition, Addison-Wesley, Reading Mass., 1974.

[4] W. Balser, Formal power series and linear systems of meromorphic ordinary differential equations, Universitext. Springer-Verlag, New York, 2000.

[5] S. Bank, On analytic and meromorphic functions satisfying nth order algebraic differential equations in regions of the plane, J. reine angew. Math. 253 (1972), 87-97.

[6] S. Bank, G. Frank and I. Laine, Über die Nullstellen von Lösungen linearer Differentialgleichungen, Math. Zeit. 183 (1983), 355-364.

[7] S. Bank and R. Kaufman, On meromorphic solutions of first-order differential equations, Comment. Math. Helv. 51 (1976), 289-299.

[8] S. Bank and I. Laine, On the oscillation theory of $f'' + Af = 0$ where A is entire, Trans. Amer. Math. Soc. 273 (1982), 351-363.

[9] S. Bank and I. Laine, Representations of solutions of periodic second order linear differential equations, J. reine angew. Math. 344 (1983), 1-21.

[10] S. Bank and I. Laine, On the zeros of meromorphic solutions of second-order linear differential equations, Comment. Math. Helv. 58 (1983), 656-677.

[11] S. Bank, I. Laine and J.K. Langley, On the frequency of zeros of solutions of second order linear differential equations, Results Math. 10 (1986), 8-24.

[12] S. Bank and J.K. Langley, Oscillation theory for higher order linear differential equations with entire coefficients, Complex Variables 16 (1991), 163-175.

[13] P.D. Barry, The minimum modulus of small integral and subharmonic functions, Proc. London Math. Soc. (3) 12 (1962), 445-495.

[14] P.D. Barry, On a theorem of Besicovitch, Quart. J. of Math. Oxford (2) 14 (1963), 293-302.

[15] P.D. Barry, On a theorem of Kjellberg, Quart. J. Math. Oxford (2) 15 (1964), 179-191.

[16] P.D. Barry, Some theorems related to the $cos\pi\rho$ theorem, Proc. London Math. Soc. (3) 21 (1970), 334-360.

[17] A.F. Beardon, A primer on Riemann surfaces, LMS Lecture Notes 78, Combridge University Press, 1984.

[18] A. F. Beardon, Iteration of rational functions, Springer, New York, Berlin, Heidelberg, 1991.

[19] W. Bergweiler, Iteration of meromorphic functions, Bull. Amer. Math. Soc. 29 (1993), 151-188.

[20] W. Bergweiler and A. Eremenko, On the singularities of the inverse to a meromorphic function of finite order, Rev. Mat. Iberoamericana 11 (1995), 355-373.

[21] L. Bieberbach, Theorie der gewöhnlichen Differentialgleichungen, 2. Auflage, Springer, Berlin, 1965.

[22] L. Brickman and E.S. Thomas, Conformal equivalence of analytic flows, *J. Differential Equations* 25 (1977), 310-324.

[23] K.A. Broughan, The structure of sectors of zeros of entire flows, *Topology Proceedings* 27 (2003), 379-394.

[24] A. Browder, Topology in the complex plane, Amer. Math. Monthly 107 (2000), 393-401.

[25] J. Clunie, A. Eremenko and J. Rossi, On equilibrium points of logarithmic and Newtonian potentials, J. London Math. Soc. (2) 47 (1993), 309-320.

[26] A. Eremenko, J.K. Langley and J. Rossi, On the zeros of meromorphic functions of the form $\sum_{k=1}^{\infty} \frac{a_k}{z-z_k}$, J. d'Analyse Math. 62 (1994), 271-286.

[27] A. Garijo, A. Gasull and X. Jarque, Local and global phase portrait of equation $\dot{z} = f(z)$, *Discrete Contin. Dyn. Syst.* 17 (2007), no. 2, 309-329.

[28] A.A. Gol'dberg and I.V. Ostrowski, Distribution of values of meromorphic functions, Nauka, Moscow 1970.

[29] G. Gundersen, Estimates for the logarithmic derivative of a meromorphic function, plus similar estimates, J. London Math. Soc. (2) 37 (1988), 88-104.

[30] O. Hajek, Notes on meromorphic dynamical systems. I. *Czechoslovak Math. J.* 16 (91), 1966, 14-27.

[31] O. Hajek, Notes on meromorphic dynamical systems. II. *Czechoslovak Math. J.* 16 (91), 1966, 28-35.

[32] W.K. Hayman, Picard values of meromorphic functions and their derivatives, Ann. of Math. 70 (1959), 9-42.

[33] W.K. Hayman, Meromorphic functions, Oxford at the Clarendon Press, 1964.

[34] W.K. Hayman, On the characteristic of functions meromorphic in the plane and of their integrals, Proc. London Math. Soc. (3) 14A (1965), 93-128.

[35] W.K. Hayman, Research problems in function theory, Athlone Press, London, 1967.

[36] W.K. Hayman, The local growth of power series: a survey of the Wiman-Valiron method, Canad. Math. Bull. 17 (1974), 317-358.

[37] W.K. Hayman, On Iversen's theorem for meromorphic functions with few poles, Acta Math. 141 (1978), 115-145.

[38] W.K. Hayman, Subharmonic functions Vol. 2, Academic Press, London, 1989.

[39] W.K. Hayman, Multivalent functions, 2nd edition, Cambridge Tracts in Mathematics 110, Cambridge University Press, Cambridge 1994.

[40] W.K. Hayman and P.B. Kennedy, Subharmonic functions Vol. 1, Academic Press, London, 1976.

[41] S. Hellerstein, J. Miles and J. Rossi, On the growth of solutions of $f'' + gf' + hf = 0$, Trans. Amer. Math. Soc. 324 (1991), 693-706.

[42] S. Hellerstein, J. Miles and J. Rossi, On the growth of solutions of certain linear differential equations, Ann. Acad. Sci. Fenn. Ser. A. I. Math. 17 (1992), 343-365.

[43] H. Herold, Ein Vergleichsatz für komplexe lineare Differentialgleichungen, Math. Z. 126 (1972), 91-94.

[44] E. Hille, Lectures on ordinary differential equations, Addison-Wesley, Reading, Mass., 1969.

[45] E. Hille, Ordinary differential equations in the complex domain, Wiley, New York, 1976.

[46] J.D. Hinchliffe, The Bergweiler-Eremenko theorem for finite lower order, Result. Math. 43 (2003), 121-128.

[47] E.L. Ince, Ordinary differential equations, Dover, New York, 1956.

[48] G. Jank and L. Volkmann, Einführung in die Theorie der ganzen und meromorphen Funktionen mit Anwendungen auf Differentialgleichungen, Birkhäuser, Basel, 1985.

[49] W. Jurkat, Meromorphe Differentialgleichungen, Lecture Notes in Mathematics 637, Springer, Berlin, 1978.

[50] Needham, D. J. and King, A. C., On meromorphic complex differential equations. *Dynam. Stability Systems* 9 (1994), no. 2, 99-122.

[51] I. Laine, Nevanlinna theory and complex differential equations, de Gruyter Studies in Math. 15, Walter de Gruyter, Berlin/New York 1993.

[52] J.K. Langley, Quasiconformal modifications and Bank-Laine functions, Archiv der Math. 71 (1998), 233-239.

[53] J.K. Langley, Bank-Laine functions with sparse zeros, Proc. Amer. Math. Soc. 129 (2001), 1969-1978.

[54] J.K. Langley, Linear differential equations with entire coefficients of small growth, Arch. Math. (Basel) 78 (2002), 291-296.

[55] J.K. Langley and John Rossi, Critical points of certain discrete potentials, Complex Variables 49 (2004), 621-637.

[56] J.K. Langley and D.F. Shea, On multiple points of meromorphic functions, J. London Math. Soc. (2) 57 (1998), 371-384.

[57] O. Lehto and K. Virtanen, Quasiconformal mappings in the plane, 2nd edn., Springer, Berlin, 1973.

[58] J. Miles and J. Rossi, Linear combinations of logarithmic derivatives of entire functions with applications to differential equations, Pacific J. Math. 174 (1996), 195-214.

[59] R. Nevanlinna, Eindeutige analytische Funktionen, 2. Auflage, Springer, Berlin, 1953.

[60] C. Pommerenke, Boundary behaviour of conformal maps, Grundlehren der Mathematischen Wissenschaften 299, Springer, Berlin 1992.

[61] T. Ransford, Potential theory in the complex plane, L.M.S. Student Texts 28, Cambridge, 1995.

[62] J. Rossi, Second order differential equations with transcendental coefficients, Proc. Amer. Math. Soc. 97 (1986), 61-66.

[63] H.L. Royden, Real analysis, 2nd Edn., MacMillan, 1968.

[64] W. Rudin, Real and complex analysis, McGraw-Hill, 1966.

[65] W. Schwick, Normality criteria for families of meromorphic functions, J. Analyse Math. 52 (1989), 241-289.

[66] L.C. Shen, Solution to a problem of S. Bank regarding the exponent of convergence of the solutions of a differential equation $f'' + Af = 0$, Kexue Tongbao 30 (1985), 1581-1585.

[67] L.C. Shen, Construction of a differential equation $y'' + Ay = 0$ with solutions having prescribed zeros, Proc. Amer. Math. Soc. 95 (1985), 544-546.

[68] N. Steinmetz, Eine Verallgemeinerung des zweiten Nevanlinnaschen Hauptsatzes, J. für die reine u. angewandte Mathematik 368 (1986), 134-141.

[69] E.C. Titchmarsh, The theory of functions, 2nd edn., Oxford University Press, 1939.

[70] M. Tsuji, Potential theory in modern function theory, Maruzen, Tokyo, 1959.

[71] G. Valiron, Lectures on the general theory of integral functions, Edouard Privat, Toulouse, 1923.

[72] W. Wasow, Asymptotic expansions for ordinary differential equations, Dover, New York, 1987.

[73] A. Weitsman, A theorem on Nevanlinna deficiencies, Acta Math. 128 (1972), 41-52.

[74] H. Wittich, Neuere Untersuchungen über eindeutige analytische Funktionen, Ergebnisse der Mathematik und ihrer Grenzgebiete 8 (Springer, Berlin, Göttingen, Heidelberg, 1955).

[75] L. Zalcman, A heuristic principle in complex function theory, Amer. Math. Monthly 82 (1975), 813-817.

[76] L. Zalcman, Normal families: new perspectives, Bull. Amer. Math. Soc. (N.S.) 35 (1998), 215-230.

School of Mathematical Sciences, University of Nottingham, NG7 2RD.