



Leif Mejlbro

# Stability, Riemann Surfaces, Conformal Mappings

Complex Functions Theory a-3

---

Leif Mejlbro

---

Stability, Riemann Surfaces, Conformal Mappings  
Complex Functions Theory a-3

---

---

Stability, Riemann Surfaces, Conformal Mappings – Complex Functions Theory a-3  
© 2014 Leif Mejlbro & [bookboon.com](http://bookboon.com)  
ISBN 978-87-7681-702-2

# Contents

	<b>Introduction</b>	<b>7</b>
<b>1</b>	<b>The argument principle, and criteria of stability</b>	<b>9</b>
1.1	The argument principle	9
1.2	Criteria of Stability	25
<b>2</b>	<b>Many-valued functions and Riemann surfaces</b>	<b>38</b>
2.1	Inverse mapping	38
2.2	The two-valued function	42
2.3	The $n$ -valued function	45
2.4	The infinite-valued function $\log z$	46
2.5	The many-valued functions $az$ and $z^a$	49
2.6	The many-valued function $\arcsin z$	51
2.7	The many-valued arcus functions and area functions	58
2.7.1	The infinite-valued arcus function $\arcsin z$	58
2.7.2	The infinite-valued arcus function $\arccos z$	58
2.7.3	The infinite-valued arcus function $\arctan z$	58
2.7.4	The infinite-valued arcus function $\operatorname{arccot} z$	59
2.7.5	The infinite-valued area function $\operatorname{arsinh} z$	59
2.7.6	The infinite-valued area function $\operatorname{arcosh} z$	60

---

2.7.7	The infinite-valued area function $\operatorname{artanh}z$	62
2.7.8	The infinite-valued area function $\operatorname{arcoth}z$	62
2.7.9	Some further remarks on the arcus and area functions	63
2.8	Example	63
2.9	Simple examples of potential flows in Physics	67
2.9.1	Sources and sinks	67
2.9.2	Vortices	69
2.9.3	Vortex sources/sinks	69
2.9.4	Dipoles	70
<b>3</b>	<b>Conformal mappings and the Dirichlet problem</b>	<b>72</b>
3.1	Conformal mappings	72
3.2	The Möbius transformations	74
3.3	Some special Möbius transformations	83
3.3.1	Unit disc onto unit disc	83
3.3.2	Unit disc onto complementary set of the unit disc	83
3.3.3	Upper half plane onto upper half plane	83
3.3.4	Upper half plane onto unit disc	83
3.3.5	Unit disc onto upper half plane	83
3.4	The transformation $w = z^2$	84
3.5	The exponential transformation $w = \exp z$	87
3.6	The sine transformation $w = \sin z$	89
3.7	Zhukovski's function	91

3.8	The Schwarz-Christoffel transformation	95
3.9	The Dirichlet problem	104
3.10	Some special Dirichlet problems for the upper half plane	107
	<b>Index</b>	<b>114</b>

## Introduction

We have in *Ventus: Complex Functions Theory a-1* characterized the analytic functions by their complex differentiability and by *Cauchy-Riemann's equation*. We obtained a lot of important results by arguing on line integrals in  $\mathbb{C}$ . In this way we proved the *Cauchy's Integral Theorem* and *Cauchy's Integral Formula*.

In *Ventus: Complex Functions Theory a-2* we followed an alternative approach by proving that *locally* every analytic function is described by its *Taylor series*. Historically this was the original definition of an analytic function, introduced by *Lagrange* as early as in 1797. The advantage of this approach is that it is easy to calculate on series. The disadvantage is that this approach is not global.

By combining the two aspects of analytic functions it is possible in the following to use *Cauchy-Riemann's equations*, when they are most convenient, and series when these give a better description, so we can benefit from that we have two equivalent, though different theories of the analytic functions.

We collect in the present volume *Ventus: Complex Functions Theory a-3* some applications which are not natural to include in the former two volumes.

We prove the *argument principle* in Chapter 1 and show some powerful and unexpected applications. Then we briefly sketch some *criteria of stability* of dynamic systems. These are important in some engineering sciences.

In Chapter 2 we introduce the so-called *many-valued functions*, and we visualize them by introducing their corresponding *Riemann surfaces*.

Finally, in Chapter 3 we turn to more geometric applications of *conformal mappings*, which sometimes can be used to solve some *Dirichlet problems* in an elegant way.

*Complex Functions Theory* is here described in an **a** series and a **c** series. The **c** series gives a lot of supplementary and more elaborated examples to the theory given in the **a** series, although there are also some simpler examples in the **a** series. When reading a book in the **a** series the reader is therefore recommended also to read the corresponding book in the **c** series. The present **a** series is divided into three successive books, which will briefly be described below.

**a-1** The book *Elementary Analytic Functions* is defining the battlefield. It introduces the analytic functions using the *Cauchy-Riemann equations*. Furthermore, the powerful results of the *Cauchy Integral Theorem* and the *Cauchy Integral Formula* are proved, and the most elementary analytic functions are defined and discussed as our building stones. The important applications of Cauchy's two results mentioned above are postponed to **a-2**.

**a-2** The book *Power Series* is dealing with the correspondence between an analytic function and its complex power series. We make a digression into the theory of *Harmonic Functions*, before we continue with the *Laurent series* and the *Residue Calculus*. A handful of simple rules for computing the residues is given before we turn to the powerful applications of the residue calculus in computing certain types of *trigonometric integrals*, *improper integrals* and the sum of some not so simple series.

**a-3** The book *Stability, Riemann surfaces, and Conformal maps* starts with pointing out the connection between analytic functions and *Geometry*. We prove some classical criteria for stability in *Cybernetics*. Then we discuss the inverse of an analytic function and the consequence of extending

this to the so-called *multi-valued functions*. Finally, we give a short review of the *conformal maps* and their importance for solving a *Dirichlet problem*.

The author is well aware of that the topics above only cover the most elementary parts of *Complex Functions Theory*. The aim with this series has been hopefully to give the reader some knowledge of the mathematical technique used in the most common technical applications.

Leif Mejlbro  
11th November 2014



# 1 The argument principle, and criteria of stability

## 1.1 The argument principle

Roughly speaking, the argument principle along a *simple* closed curve  $C$  of an analytic function  $f$  having only poles and no essential singularity relates the *winding number* of the image curve  $f(C)$  with respect to  $w_0 = 0$  with the difference between the number of zeros and number of poles of  $f$  inside  $C$ . This is an unexpected and strange property of an analytic function (without essential singularities inside  $C$ ), so we may expect some very powerful applications of it, and it is indeed so.

Let  $f : M \rightarrow \mathbb{C}$  be a function on any set  $M$  with complex values. By an *argument function*  $\arg f$  we (slightly incorrectly) denote any function on  $M$ , which for  $t \in M$  selects precisely one of the possible values of  $\arg f(t)$ , a many-valued function already defined in *Ventus: Complex Functions Theory a-1*. Notice that we do not at all assume that  $f$  is analytic.

**Theorem 1.1.1** *Assume that  $f : [a, b] \rightarrow \mathbb{C} \setminus \{0\}$  is continuous and never 0. Then  $f$  has a continuous argument function  $\arg f(t)$ ,  $t \in [a, b]$ .*

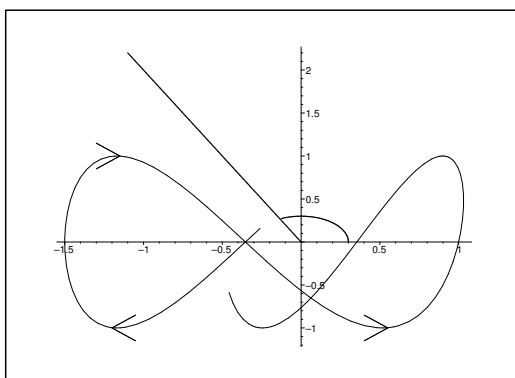


Figure 1: Proof of Theorem 1.1.1, part a).

PROOF. a) If there exists an  $\alpha$ , such that

$$f(t) \in \mathbb{C} \setminus \{z = r e^{i\alpha} \mid r \geq 0\} \quad \text{for all } t \in [a, b],$$

cf. Figure 1, then we fix an argument function by

$$\alpha < \arg f(t) < \alpha + 2\pi, \quad t \in [a, b].$$

This argument function is continuous. In fact, define

$$\text{Log}_\alpha : \mathbb{C} \setminus \{z = r e^{i\alpha} \mid r \geq 0\} \rightarrow \mathbb{C}$$

by

$$\text{Log}_\alpha z := \ln |z| + i \text{Arg}_\alpha z, \quad \text{Arg}_\alpha z \in ]\alpha, \alpha + 2\pi[$$

where  $\text{Arg}_\alpha z$  denotes the uniquely determined argument of  $z$  in the interval  $]\alpha, \alpha + 2\pi[$ . Then  $\text{Log}_\alpha$  is analytic in the slit complex plane, so it is in particular continuous. This implies that the composition  $\text{Log}_\alpha \circ f$  is also continuous, and so is its imaginary part,

$$\Im \text{Log}_\alpha \circ f(t) = \text{Arg}_\alpha f(t) = \arg f(t), \quad t \in [a, b].$$

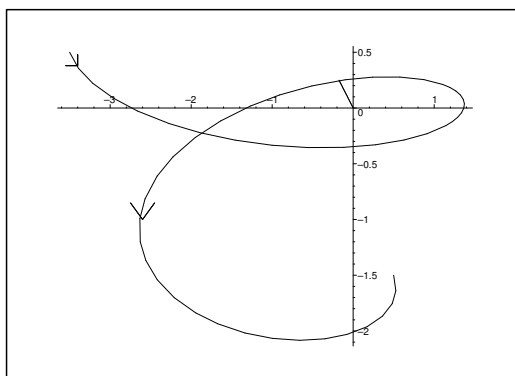


Figure 2: Proof of Theorem 1.1.1, part b).

b) Then assume that no such  $\alpha$  exists. Put

$$m := \inf\{|f(t)| \mid t \in [a, b]\}.$$

Since  $f$  is continuous, and  $[a, b]$  is compact, the image  $f([a, b])$ , where  $0 \notin f([a, b])$ , is also compact (by one of the main theorems of continuous functions), so

$$m := \inf\{|f(t)| \mid t \in [a, b]\} = \min\{|f(t)| \mid t \in [a, b]\} > 0.$$

Furthermore, by another main theorem of continuous functions,  $f$  is even *uniformly* continuous on  $[a, b]$ . Therefore, we can choose  $\delta > 0$  corresponding to  $m > 0$  (and independent of  $s$  and  $t$ ), such that

$$|f(s) - f(t)| < m, \quad \text{if } s, t \in [a, b] \text{ and } |s - t| < \delta.$$

Choose finitely many points  $a = t_0 < t_1 < \dots < t_n = b$ , such that  $|t_j - t_{j-1}| < \delta$  for all  $j = 1, \dots, n$ . For every subinterval  $[t_{j-1}, t_j]$  there exists a constant  $\alpha_j$ , such that

$$f(t) \in \mathbb{C} \setminus \{z = r \cdot \exp(i\alpha_j) \mid r \geq 0\} \quad \text{for all } t \in [t_{j-1}, t_j].$$

We may e.g. choose  $\alpha_j = \arg f(t_j) + \pi$ .

It follows from a) that there exists a continuous argument function  $\arg f$  on  $[t_{j-1}, t_j]$  for every  $j = 1, \dots, n$ . Furthermore, we can choose these argument functions such that  $\arg f(t_j)$  has the same value in both the adjacent intervals  $[t_{j-1}, t_j]$  and  $[t_j, t_{j+1}]$ , and the composed argument function is continuous.  $\square$

Given one continuous argument function  $\Theta = \arg f$ , we get other continuous argument functions by adding  $2p\pi$  for  $p \in \mathbb{Z}$ , i.e.  $\Theta(t) + 2p\pi$  is also a continuous argument function for every fixed  $p \in \mathbb{Z}$ .

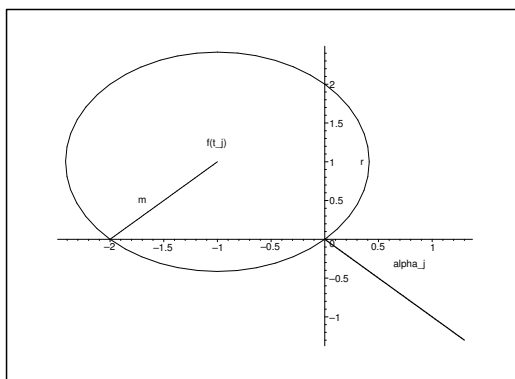


Figure 3: Proof of Theorem 1.1.1, part b).

If conversely  $\Theta(t)$  and  $\Theta_1(t)$  are continuous argument functions, then the difference  $\Theta_1(t) - \Theta(t)$  is again a continuous function on  $[a, b]$ , which for every  $t \in [a, b]$  has the value  $2p\pi$  for some function  $p = p(t) \in \mathbb{Z}$ . Since  $[a, b]$  is connected, and  $p : [a, b] \rightarrow \mathbb{Z}$  is continuous, it follows from the definition of connectedness that  $p(t) = p$  must be a constant. This implies that the difference

$$(1) \arg f(b) - \arg f(a)$$

is a constant independent of the chosen argument function  $\arg$  in (1). We call this constant (1) the *argument variation* of the function  $f$  over the interval  $[a, b]$ .

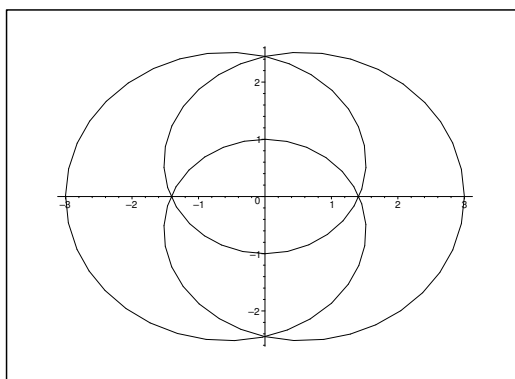


Figure 4: The argument variation of a closed curve.

If  $f(a) = f(b)$ , then the continuous curve described by the function  $z = f(t)$ ,  $t \in [a, b]$ , is a closed curve, which does not pass through 0. This means that the argument variation must be

$$(2) \arg f(b) - \arg f(a) = 2n\pi \quad \text{for some } n \in \mathbb{Z}.$$

The uniquely determined number  $n \in \mathbb{Z}$  is called the *winding number* of the function – or the curve – with respect to 0.

The *winding number* indicates the number of times the curve winds around 0 counted positive in the positive orientation of the plane, and negative in the negative orientation of the plane. Notice that positive and negative winding numbers cancel each other. This simple geometric interpretation often makes it easy to find the winding number of a given curve.

The importance of the winding number will later be made clear by the *argument principle*. We first prove

**Theorem 1.1.2** *Let  $f : [a, b] \rightarrow \mathbb{C} \setminus \{0\}$  and  $g : [a, b] \rightarrow \mathbb{C}$  be two continuous complex functions, for which  $f(a) = f(b)$  and  $g(a) = g(b)$ . If for all  $t \in [a, b]$ ,*

$$(3) \quad |g(t)| < |f(t)|,$$

*then  $f$  and  $f + g$  have the same winding number  $n$  with respect to 0.*

**Remark 1.1.1** A man is walking his dog around a monument, represented by movements in the complex plane with the monument situated at 0. The dog has the distance  $|g(t)|$  from the man, and the man has the distance  $|f(t)|$  from the monument. If the lace is so short that always  $|g(t)| < |f(t)|$  and the man and the dog start and end at the same points, then they have both walked the same number of times around the monument.

Another example is of course the Earth and the Moon. From the time when this planet/satellite system was created, the Earth and the Moon have moved exactly the same number of times around the sun.  $\diamond$

PROOF. Since

$$|f(t) + g(t)| \geq |f(t)| - |g(t)| > 0 \quad \text{for all } t \in [a, b],$$

it follows that the two continuous functions  $f$  and  $f + g$  on  $[a, b]$  are always  $\neq 0$ , and

$$f(a) = f(b), \quad \text{and} \quad (f + g)(a) = (f + g)(b).$$

The rewrite

$$f(t) + g(t) = f(t) \cdot \left\{ 1 + \frac{g(t)}{f(t)} \right\}, \quad t \in [a, b].$$

It follows from  $\left| \frac{g(t)}{f(t)} \right| < 1$  that

$$\Re \left\{ 1 + \frac{g(t)}{f(t)} \right\} \geq 1 - \left| \frac{g(t)}{f(t)} \right| \geq 0,$$

so  $1 + \frac{g(t)}{f(t)}$  lies in the right hand half plane for every  $t \in [a, b]$ . In particular, the *principal argument*  $\text{Arg} \left( 1 + \frac{g(t)}{f(t)} \right)$  is continuous for  $t \in [a, b]$ .

Choosing any continuous argument function  $\arg^* f$  for  $f$  it follows that

$$\arg^* f(t) + \text{Arg} \left\{ 1 + \frac{g(t)}{f(t)} \right\} := \arg(f + g)(t)$$

is a continuous argument function for  $f + g$ . Since

$$\text{Arg} \left\{ 1 + \frac{g(a)}{f(a)} \right\} = \text{Arg} \left\{ 1 + \frac{g(b)}{f(b)} \right\},$$

we finally conclude that

$$\arg(f + g)(b) - \arg(f + g)(a) = \arg^* f(b) - \arg^* f(a) = 2\pi n. \quad \square.$$

The importance of Theorem 1.1.2 lies in the fact that we can move closed curves slightly without changing its *winding number* with respect to 0.

We first extend the definition of the *winding number* in the following obvious way. Let  $\Omega \subseteq \mathbb{C}$  be an open domain, and let  $f : \Omega \rightarrow \mathbb{C} \setminus \{0\}$  be a continuous (not necessarily analytic) function. If  $C$  is a *simple*, closed curve in  $\Omega$ , i.e. without any double points, then the image  $f(C)$  is a closed curve, though not necessarily simple (it may have double points) in  $\mathbb{C} \setminus \{0\}$ , and the geometric definition of the *winding number* of the curve  $f(C)$  with respect to 0 makes sense.

We shall in the following restrict ourselves to analytic functions  $f : \Omega \rightarrow \mathbb{C}$ , where  $\Omega \subseteq \mathbb{C}$  is an open domain. We shall assume that the singularities of  $f$  are either *removable* or *poles*, and that  $f$  is not identically zero.

**Definition 1.1.1** *Let the function  $f : \Omega \rightarrow \mathbb{C}$  be given as above. We define its logarithmic derivative as the analytic function*

$$\frac{f'(z)}{f(z)} \quad \text{defined in the set } \{z \in \mathbb{C} \mid f(z) \neq 0\}.$$

The name “logarithmic derivative” is due to the fact that if  $f(z) \notin \mathbb{R}_- \cup \{0\}$  for all  $z \in \Omega$ , i.e.  $f(z)$  is never a real negative number, or 0, then  $\text{Log } f(z)$  is analytic in all of  $\Omega$ , and we have

$$(4) \quad \frac{d}{dz} \text{Log } f(z) = \frac{f'(z)}{f(z)}.$$

In general, there does not exist an analytic logarithm of  $f(z)$  defined in all of  $\Omega$ . However, since the right hand side of (4) is defined and analytic in  $\{z \in \Omega \mid f(z) \neq 0\}$ , and since it often occurs in the applications, we have coined this name in Definition 1.1.1.

One very important result in this section is the following

**Theorem 1.1.3** *The argument principle. Let  $f : \Omega \rightarrow \mathbb{C}$  be analytic in an open domain  $\Omega$ , where  $f$  is not identically zero, and where all isolated singularities of  $f$  in  $\overline{\Omega}$  are poles. Let  $C$  be a simple, closed curve in  $\Omega$ , not passing through any zero of  $f(z)$ , and denote by  $\omega \subset \mathbb{C}$  the bounded domain inside  $C$ .*

*Let  $Z$  denote the number of zeros (counted by multiplicity) and  $P$  the number of poles (also counted by multiplicity) in  $\omega$ . Then the difference  $Z - P$  is equal to the winding number of the closed curve  $f(C)$  with respect to 0 in the  $w$ -plane.*

*We have more precisely,*

$$(5) \quad \frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz = Z - P = \text{the winding number of } f(C) \text{ with respect to } w_0 = 0.$$

**Remark 1.1.2** Notice that  $\omega$  in general is not a subset of  $\Omega$ . We see that  $\omega \setminus \Omega$  is precisely the set of all poles of  $f$  inside the simple, closed curve  $C$ .  $\diamond$

**PROOF.** We first prove that the number of poles is finite,  $P < +\infty$ . Let us assume that instead  $P = +\infty$ . Then  $f(z)$  has infinitely many poles  $\{z_n \in \overline{\omega} \mid n \in \mathbb{N}\}$  in the compact set  $\overline{\omega}$ , and the reciprocal function  $g(z) := \frac{1}{f(z)}$  would have the infinitely many zeros  $\{z_n \in \overline{\omega} \mid n \in \mathbb{N}\}$  in the compact

set  $\bar{\omega}$ . Hence, there exists at least one cluster point  $z_0 \in \bar{\omega}$  and a subsequence  $\{z_{n_j}\} \rightarrow z_0$ , and we conclude from  $0 = g(z_{n_j}) \rightarrow g(z_0)$  that also  $g(z_0) = 0$ , so  $z_0$  is a non-isolated zero of  $g$ . This implies, cf. *Ventus: Complex Functions Theory a-1*, that  $g(z)$  is identical 0, so  $f(z)$  is identical  $\infty$ , which is not possible. Hence, the number of poles in  $\omega$  must be finite,  $P \in \mathbb{N}_0$ .

Then we prove that  $\frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz$  is equal to the winding number of  $f(C)$  with respect to  $w_0 = 0$ .

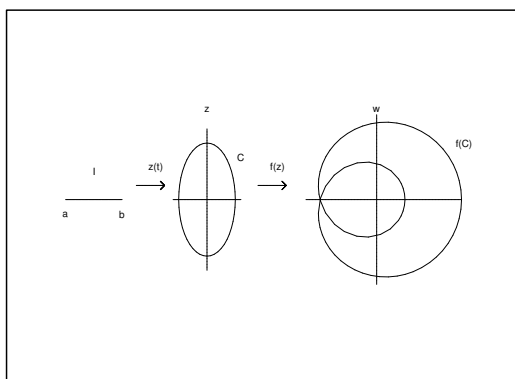


Figure 5: The winding number in the proof of the argument principle.

Assume that the curve  $C$  is given by its parametric description  $z(t)$ ,  $t \in [a, b]$ . We define by composition a continuous complex function  $g : [a, b] \rightarrow \mathbb{C} \setminus \{0\}$  by  $g(t) := f(z(t)) \neq 0$  for  $t \in [a, b]$ , and  $g(t)$  has a continuous argument function  $\arg g$ .

Using the same construction as in b) of the proof of Theorem 1.1.1 it follows that each subinterval  $[t_j, t_{j+1}]$  corresponds to a curve segment  $C_j$  of  $C$ , and

$$\begin{aligned} \frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz &= \sum_{j=0}^{n-1} \frac{1}{2\pi i} \int_{C_j} \frac{f'(z)}{f(z)} dz = \sum_{j=0}^{n-1} \frac{1}{2\pi i} \int_{t_j}^{t_{j+1}} \frac{g'(t)}{g(t)} dt \\ &= \sum_{j=0}^{n-1} \frac{1}{2\pi i} [\ln |g(t)| + i \arg g(t)]_{t_j}^{t_{j+1}} \\ &= \frac{1}{2\pi i} \sum_{j=0}^{n-1} \{\ln |g(t_{j+1})| - \ln |g(t_j)|\} + \frac{1}{2\pi} \sum_{j=0}^{n-1} \{\arg g(t_{j+1}) - \arg g(t_j)\} \\ &= \frac{1}{2\pi i} \{\ln |g(t_n)| - \ln |g(t_0)|\} + \frac{1}{2\pi} \{\arg g(t_n) - \arg g(t_0)\} \\ &= \frac{1}{2\pi i} \{\ln |g(b)| - \ln |g(a)|\} + \frac{1}{2\pi} \{\arg g(b) - \arg g(a)\} \\ &= 0 + \text{winding number of } g([a, b]) = f(C) \text{ with respect to } w_0 = 0. \end{aligned}$$

Finally, we prove that  $\frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz = Z - P$ .

The singularities of  $\frac{f'(z)}{f(z)}$  in  $\omega$  are precisely all zeroes and all poles of  $f$  in  $\omega$ . Hence by the *Residuum Theorem*, cf. *Ventus: Complex Functions Theory a-2*,

$$(6) \quad \frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz = \sum_{z_0 \text{ zero}} \operatorname{res}\left(\frac{f'}{f}; z_0\right) + \sum_{z_0 \text{ pole}} \operatorname{res}\left(\frac{f'}{f}; z_0\right).$$

Assume that  $z_0$  is a zero of multiplicity  $n$  for  $f$ . Then we have in a neighbourhood of  $z_0$ ,

$$f(z) = (z - z_0)^n \cdot h(z), \quad \text{where } h(z_0) \neq 0.$$

If instead  $z_0$  is a pole of multiplicity  $p$  for  $f$ , then similarly in a neighbourhood of  $z_0$ ,

$$f(z) = (z - z_0)^{-n} \cdot h(z), \quad \text{where } h(z_0) \neq 0.$$

The structure is the same in the two cases, so we may assume that we in some neighbourhood of  $z_0$  have

$$(7) \quad f(z) = (z - z_0)^m \cdot h(z), \quad \text{where } h(z) \neq 0 \text{ everywhere in the neighbourhood and } m \in \mathbb{Z}.$$



If  $m > 0$ , then  $z_0$  is a zero of multiplicity  $m$ , and if  $m < 0$ , then  $z_0$  is a pole of multiplicity  $-m = |m|$ .

In this neighbourhood,

$$\frac{f'(z)}{f(z)} = \frac{m(z-z_0)^{m-1} \cdot h(z) + (z-z_0)^m \cdot h'(z)}{(z-z_0)^m \cdot h(z)} = \frac{m}{z-z_0} + \frac{h'(z)}{h(z)}.$$

Since  $\frac{h'(z)}{h(z)}$  is analytic in this neighbourhood, we conclude that

$$\operatorname{res}\left(\frac{f'}{f}; z_0\right) = m.$$

Inserting this result into (6) we finally get

$$\frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz = Z + (-P) = Z - P,$$

and the argument principle is proved.  $\square$

Combining the argument principle and Theorem 1.1.2 we easily get the following importing result.

**Theorem 1.1.4** Rouché's theorem. *Let  $f, g : \Omega \rightarrow \mathbb{C}$  be analytic functions in an open domain  $\Omega \subseteq \mathbb{C}$ . Let  $C$  be a simple closed curve in  $\Omega$ , and assume that neither  $f(z)$  nor  $g(z)$  have singularities inside  $C$ . If*

$$|f(z)| > |g(z)| \quad \text{for all } z \in C,$$

*then the two functions  $f$  and  $f + g$  have the same number of zeros inside  $C$  (counted by their multiplicities), and it is equal to the winding number of either of the two functions  $f$  and  $f + g$ .*

PROOF. It follows from Theorem 1.1.2 that  $f$  and  $f + g$  have the same winding number with respect to  $w_0 = 0$ , hence

$$Z_{f+g} - P_{f+g} = Z_f - P_f.$$

According to the assumptions of the theorem,  $P_{f+g} = P_f = 0$ , thus  $Z_{f+g} = Z_f$ , and the theorem is proved.  $\square$

**Remark 1.1.3** In the proof of *Rouché's theorem* we only used that  $P_{f+g} = P_f$  and not that their common value is 0. We can therefore relax this theorem to the weaker requirement that the "larger" function  $f(z)$  is allowed to have poles, while only the perturbation  $g(z)$  (the "smaller" function) must not have poles inside  $C$ . Neither of them must have essential singularities. We notice that if  $P_f > 0$ , then the number of zeros  $Z_{f+g} = Z_f = P_f + \text{the winding number}$ , is clearly bigger than the winding number.  $\diamond$

Notice that when we add the perturbation  $g(z)$  (without poles) the zeros of  $f + g$  may lie far away from the zeros of  $f$ , even though their numbers are equal. We shall demonstrate this by the following example.

**Example 1.1.1** Choose a constant  $a \in \mathbb{C}$ , such that  $0 < |a| < 1$ , i.e.  $a$  lies in the deleted disc  $D(0; 1)$ . Put  $f(z) = z^n$  and  $g(z) = -a$ . Let  $C$  be the unit circle  $|z| = 1$ . Then we have the estimate

$$|f(z)| = |z^n| = 1 > |a| = |g(z)| \quad \text{for } |z| = 1,$$

and we conclude from *Rouché's theorem* that the two functions  $f(z) = z^n$  and  $f(z) + g(z) = z^n - a$  must have the same number of zeros inside  $C$ , because we trivially have no poles involved.

The function  $f(z) = z^n$  has the  $n$ -tuple zero  $z_0 = 0$ , while the perturbed function  $f(z) + g(z) = z^n - a$  has  $n$  simple roots, all lying on the circle  $|z| = \sqrt[n]{|a|}$  of radius  $\sqrt[n]{|a|}$  and centre 0. Here we use a solution formula of the binomial equation from *Ventus: Complex Functions Theory a-1*. The addition of even a very small constant  $a \neq 0$  will cause the (unstable) zero at  $z_0 = 0$  of multiplicity  $n$  to bifurcate into  $n$  simple roots “relatively far away from  $z_0 = 0$ .”  $\diamond$

It is possible to find the exact value of more or less “impossible” line integrals of the form  $\oint_C \frac{h'(z)}{h(z)} dz$  by applying *Rouché's theorem* and the *argument principle*. The trick is to write  $h(z)$  as a sum,  $h(z) = f(z) + g(z)$ , where the dominating term  $f(z)$  should be chosen fairly simple. We shall demonstrate the technique by the following deliberately laboured example.

**Example 1.1.2** We shall find the exact value of the following line integral

$$\frac{1}{2\pi i} \oint_{|z|=2} \frac{10z^9 + 25 \sinh z + \frac{1}{2} \exp\left(\frac{z}{2}\right) \cdot \exp\left(\exp\left(\frac{z}{2}\right)\right)}{z^{10} + 25 \cosh z + \exp\left(\exp\left(\frac{z}{2}\right)\right)} dz.$$

The usual method of inserting a parametric description and hope for some divine inspiration is clearly doomed to failure. Instead we put the denominator

$$h(z) := z^{10} + 25 \cosh z + \exp\left(\exp\left(\frac{z}{2}\right)\right),$$

and then notice that the line integral has the structure

$$\frac{1}{2\pi i} \oint_{|z|=2} \frac{h'(z)}{h(z)} dz = Z_h - P_h = Z_h,$$

because  $h(z)$  does not have poles, so  $P_h = 0$ .

We estimate each of the three terms of  $h(z)$  on the circle  $|z| = 2$ .

First,

$$|z^{10}| = 2^{10} = 1024 \quad \text{for } |z| = 2.$$

Then use that  $|e^z| = e^x \leq e^2 < 8$  to get

$$|25 \cosh z| = 25 \left| \frac{1}{2} \{e^z + e^{-z}\} \right| \leq \frac{25}{2} \cdot 2 \cdot e^2 \leq 200 \quad \text{for } |z| = 2,$$

where we of course can obtain a better estimate by using a pocket calculator. However, the estimate above will later prove to be sufficient.

Finally,

$$\left| \exp \left( \exp \left( \frac{z}{2} \right) \right) \right| \leq e^e < 3^3 = 27 \quad \text{for } |z| = 2.$$

If we put  $f(z) = z^{10}$  and  $g(z) = 25 \cosh z + \exp \left( \exp \left( \frac{z}{2} \right) \right)$ , then we get for  $|z| = 2$  the following estimates

$$|g(z)| \leq |25 \cosh z| + \left| \exp \left( \exp \left( \frac{z}{2} \right) \right) \right| < 200 + 27 < 1024 = |z^{10}| = |f(z)|.$$

It follows from *Rouché's theorem* that

$$\frac{1}{2\pi i} \oint_{|z|=2} \frac{10z^9 + 25 \sinh z + \frac{1}{2} \exp \left( \frac{z}{2} \right) \cdot \exp \left( \exp \left( \frac{z}{2} \right) \right)}{z^{10} + 25 \cosh z + \exp \left( \exp \left( \frac{z}{2} \right) \right)} dz = Z_h = Z_{f+g} = Z_f = 10,$$

because  $f(z) = z^{10}$  has a zero of multiplicity 10 at  $z_0 = 0$  and no other zeros.  $\diamond$

When we choose the dominating term  $f(z)$  in the splitting  $h(z) = f(z) + g(z)$ , we often take it as the polynomial part of  $h(z)$ . This is, however, not always the right thing to do, which is shown by the following example.

**Example 1.1.3** We shall find the number of zeros of the analytic function  $z^5 + 5 \sin z$  within the rectangle of the corners  $\pm \frac{\pi}{2} \pm i$ , cf. Figure 6. It follows by inspection that  $z = 0$  is a (simple) zero. The problem is, if there are other zeros inside  $C$ .

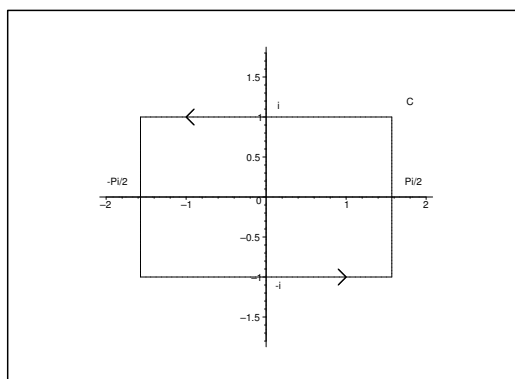


Figure 6: The curve  $C$  of Example 1.1.3.

Using a result from *Ventus: Complex Functions Theory a-1* we get

$$|\sin z|^2 = |\sin x \cdot \cosh y + i \cos x \cdot \sinh y| = \sin^2 x + \sinh^2 y,$$

from which we conclude that

$$|\sin z| \geq \left| \sin \left( \pm \frac{\pi}{2} \right) \right| = 1 \quad \text{for } z = \pm \frac{\pi}{2} + iy, \quad y \in [-1, 1],$$

(the vertical line segments of  $C$ ), and

$$|\sin z| \geq \sinh 1 > 1 \quad \text{for } z = x \pm i, \quad x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],$$

(the horizontal line segments of  $C$ ).

Then choose  $f(z) = 5 \sin z$  and  $g(z) = z^2$ , where

$$|f(z)| = |5 \sin z| \geq 5 \quad \text{for } z \in C,$$

(the assumed “larger” term is estimated *downwards*), and by a geometric inspection (choose the most distant points on  $C$  from 0; cf. Figure 6)

$$|g(z)| = |z^2| \leq \left|\frac{\pi}{2} + i\right|^2 = \frac{\pi^2}{4} + 1 < \frac{10}{4} + 1 < 5 \quad \text{for } z \in C,$$

(the assumed “smaller” term is estimated *upwards*).

Since  $|g(z)| < |f(z)|$  on  $C$ , it follows from *Rouché’s theorem* that  $Z_{f+g} = Z_f$ .

The (simple) zeros of the function  $f(z) = 5 \sin z$  are given by the set  $\{p\pi \mid p \in \mathbb{Z}\}$ . Clearly, among these only  $z_0 = 0$  lies inside  $C$ . Hence,  $z_0 = 0$  is the only zero of  $z^2 + 5 \sin z$  inside  $C$ , i.e. in the rectangle.

It is left to the reader to fill in the details of the proof of the following:  
Assume that  $|\lambda| < 4 - \left(\frac{\pi}{2}\right)^2$ . Then the transcendent equation

$$z^2 + 5 \sin z = \lambda$$

has precisely one solution in the rectangle inside  $C$ .

Sketch of proof: Just repeat the proof above with the modification that now  $g(z) = z^2 - \lambda$ .  $\diamond$

Another application of *Rouché's theorem* is an alternative proof of the *Fundamental Theorem of Algebra*.

**Theorem 1.1.5** The Fundamental Theorem of Algebra. *Every polynomial*

$$P(z) = a_0 + a_1z + \cdots + a_nz^n, \quad a_n \neq 0,$$

*of degree  $n \in \mathbb{N}$  has precisely  $n$  complex roots (counted by their multiplicities).*

PROOF. Choose  $R > 0$  so big that

$$|a_n| r^n > |a_0| + |a_1| r + \cdots + |a_{n-1}| r^{n-1} \quad \text{for every } r \geq R.$$

This follows from

$$\frac{1}{r} \left\{ |a_{n-1}| + |a_{n-2}| \cdot \frac{1}{r} + \cdots + |a_0| \cdot \frac{1}{r^{n-1}} \right\} \rightarrow 0 \quad \text{for } r \rightarrow +\infty,$$

so there exists such an  $R > 0$ .

If we put  $f(z) = a_nz^n$  and  $g(z) = a_0 + a_1z + \cdots + a_{n-1}z^{n-1}$ , then  $|f(z)| > |g(z)|$  on every circle  $C_r$  of centre 0 and radius  $r \geq R$ , and it follows from *Rouché's theorem* that  $P(z) = f(z) + g(z)$  and  $f(z) = a_nz^n$  have the same number of zeros inside every circle  $C_r$ , provided that  $r \geq R$ , i.e. in all of  $\mathbb{C}$ . Since  $f(z) = a_nz^n$  has the  $n$ -tuple zero  $z_0 = 0$  and no other zero, the theorem is proved.  $\square$

**Example 1.1.4** Given the polynomial

$$P(z) = z^3 + 2z^2 - 50z + 100.$$

It follows from the *Fundamental Theorem of Algebra* that  $P(z)$  has three roots in  $\mathbb{C}$ . We shall try roughly to find where they are situated in  $\mathbb{C}$ .

We first prove that all three roots lie inside the circle  $|z| = 9$ . Put

$$f_1(z) = z^3 \quad \text{and} \quad g_1(z) = 2z^2 - 50z + 100.$$

Then  $P(z) = f_1(z) + g_1(z)$  and

$$|f_1(z)| = 729 \quad \text{and} \quad |g_1(z)| \leq 2 \cdot 9^2 + 50 \cdot 9 + 100 = 712 \quad \text{for } |z| = 9,$$

so  $f_1(z)$  and  $P(z) = f_1(z) + g_1(z)$  have the same number of roots inside  $|z| = 9$ , namely three.

Then we prove that we have only one root lying inside the circle  $|z| = 4$ . In this case we choose  $f_2(z) = -50z$  and  $g_2(z) = z^3 + 2z^2 + 100$ , and we get for  $|z| = 4$ ,

$$|f_2(z)| = |-50z| = 200 \quad \text{and} \quad |g_2(z)| = |z^3 + 2z^2 + 100| \leq 64 + 32 + 100 = 196,$$

so  $f_2(z) = -50z$  and  $P(z) = f_2(z) + g_2(z) = z^3 + 2z^2 - 50z + 100$  have the same number of roots inside  $|z| = 4$ , namely one.

Finally we prove that there are no roots inside  $|z| = \frac{7}{4}$ . In this case we choose  $f_3(z) = 100$  and  $g_3(z) = z^3 + 2z^2 - 50z$ , so we get for  $|z| = \frac{7}{4}$  the estimate

$$|g_3(z)| = |z^3 + 2z^2 - 50z| \leq \left\{2 + \frac{7}{4}\right\} \cdot \left\{\frac{7}{4}\right\}^2 + 50 \cdot \frac{7}{4} = \frac{15}{4} \cdot \frac{49}{16} + \frac{7}{4} \cdot 50 < \left\{\frac{1}{4} + \frac{7}{4}\right\} \cdot 50 = 100 = |f_3(z)|,$$

so  $f_3(z) = 100$  (no roots) dominates, and the claim follows.

As a check we mention that the roots of  $z^3 + 2z^2 - 50z + 100$  are approximately

$$-8.889\,794\,306, \quad 2.658\,473\,477, \quad 4.231\,320\,828,$$

i.e. it has three real roots.

Then we check what happens if we change the minus sign in the polynomial to a plus sign, i.e. we consider the polynomial

$$P_1(z) = z^3 + 2z^2 + 50z + 100.$$

The coefficients of  $P_1(z)$  are *numerically* the same as the coefficients of  $P(z)$ , so we can with no change repeat the argument above. Thus,

- all three roots lie inside  $|z| = 9$ ;
- there is only one root inside  $|z| = 4$  (and hence two roots in the annulus  $4 < |z| < 9$ );
- there are no roots inside  $|z| = \frac{7}{4}$ , thus precisely one root in the annulus  $\frac{7}{4} < |z| < 4$ .

There is nothing wrong with this analysis, except for the fact that we can directly by inspection find the roots of  $P_1(z)$ , because we have the factorization

$$P_1(z) = z^3 + 2z^2 + 50z + 100 = (z + 2)(z^2 + 50),$$

so the roots are in this case

$$-2, \quad 5\sqrt{2}i, \quad -5\sqrt{2}i.$$

We notice that we get one real root and two complex conjugated roots of  $P_1(z)$ .  $\diamond$

**Example 1.1.5** We shall in this example show that we are now able to compute line integrals like

$$(8) \oint_{|z|=1} \frac{z-1}{e^z-3z} dz,$$

which does not fall into any of the standard categories considered previously.

We shall aim at an application of the *residuum formula*, so we shall first find the zeros of the denominator  $h(z) = e^z - 3z$  inside  $|z| = 1$ . A direct attack on the transcendental equation  $h(z) = e^z - 3z = 0$  does not look promising, if we want an exact solution, and not just an approximation. Let us first more modestly find the *number* of zeros inside  $|z| = 1$ . It follows by choosing  $f(z) = -3z$  and  $g(z) = e^z$  that

$$|g(z)| = |e^z| = e^x \leq e < 3 = |-3z| = |f(z)| \quad \text{for } |z| = 1,$$

so  $f(z) = -3z$  and  $f(z) + g(z) = h(z) = e^z - 3z$  have the same number of zeros inside  $|z| = 1$ , namely one. Let  $z_0$  denote the zero of  $h(z)$  in this set. Then it is easily seen that  $z_0$  is real and  $0 < z_0 < 1$ , so an approximate value can be found either by a pocket calculator or by an iteration procedure. We shall here use another method, because we want to see if we can find the *exact value* of (8) and not just an approximate one. We note so far that we have

$$\exp(z_0) = 3z_0 \quad \text{and} \quad |z_0| < 1 \quad (\text{and even } z_0 \text{ real and } 0 < z_0 < 1).$$

We shall need the following simple result from residuum calculus:

Let  $z_0$  be a zero of order  $m$ , or a pole of order  $-m$ , of an analytic function  $h(z)$  in a deleted neighbourhood of  $z_0$ , and let  $\varphi(z)$  be analytic in some neighbourhood of  $z_0$ . Then

$$(9) \quad \text{res} \left( \varphi(z) \cdot \frac{h'(z)}{h(z)}; z_0 \right) = m \cdot \varphi(z_0).$$

The proof is simple, because a *Laurent series expansion* of  $\varphi(z) \cdot \frac{h'(z)}{h(z)}$  from  $z_0$  must have the form

$$\varphi(z) \cdot \frac{h'(z)}{h(z)} = \varphi(z_0) \cdot \frac{m}{z - z_0} + \dots,$$

where the dots indicate a power series in  $z - z_0$ . Since the residuum at  $z_0$  is defined as the coefficient of  $(z - z_0)^{-1}$  in this Laurent series expansion, the claim follows.

The result (9) indicates that we should try to rewrite the integrand of (8) in the form

$$(10) \quad \frac{z - 1}{e^z - 3z} = \frac{z - 1}{e^z - 3} \cdot \frac{e^z - 3}{e^z - 3z}$$

with

$$h(z) = e^z - 3z, \quad h'(z) = e^z - 3, \quad \text{and} \quad \varphi(z) = \frac{z - 1}{e^z - 3}.$$

This is legal, if we can prove that we never divide by zero, so we require that  $h'(z) = e^z - 3 \neq 0$  for  $|z| = 1$ . This is, however, obvious, because

$$|e^z| \leq e^x \leq e < 3 \quad \text{for } |z| \leq 1.$$

Using (10) it then follows from (9) and (8) and  $\exp(z_0) = 3z_0$  for the only zero (of multiplicity  $m = 1$ )

$$\begin{aligned} \oint_{|z|=1} \frac{z - 1}{e^z - 3z} dz &= \oint_{|z|=1} \frac{z - 1}{e^z - 3} \cdot \frac{e^z - 3}{e^z - 3z} dz = \oint_{|z|=1} \frac{z - 1}{e^z - 3} \cdot \frac{h'(z)}{h(z)} dz \\ &= 2\pi i \text{res} \left( \frac{z - 1}{e^z - 3} \cdot \frac{h'(z)}{h(z)}; z_0 \right) = 2\pi i \cdot \frac{z_0 - 1}{\exp(z_0) - 3} \cdot 1 = 2\pi i \cdot \frac{z_0 - 1}{3z_0 - 3} = \frac{2\pi i}{3}. \end{aligned}$$

Alternatively, one may directly apply one of the simple rules of computing the residuum given in *Ventus: Complex Functions Theory a-2*, where we again use that  $z_0$  given by  $\exp(z_0) = 3z_0$  and  $|z_0| < 1$  is a simple pole. In this approach we get instead

$$\begin{aligned}\oint_{|z|=1} \frac{z-1}{e^z-3z} dz &= 2\pi i \lim_{z \rightarrow z_0} \frac{z-1}{\frac{d}{dz}\{e^z-3z\}} = 2\pi i \cdot \lim_{z \rightarrow z_0} \frac{z-1}{e^z-3} \\ &= 2\pi i \cdot \frac{z_0-1}{\exp(z_0)-3} = 2\pi i \cdot \frac{z_0-1}{3z_0-3} = \frac{2\pi i}{3}.\end{aligned}$$

It is remarkable that we in neither of the two solutions methods need to find the zero  $z_0$  of the denominator explicitly. We just eliminaten  $z_0$  by the equation  $\exp(z_0) = 3z_0$ .  $\diamond$



## 1.2 Criteria of Stability

When we consider the question of stability of mechanical or electric systems it is of paramount importance to decide whether all zeros of a certain polynomial connected with the system lie in the left half plane. As an example, the simplest oscillation problem will lead to a model, which is described by a *linear differential equation* of constant coefficients

$$(11) \quad \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \cdots + a_{n-1} \frac{dx}{dt} + a_n x = 0.$$

The corresponding *characteristic polynomial*, which is also called the *transition function*, and which sometimes in the technical literature is found by taking the *Laplace transform* of (11), which is not necessary in this simple case, because it is simply obtained by replacing each derivative  $\frac{d^j x}{dt^j}$  by the monomial  $\lambda^j$  in (11), is given by

$$(12) \quad P(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n.$$

In general we are of course forced to use the Laplace transform, see below.

To illustrate what we are aiming at, we here for *simplicity* – it is not necessary, but very convenient – assume that the complex roots  $\lambda_1, \dots, \lambda_n$  of (12) are mutually different,  $\lambda_i \neq \lambda_j$ , whenever  $i \neq j$ . Then the solution of (11) is given by

$$(13) \quad x = \sum_{j=1}^n C_j \exp(\lambda_j t), \quad \text{where } C_1, \dots, C_n \in \mathbb{C} \text{ are arbitrary constants.}$$

Let us consider the solution in (13) corresponding to the index  $j$ . We split  $\lambda_j$  into its real and imaginary parts,  $\lambda_j = s_j + i\sigma_j$ , where  $s_j, \sigma_j \in \mathbb{R}$ . Then the solution is the oscillation

$$\exp(\lambda_j t) = \exp(s_j t) \cdot \{\cos(\sigma_j t) + i \sin(\sigma_j t)\}$$

of the frequency  $\sigma_j$  (if  $\sigma_j \neq 0$ ). This oscillation is damped if  $s_j < 0$ , and the amplitude tends to infinity for  $t \rightarrow +\infty$  if  $s_j > 0$ , while the amplitude is 1, if  $s_j = 0$ . We conclude that if the solution (13) is stable, i.e. it is converging, then all the roots necessarily *must* lie in the open left half plane.

In more complicated cases we are forced to apply the *Laplace transform*, because (11) is then replaced by a more general linear *convolution equation*, typically of the form

$$(14) \quad y(t) = a_n \frac{d^n f}{dt^n} + \cdots + a_0 f(t) + b \int_0^t \varphi(t - \tau) f(\tau) d\tau,$$

where  $\varphi(t) = 0$  and  $f(t) = 0$  for  $t < 0$ , where  $\varphi(t)$  is a known function, while  $f(t)$  denotes the unknown function with some given initial conditions  $f(0), \dots, f^{(n-1)}(0)$ .

When we use the *Laplace transform* on the equation (14) with the given initial conditions, we get an equation of the form

$$(15) \quad Y(s) = G(s) \cdot F(s) + H(s), \quad \text{for } \Re s > \sigma,$$

where  $Y(s) := \mathcal{L}\{y\}(s)$  and  $F(s) := \mathcal{L}\{f\}(s)$  are shorthand for the Laplace transforms of the functions  $y(t)$  and  $f(t)$ , and the *transition function*  $G(s)$  depends on the system and the initial conditions of  $f$ .

We shall postpone the discussion of the *Laplace transform* itself to *Ventus: Complex Functions Theory a-4* and just note that (15) is a typical equation in *Cybernetics*. The solution method is then to solve (15) with respect to the Laplace transform  $F(s) = \mathcal{L}\{f\}(s)$  of the unknown function  $f(t)$ ,

$$(16) \quad F(s) = \frac{Y(s) - H(s)}{G(s)} \quad \text{for } \Re s > \sigma,$$

and then apply the *inverse Laplace transform* on (16) *formally* to get the solution

$$(17) \quad f(t) = \mathcal{L}^{-1} \left\{ \frac{Y(s) - H(s)}{G(s)} \right\} (t),$$

and then hope for some residuum formula, because the function in (16) is analytic. Such residuum formulæ exist, but they are too complicated to bring here, so they are postponed to *Ventus: Complex Functions Theory a-4*.

In principle, this seems to be a straightforward solution procedure, but in practice (17) may be very difficult to compute, even with the announced residuum formula. We shall therefore instead consider a simpler problem, namely *given the transition function*  $G(s)$  [derived e.g. by means of the Laplace transform], *what can be said about the stability of the solution (17) without computing (17)?*

First note that in (16) only the transition function  $G(s)$  occurs in the denominator. This means that the *poles* of  $F(s)$  in (16) are precisely the *zeros* of the transition function  $G(s)$ .

We saw above that the simple differential equation (11) has stable solutions, if and only if all roots of the characteristic polynomial (12) lie in the open left half plane,  $\Re z < 0$ . The same holds in general, so we shall search for criteria which guarantee that all zeros of  $G(s)$  lie in the open left half plane, which will imply stability of the solution (17) without computing it explicitly.

In the loose motivation above we have allowed ourselves to use the Laplace transform in spite of the fact that it has not yet been formally introduced. For historical reasons we have used the traditional notation  $s \in \mathbb{C}$  of the variable of the Laplace transform. Since we in the remaining part of this section no longer shall use the Laplace function and only focus on the given transition functions as *analytic functions*, we switch back to the usual complex variable  $z \in \mathbb{C}$ . We shall furthermore assume, which is customary in *Cybernetics*, that the transition function  $G(z)$  is a *fractional function*, i.e. a fraction of two *polynomials*,

$$G(z) = \frac{P(z)}{Q(z)} = \frac{a_0 z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n}{b_0 z^m + b_1 z^{m-1} + \dots + b_{m-1} z + b_m},$$

where  $P(z)$  and  $Q(z)$  are never zero at the same points. In fact, if  $z_0$  was a zero of both  $P(z)$  and  $Q(z)$ , then  $z - z_0$  would be a factor in both polynomials, so it can be canceled.

Having made these assumptions we see that the zeros of  $G(z)$  are precisely the zeros of the numerator  $P(z)$ , so it is no longer necessary to include the denominator  $Q(z)$ . Therefore, we need only consider the numerator

$$(18) \quad P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n$$

in the following.

Since the conjugated polynomial

$$\overline{P}(z) := \overline{a_0} z^n + \overline{a_1} z^{n-1} + \dots + \overline{a_{n-1}} z + \overline{a_n} = \overline{a_0 \overline{z}^n + a_1 \overline{z}^{n-1} + \dots + a_n}$$

has the complex conjugated roots of the roots of  $P(z)$ , we conclude that the roots of the product polynomial  $P(z)\overline{P}(z)$  must either be real or can be paired as complex conjugated roots. This implies that  $P(z)\overline{P}(z)$  must have real coefficients. Since complex conjugation maps the left (or right) half plane into itself, it follows that we without loss of generality in the following may assume that the polynomial (18) has only real coefficients,  $a_0, \dots, a_n \in \mathbb{R}$ .

We introduce

**Definition 1.2.1** Let  $P(z)$  be a polynomial of real coefficients. If all its zeros lie in the open left half plane  $\Re z < 0$ , then  $P(z)$  is called a Hurwitz polynomial.

The importance of Hurwitz polynomials in the engineering applications is obvious from the above. If the rational function  $G(z) = \frac{P(z)}{Q(z)}$  of real coefficients is a transition function of a dynamical system, then its solution is stable if and only if the numerator  $P(z)$  is a Hurwitz polynomial.

By the Fundamental Theorem of Algebra,

$$(19) \quad P(z) = a_0 z^n + \dots + a_n = a_0 (z - \lambda_1) \dots (z - \lambda_n), \quad z \in \mathbb{C}.$$

Assume that  $P(z)$  is a Hurwitz polynomial and that  $\alpha + i\beta$ ,  $\beta \neq 0$ , is a complex root, necessarily lying in the left half plane according to Definition 1.2.1, so  $\alpha < 0$ . Since  $P(z)$  has real coefficients,  $\alpha - i\beta$  is also a root (of the same multiplicity). It follows from  $\alpha < 0$  and

$$(z - \alpha - i\beta)(z - \alpha + i\beta) = (z - \alpha)^2 + \beta^2 = z^2 - 2\alpha z + (\alpha^2 + \beta^2)$$

that this polynomial has positive coefficients. This is also trivially the case, if  $\lambda < 0$  is a real root. When we multiply all these factors of positive coefficients, we again obtain a polynomial of positive coefficients. We have therefore proved

**Theorem 1.2.1** Given a polynomial

$$P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_n, \quad a_0, a_1, \dots, a_n \in \mathbb{R},$$

of real coefficients. A necessary condition that  $P(z)$  is a Hurwitz polynomial is that  $a_0, a_1, \dots, a_n$  are all positive, or all are negative.

We only need to add to the proof above that if  $a_0 < 0$ , then

$$\frac{1}{a_0} P(z) = z^n + \frac{a_1}{a_0} z^{n-1} + \dots + \frac{a_n}{a_0}$$

has only positive coefficients, so  $a_0, a_1, \dots, a_n$  have all the same sign – either all positive, or all negative.

**Remark 1.2.1** If  $n = 1$  or  $n = 2$ , then the necessary condition of Theorem 1.2.1 is also sufficient. For  $n \geq 3$  this is no longer true. The polynomial

$$16z^3 + 8z^2 + 9z + 17$$

of third degree has only positive coefficients, and its roots are  $-1, \frac{1}{4} \pm i$ , so it is not a Hurwitz polynomial, because it has two roots in the right half plane.  $\diamond$

**Theorem 1.2.2** *A necessary and sufficient condition that every root of a polynomial of degree 3,*

$$P(z) = a_0z^3 + a_1z^2 + a_2z + a_3$$

*is a Hurwitz polynomial, is that all coefficients have the same sign and that*

$$a_1a_2 - a_0a_3 > 0.$$

PROOF. We need the following general result, that if the roots are denoted by  $r_1$ ,  $r_2$  and  $r_3$ , then

$$\begin{aligned} a_0z^3 + a_1z^2 + a_2z + a_3 &= a_0(z - r_1)(z - r_2)(z - r_3) \\ (20) \qquad \qquad \qquad &= a_0z^3 - a_0(r_1 + r_2 + r_3)z^2 + a_0(r_1r_2 + r_2r_3 + r_3r_1)z - r_1r_2r_3. \end{aligned}$$

We either have one real root  $r$  and one pair of conjugate complex roots  $p \pm iq$ ,  $q \neq 0$ , or three real roots.

First assume that we have the roots  $r \in \mathbb{R}$  and  $p \pm iq$ , where  $p, q \in \mathbb{R}$  and  $q \neq 0$ .

We get by identifying the coefficients of (20),

$$(21) \quad \frac{a_1}{a_0} = -(r_1 + r_2 + r_3) = -(r + 2p),$$

$$(22) \quad \frac{a_2}{a_0} = r_1 r_2 + r_2 r_3 + r_3 r_1 = p^2 + q^2 + 2pr,$$

$$(23) \quad \frac{a_3}{a_0} = -r_1 r_2 r_3 = -r(p^2 + q^2).$$

Since all coefficients have the same sign, (21), (22) and (23) must all be positive. It follows in particular from (23) that  $r < 0$ , and we have proved that the real root lies in the left half plane.

Then we assume that  $a_1 a_2 - a_0 a_3 > 0$  and shall prove that the two roots  $p \pm iq$  lie in the left half plane, i.e. we shall only prove that  $p < 0$ . When the assumption is divided by  $a_0^2 > 0$ , we get by using (21), (22) and (23) that

$$(24) \quad \begin{aligned} 0 < \frac{a_1}{a_0} \cdot \frac{a_2}{a_0} - \frac{a_3}{a_0} &= -(r + 2p)(p^2 + q^2 + 2pr) + r(p^2 + q^2) \\ &= -2p\{(p^2 + q^2 + 2pr) + r^2\} = -2p\left\{\frac{a_2}{a_0} + r^2\right\}. \end{aligned}$$

Since (22) is positive, it follows from (24) that  $p < 0$ , and we have proved that  $P(z)$  is a Hurwitz polynomial in this case.

If all three roots are real, it simply follows from the assumption of the  $a_j$  having the same sign that the value of  $P(x)$  along the positive real axis, as well as at 0 has the same sign, so all three real roots must be negative, and  $P(z)$  is a Hurwitz polynomial.

Then assume that the three roots  $r$  and  $p \pm iq$ , one real and a pair of conjugate complex roots, all lie in the left half plane. Then we already know from Theorem 1.2.1 that all the coefficients of the polynomial have the same sign. We shall prove that  $a_1 a_2 - a_0 a_3 > 0$ . This follows from the reverse computations, cf. (24),

$$\begin{aligned} 0 < -2p\left\{\frac{a_2}{a_0} + r^2\right\} &= -2p\{(p^2 + q^2 + 2pr) + r^2\} \\ &= -(r + 2p)(p^2 + q^2 + 2pr) + r(p^2 + q^2) = \frac{a_1}{a_0} \cdot \frac{a_2}{a_0} - \frac{a_3}{a_0} = \frac{a_1 a_2 - a_0 a_3}{a_0^2}, \end{aligned}$$

and  $a_1 a_2 - a_0 a_3 > 0$  follows.

If all three roots,  $r$  and  $p \pm iq$  are real and negative, then  $r < 0$  and  $|q| < |p| = -p$ , and we still get (21), i.e.

$$\frac{a_1}{a_0} = -(r + 2p) > 0.$$

Formulæ (22) and (23) are replaced by a change of sign on  $q^2$ , so we get

$$\frac{a_2}{a_0} = p^2 - q^2 + 2pr > 0 \quad \text{and} \quad \frac{a_3}{a_0} = -r(p^2 - q^2) > 0.$$

Then the reverse of (24) is modified in the following way,

$$\begin{aligned} 0 < -2p \left\{ \frac{a_2}{a_0} + r^2 \right\} &= -2p \{ (p^2 - q^2 + 2pr) + r^2 \} \\ &= -(r + 2p) (p^2 - q^2 + 2pr) + r (p^2 - q^2) = \frac{a_1}{a_0} \cdot \frac{a_2}{a_0} - \frac{a_3}{a_0} = \frac{a_1 a_2 - a_0 a_3}{a_0^2}, \end{aligned}$$

and  $a_1 a_2 - a_0 a_3 > 0$  also follows in this case, and the theorem is proved.  $\square$

We shall now prove another result on Hurwitz polynomials. Given the factorization (19), i.e.

$$P(z) = a_0 z^n + \cdots + a_n = a_0 (z - \lambda_1) \cdots (z - \lambda_n), \quad z \in \mathbb{C}.$$

If  $\lambda_j = \alpha + i\beta$  and  $\Re \lambda_j = \alpha < 0$ , then for  $z = x + iy$ ,

$$|z - \lambda_j|^2 = (x - \alpha)^2 + (y - \beta)^2.$$

Thus, for  $x > 0$  and  $\alpha < 0$ ,

$$|z - \lambda_j|^2 = (x - \alpha)^2 + (y - \beta)^2 > (-x - \alpha)^2 + (-y + \beta)^2 = |-z - \bar{\lambda}_j|^2.$$

This implies that if  $\beta = 0$ , then

$$|z - \alpha| > |-z - \alpha| \quad \text{for } \Re z > 0,$$

and if  $\beta \neq 0$ , then

$$|(z - \lambda_j)(z - \bar{\lambda}_j)| > |(-z - \bar{\lambda}_j)(-z - \lambda_j)| \quad \text{for } \Re z > 0.$$

When we multiply all the factors, we conclude that if  $P(z)$  is a Hurwitz polynomial, then

$$|P(z)| > |P(-z)| \quad \text{for } \Re z > 0,$$

and replacing  $z$  by  $-z$ ,

$$|P(z)| < |P(-z)| \quad \text{for } \Re z < 0.$$

We notice that if  $x = 0$ , then

$$|P(z)| = |P(iy)| = |P(-iy)| = |P(-z)| \quad \text{for } \Re z = 0.$$

Conversely, assume that  $P(z)$  is a polynomial of real coefficients, such that

$$|P(z)| > |P(-z)| \quad \text{for } \Re z > 0.$$

Then in particular  $P(z) \neq 0$  for  $\Re z > 0$ , so the zeros of  $P(z)$  must satisfy  $\Re z \leq 0$ . Hence, if we require that there are no zeros on the imaginary axis, i.e. we assume that  $P(iy) \neq 0$  for all  $y \in \mathbb{R}$ , then  $P(z)$  is a Hurwitz polynomial, and we have proved

**Theorem 1.2.3** *Let  $P(z)$  be a polynomial of real coefficients. Then  $P(z)$  is a Hurwitz polynomial, if and only if*

$$\begin{cases} |P(z)| > |P(-z)| & \text{for } \Re z > 0, \\ P(iy) \neq 0 & \text{for } y \in \mathbb{R}. \end{cases}$$

Assume that  $P(z)$  is a Hurwitz polynomial, and put

$$R(z) := P(1)P(z) - P(-1)P(-z).$$

When Theorem 1.2.3 is applied on  $P$ , we get

$$|P(1)P(z)| > |P(-1)P(-z)| \quad \text{for } \Re z \geq 0,$$

so we conclude that  $R(z) \neq 0$  for  $\Re z \geq 0$ , and  $R(z)$  is a Hurwitz polynomial.

A trivial root of  $R(z)$  is  $z = -1$ , so the quotient  $\frac{R(z)}{z+1}$  must be a polynomial, and it follows from the above that

$$(25) \quad Q(z) := \frac{R(z)}{z+1} = \frac{P(1)P(z) - P(-1)P(-z)}{z+1}$$

is a Hurwitz polynomial.

Then assume that  $P(z)$  is *not* a Hurwitz polynomial and that all its coefficients have the same sign. We shall prove that  $Q(z)$  given by (25) is *not* a Hurwitz polynomial. If we can prove this, then it follows that  $P(z)$  is a Hurwitz polynomial, if and only if  $Q(z)$  is a Hurwitz polynomial.

First assume that  $P(z)$  has a root  $iy_0$  on the imaginary axis. Then  $-iy_0$  is also a root, because  $P(z)$  has real coefficients, and it follows by insertion into (25) that

$$Q(iy_0) = \frac{P(1)P(iy_0) - P(-1)P(-iy_0)}{iy_0 + 1} = \frac{0 - 0}{iy_0 + 1} = 0,$$

so  $Q(z)$  is not a Hurwitz polynomial in this case.

Then assume that  $P(z_0) = 0$  for some  $z_0$  in the right half plane  $\Re z_0 > 0$ . Then

$$(z_0 + 1)Q(z_0) = -P(-1)P(-z_0) \quad \text{and} \quad (-z_0 + 1)Q(-z_0) = P(1)P(-z_0).$$

If also  $P(-z_0) = 0$ , then clearly  $Q(z_0) = 0$ , because  $z_0 + 1 \neq 0$  for  $\Re z_0 > 0$ . Hence,  $Q(z)$  is not a Hurwitz polynomial in this case.

The remaining possibility is that  $P(-z_0) \neq 0$ . Since we have assumed that the real constants  $a_0, \dots, a_n$  all have the same sign, either positive or negative, we get

$$|P(-1)| = |(-1)^n a_0 + (-1)^{n-1} a_1 + \dots + a_n| < |a_0 + a_1 + \dots + a_n| = |P(1)|,$$

from which follows that

$$|(z_0 + 1)Q(z_0)| = |P(-1)| \cdot |P(-z_0)| < |P(1)| \cdot |P(-z_0)| = |(-z_0 + 1)Q(-z_0)|,$$

from which

$$|Q(z_0)| < \left| \frac{1 - z_0}{1 + z_0} \right| \cdot |Q(-z_0)| < |Q(-z_0)|,$$

because  $\Re z_0 > 0$  by assumption. Since  $z_0$  lies in the right half plane, this estimate shows that  $Q(z)$  cannot be a Hurwitz polynomial, and the claim is proved.

Summing up, we have shown

**Theorem 1.2.4** Schur's criterion. *The polynomial  $P(z)$  of real coefficients is a Hurwitz polynomial, if and only if the following two conditions are fulfilled,*

- 1) *all coefficients of  $P(z)$  have the same sign, positive or negative,*
- 2) *the polynomial*

$$Q(z) = \frac{P(1)P(z) - P(-1)P(-z)}{z + 1}$$

*is a Hurwitz polynomial.*

Since  $Q(z)$  clearly is of lower degree than  $P(z)$ , we can repeat the application of *Schur's criterion* and thus get a sequence of polynomials  $Q_j(z)$  of decreasing degrees, which either all are Hurwitz polynomials, or none of them are Hurwitz polynomials. This process stops after at most  $n - 2$  steps.

**Example 1.2.1** Consider the polynomial

$$P(z) = z^3 + 2z^2 + 3z + 1$$

of positive coefficients, so condition 1) of *Schur's criterion* is fulfilled.



Then compute

$$Q(z) = \frac{1}{z+1} \{7(z^3+2z^2+3z+1) - (-1)(-z^3+2z^2-3z+1)\} = 6z^2 + 10z + 8.$$

We see that  $Q(z)$  is a polynomial of degree 2 with only positive coefficients, so it is a Hurwitz polynomial, and we conclude from *Schur's criterion* that  $P(z) = z^3 + 2z^2 + 3z + 1$  is also a Hurwitz polynomial.  $\diamond$

**Example 1.2.2** In case of the polynomial

$$P(z) = z^3 + 2z^2 + z + 3$$

we find

$$Q(z) = \frac{1}{z+1} \{7(z^3+2z^2z+3) - 3(-z^3+2z^2-z+3)\} = 10z^2 - 2z + 12.$$

Since  $Q(z)$  does not have all coefficients of the same sign, it is not a Hurwitz polynomial, so  $P(z)$  is not a Hurwitz polynomial.  $\diamond$

The next result is one of the most famous theorems in the theory of stability. Unfortunately, its proof is very long, so it is not reasonable to give it here. Furthermore, we have already complete results for  $n = 1$  and  $n = 2$ , where Theorem 1.2.1 both gives necessary and sufficient conditions for Hurwitz polynomials, and for  $n = 3$  where Theorem 1.2.2 does the same, so in practice it is only needed for  $n \geq 4$ , and when  $n$  becomes too big, the computations become large.

**Theorem 1.2.5** Routh-Hurwitz Stability Criterion (1895). *Assume that the polynomial  $P(z) = a_0z^n + a_1z^{n-1} + \dots + a_{n-1}z + a_n$  has positive coefficients,  $a_0, \dots, a_n > 0$ . Then  $P(z)$  is a Hurwitz polynomial, if and only if the following inequalities hold,*

$$D_1 = a_1 > 0, \quad D_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix} > 0, \quad D_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix} > 0, \quad \dots,$$

$$D_n = \begin{vmatrix} a_1 & a_0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{2n-1} & a_{2n-2} & a_{2n-3} & \dots & a_n \end{vmatrix} > 0,$$

where we put  $a_k = 0$ , whenever  $k > n$ .

**Example 1.2.3** In spite of what was mentioned above we shall here in the first two examples only consider polynomials of degree 3. We get for the polynomial  $z^3 + 2z^2 + 3z + 1$ , cf. Example 1.2.1,

$$D_1 = 2, \quad D_2 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5, \quad D_3 = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 5,$$

and it follows from *Routh-Hurwitz criterion* that all roots lie in the left half plane  $\Re z < 0$ .  $\diamond$

**Example 1.2.4** Then consider the polynomial  $z^3 + 2z^2 + z + 3$ , cf. also Example 1.2.2. In this case we get

$$D_2 = \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = -1 < 0,$$

(no need to compute  $D_3$ ), so the polynomial has at least one root in the closed right half plane  $\Re z \geq 0$ . The values on the imaginary axis are

$$P(iy) = \{3 - 2y^2\} + iy(1 - y^2) \neq 0 \quad \text{for all } y \in \mathbb{R},$$

so we must have at least one root in the right half plane.

All coefficients are positive, so this root cannot be positive, and the polynomial must have two complex conjugated roots in the right half plane. The third root must be real and negative.

By means of a pocket calculator we find the approximate values of the three roots,

$$-2.17455941 \quad \text{and} \quad 0.0872797 \pm 1.1713121 \cdot i. \quad \diamond$$

**Example 1.2.5** Finally, let us consider a polynomial of positive coefficients and degree 5,

$$P(z) = z^5 + z^4 + 2z^3 + z^2 + z + 2.$$

We are in the situation that the roots can easily be found by inspection, because we immediately get the factorization

$$P(z) = (z^2 + z + 2)(z^3 + 1),$$

so its roots are

$$z_1 = -1, \quad \left. \begin{matrix} z_2 \\ z_3 \end{matrix} \right\} = \frac{1}{2} \pm i \frac{\sqrt{3}}{2}, \quad \left. \begin{matrix} z_4 \\ z_5 \end{matrix} \right\} = -\frac{1}{2} \pm i \frac{\sqrt{7}}{2},$$

and we see that  $z_2$  and  $z_3$  lie in the right half plane, so  $P(z)$  is not a Hurwitz polynomial.

We shall now instead apply *Routh-Hurwitz criterion*, and this time we shall do it to the very end, even if we could stop at  $D_4 < 0$ . The easy part is to compute

$$D_1 = 1 > 0, \quad D_2 = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 > 0, \quad D_3 = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{vmatrix} = 2 > 0.$$

Then we continue with

$$D_4 = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{vmatrix} = -2 \cdot \begin{vmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{vmatrix} = -2 - 2 = -4 < 0,$$

and we conclude that  $P(z)$  is not a Hurwitz polynomial. However, for completeness we also compute

$$D_5 = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 1 & 0 \\ 2 & 1 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{vmatrix} = 2 \begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{vmatrix} = -2 \cdot 2 \begin{vmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{vmatrix} + 2 \cdot 1 \begin{vmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{vmatrix} = -12 + 4 = -8 < 0,$$

demonstrating that the computations become big, when  $n$  is large.  $\diamond$

The stability criteria above are all algebraic. There also exist geometrical stability criteria, of which we shall only mention *Nyquist's criterion*.

Consider a control system with a simple feedback. It can be shown that the corresponding *transition function* is given by

$$f(z) = \frac{H(z)}{1 + H(z)} = 1 - \frac{1}{1 + H(z)},$$

where  $H(z)$  is some fractional function.

The poles of  $f(z)$  are precisely the zeros of  $H(z) + 1$ .

Denote by  $P$  the number of poles of  $f$  in the right half plane, and consider the curve  $C_R$  on Figure 7. In engineering sciences,  $C_R$  is conventionally always given the opposite orientation of that known in Mathematics, cf. the arrows on Figure 7. This conventional change of orientation will in the following cause a trivial change of the sign.

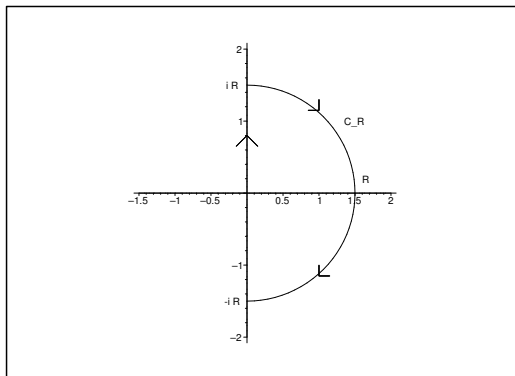


Figure 7: The closed curve  $C$  in the proof of Nyquist's criterion. Notice the reverse orientation of the curve.

It follows from the *argument principle* that  $f(z)$  does *not* have poles in the right half plane, if and only if there exists an  $R_0 > 0$ , such that

$$\Delta_R \arg f(z) = 2\pi P \quad \text{for every } R \geq R_0,$$

where  $\Delta_R \arg f(z)$  denotes the growth of the argument of  $f(z)$ , when  $z$  runs through  $C_R$  in the direction given on Figure 7, i.e. in the mathematically *negative* sense.

On the other hand,

$$\frac{1}{2\pi} \Delta_R \arg f(z) \quad [= P] \quad \text{for } R \geq R_0,$$

is equal to the *winding number* of  $f(z)$  with respect to 0 in the  $w$ -plane (counted *positive* when traversed in the mathematically *negative* sense), when  $z$  runs once through  $C_R$ , also in the negative sense as indicated on Figure 7. It follows from the identity

$$1 - f(z) = \frac{1}{1 + H(z)}$$

that this winding number is equal to the winding number of  $\frac{1}{1 + H(z)}$  with respect to  $w = +1$ . Finally, by taking the limit  $R \rightarrow +\infty$  we have proved,

**Theorem 1.2.6** Nyquist's criterion (1932). *Let  $H(z) = \frac{P(z)}{Q(z)}$  be a rational function, where  $P(z)$  and  $Q(z)$  are polynomials, and assume that  $H(z) \neq -1$  on the imaginary axis. Let  $P$  denote the number of poles of  $\frac{1}{1 + H(z)}$  in the right half plane.*

*The control system of simple feedback corresponding to the transition function  $f(z) = \frac{H(z)}{1 + H(z)}$  is stable, if and only if the winding number with respect to the point  $w = +1$  of  $\frac{1}{1 + H(z)}$ , when  $z$  runs through the imaginary axis, is equal to this number of poles  $P$  in the right half plane.*

The transition function  $f(z) = \frac{H(z)}{1 + H(z)}$  is a special way of writing transition functions of the type

$$f(z) = \frac{1}{K \cdot G(z)} + 1,$$

where  $G(z)$  is a rational function, and where the constant  $K$  is called the *amplification factor* of the control system.

It is simple algebra to see that we have the correspondences

$$H(z) = -1 - K \cdot G(z) \quad \text{and} \quad G(z) = -\frac{1}{K} \{1 + H(z)\}.$$

Hence, we easily derive the following generalized criterion of stability.

**Theorem 1.2.7** Nyquist–Michailow's criterion. *Assume that  $G(z)$  is a rational function with no zeros on the imaginary axis, and with  $N$  zeros in the right half plane.*

*The control system corresponding to the transition function*

$$f(z) = \frac{1}{K \cdot G(z)} + 1$$

*is stable, if and only if  $\frac{1}{G(z)}$  has the winding number  $N$  with respect to the point  $w = -K$ , when  $z$  runs through the imaginary axis from  $-i\infty$  to  $+i\infty$ .*

## 2 Many-valued functions and Riemann surfaces

### 2.1 Inverse mapping

We shall start this section by proving the following important theorem.

**Theorem 2.1.1** *Let  $f$  be analytic in an open domain  $\Omega$ , and assume that  $f : \Omega \rightarrow f(\Omega)$  is bijective. Then the inverse map  $f^{-1} : f(\Omega) \rightarrow \Omega$  exists and is analytic in  $f(\Omega)$  and its derivative is given by*

$$(26) \quad \frac{d}{dw} f^{-1}(w) = \frac{1}{f'(f^{-1}(w))}.$$

PROOF. The inverse  $f^{-1}$  trivially exists, because  $f$  is bijective. We shall prove that  $f^{-1}$  is also analytic in  $f(\Omega)$ .

It follows from a theorem on *local inverse maps* in Ventus, *Complex Functions Theory a-1*, Chapter 3, that the inverse function  $f^{-1}$  is analytic in a neighbourhood of every point  $w_0 = f(z_0)$ , for which  $f'(z_0) \neq 0$ , and if so, then its derivative at  $w_0$  is given by (26). Thus, we shall only prove that  $f'(z) \neq 0$  everywhere in  $\Omega$ .

We shall prove this claim by *contraposition*. Proving the claim is equivalent to prove that if  $f'(z_0) = 0$  for some  $z_0 \in \Omega$ , then  $f$  is *not* injective in any neighbourhood of  $z_0$ . Therefore, we shall assume that  $f'(z_0) = 0$  in the following, and we shall find a neighbourhood of  $z_0$ , such that  $f$  is not injective in any smaller neighbourhood of  $z_0$ .

Then we apply a result from Ventus, *Complex Functions Theory a-2* on series, by which there is a smallest integer  $n > 1$ , such that  $f^{(n)}(z_0) \neq 0$ . If we write  $w_0 = f(z_0)$ , then we have in a neighbourhood of  $z_0$  that the series expansion of  $f(z)$  is locally given by

$$w = f(z) = w_0 + a_n (z - z_0)^n + a_{n+1} (z - z_0)^{n+1} + \dots = w_0 + (z - z_0)^n g(z),$$

where

$$g(z) = a_n + a_{n+1} (z - z_0) + \dots$$

is analytic in a neighbourhood of  $z_0$ , and  $g(z_0) = a_n \neq 0$ .

Choose any disc  $B(z_0, r)$  contained in this neighbourhood, such that also

$$\left| 1 - \frac{1}{a_n} g(z) \right| < 1 \quad \text{for } z \in B(z_0, r).$$

The use a result from Ventus, *Complex Functions Theory a-2* on series to define an analytic function  $h(z) = \sqrt[n]{g(z)}$  by

$$\begin{aligned} h(z) &= \sqrt[n]{g(z)} = \sqrt[n]{a_n} \cdot \sqrt[n]{1 + \left\{ \frac{1}{a_n} g(z) - 1 \right\}} := b_0 \sum_{k=0}^{+\infty} \binom{\frac{1}{2}}{k} \cdot \left\{ \frac{1}{a_n} g(z) - 1 \right\}^k \\ &= \sum_{k=0}^{+\infty} b_k (z - z_0)^k \quad \text{for } z \in B(z_0, r), \end{aligned}$$

where  $b_0$  is any one of the  $n$  solutions of the binomial equation  $z^n = a_n$ . Then  $h(z)^n = g(z)$ .

If we put

$$k(z) = (z - z_0)h(z) = b_0(z - z_0) + b_1(z - z_0)^2 + \dots,$$

then

$$w = f(z) = w_0 + \{k(z)\}^n \quad \text{for } z \in B(z_0, r).$$

Therefore, the map  $w = f(z) : B(z_0, r) \rightarrow \mathbb{C}$  can be written as the composition of the three analytic maps

$$t = k(z), \quad s = t^n, \quad w = w_0 + s.$$

Since  $k'(z_0) = b_0 \neq 0$ , it follows from the already quoted inverse theorem in *Ventus, Complex Functions Theory a-1*, Chapter 3, that there exists a neighbourhood  $\Omega_0 \subseteq B(z_0, r)$  of  $z_0$ , which by  $t = k(z)$  is mapped bijectively onto a neighbourhood  $T_0$  of  $k(z_0) = 0$ . We may of course choose  $\Omega_0$ , such that  $T_0 = B(0, \alpha)$  is an open disc of centre 0 and radius  $\alpha > 0$ .

Then using the map  $s = t^n$ , where  $n > 1$ , the disc  $B(0, \alpha_1)$  is mapped surjectively, but not injectively onto the disc  $B(0, \alpha_1^n)$  for every  $0 < \alpha_1 \leq \alpha$ . Just notice that every number  $\exp\left(p \cdot \frac{2i\pi}{n}\right) \cdot t$  for  $p = 1, \dots, n$  and  $0 < t < \alpha$ , is mapped into the same point  $t^n$ . We therefore conclude that if  $f'(z_0) = 0$ , then the map cannot be injective in any neighbourhood of  $z_0$ , because such a neighbourhood contains some  $B(0, \alpha_1)$ , where  $0 < \alpha_1 \leq \alpha$ , and the theorem is proved.  $\square$

In order to find the inverse map we note that it is important first to find a subset  $\Omega$  of the  $z$ -plane, on which  $f$  is *injective*. We know already that  $\exp \mathbb{C} = \mathbb{C} \setminus \{0\}$ , and that the points  $z + 2i\pi$ ,  $p \in \mathbb{Z}$ , are all mapped into the same point  $e^z$  by the exponential, so  $\exp$  is not injective, unless we restrict ourselves to suitable smaller subsets of  $\mathbb{C}$ . On the other hand, we know that all solutions of the equation  $e^w = z$  for  $z \in \mathbb{C} \setminus \{0\}$  are given by

$$(27) \quad w = \text{Log } z + 2ip\pi, \quad p \in \mathbb{Z},$$

so even if there are infinitely many solutions in (27), they lie at a mutual distance  $\geq 2\pi$  from each other. It is therefore in practice possible to separate them.

Geometrically this is done by letting each  $p \in \mathbb{Z}$  in (27) correspond to precisely one copy of  $f(\Omega_p) = \mathbb{C} \setminus \{0\}$ , and then put all these  $f(\Omega_p)$ ,  $p \in \mathbb{Z}$ , above each other, slice them along the real negative axis  $\mathbb{R}_-$  and then glue them together along these slits. In the present case we get a model of an infinite winding staircase, which is called the *Riemann surface* of (the inverse map of) the exponential. In other words, we replace the image  $f(\Omega)$  by this *Riemann surface* and obtain that  $\Omega = \mathbb{C}$  by  $\exp$  is in a bijective correspondence with this Riemann surface, i.e. every point in  $\mathbb{C}$  corresponds precisely to one point on the Riemann surface – or to one point in just one sheet  $f(\Omega_p)$  for some  $p \in \mathbb{Z}$ .

Riemann surfaces may be extremely complicated, so only the simplest ones are described in the following sections, with the exception of Section 2.8, where we demonstrate the general technique by a more complicated example.

One should also mention that instead of creating a Riemann surface of the inverse of  $\exp$  we could consider (27) as a definition of the *many-valued* inverse function, so the image of one point is given by several complex numbers. We shall mainly adopt this aspect in the following.

Since the topic of this chapter is the inverse of a map, it is quite natural in the following to interchange  $z$  and  $w$ . Thus we shall be given an analytic function  $f(w)$  in the  $w$ -space, and we shall find a reasonable expression of the inverse  $f^{-1}(z)$ . This is the same as solving the equation  $z = f(w)$  with respect to  $w$  for given  $z \in f(\Omega)$ .

We used above in Theorem 2.1.1 that  $f^{-1}$  is always *locally* defined in some neighbourhood of  $f(w)$ , if  $f'(w) \neq 0$ , and that this is *never* possible, if instead  $f'(w) = 0$ . Here we used a result from *Ventus, Complex Functions Theory a-1*, Chapter 3. This observation leads to the following natural definition.

**Definition 2.1.1** *Let  $f : \Omega \rightarrow \mathbb{C}$  be analytic in an open domain  $\Omega$  in the  $w$ -plane. A point  $w_0 \in \Omega$  is called a singular point of the map  $f(w)$ , if  $f'(w_0) = 0$ .*

The *image*  $z_0 = f(w_0)$  of a singular point  $w_0$  becomes a *branching point* of the corresponding Riemann surface. Such branching points are easy to handle, which we shall see in examples in the following.

There also exists another type of branching points which is far more difficult to treat. One example is given by (27), where the point 0 does not belong to the domain of the right hand side for any  $p \in \mathbb{Z}$ . We shall say that  $z_0$  is a *logarithmic branching point*, if  $z_0$  is an isolated point of  $\mathbb{C} \setminus f(\Omega)$ , and if furthermore we cannot find any open domain  $\Omega_1 \supset \Omega$  and any analytic function  $f_1 : \Omega_1 \rightarrow \mathbb{C}$ , such that  $f_1(w) = f(w)$  for  $w \in \Omega$  and  $z_0 \in f_1(\Omega_1)$ .



The program of solving the equation  $z = f(w)$  with respect to  $w \in \Omega$  can now be described in the following way:

- 1) First find the *singular points*  $w_0 \in \Omega$ , i.e. the solutions of the equation  $f'(w) = 0$ .
- 2) Then find the *branching points*  $z_0 = f(w_0)$ , and notice that  $f$  does not establish any bijective connection between any neighbourhood of  $w_0$  and  $z_0$ .
- 3) Finally, check if the geometry of  $f(\Omega)$  allows the possibility of *logarithmic branching points*.
- 4) Once  $f(\Omega)$  has been identified, find the largest possible *open* subdomains  $\omega \subseteq \Omega$ , such that  $f : \omega \rightarrow f(\omega)$  is bijective. Every such maximum open subdomain is called a *fundamental domain*.
- 5) If the boundary of a fundamental domain  $\omega$  is composed of piecewise differentiable curves contained in  $\Omega$ , then the boundary of  $f(\Omega) \setminus f(\omega)$  is composed of curves between either *branching points* (logarithmic or not) or  $\infty$ . Every such curve in the  $z$ -plane is called a *branch cut*, and the uniquely determined inverse map  $f^{-1} : f(\omega) \rightarrow \omega$ , where  $\omega$  is a *fundamental domain*, is called a *branch* of  $f$ . Thus a *branch* is always a *single-valued* analytic function.

Notice that if  $f$  is not bijective then the fundamental domains and the branch cuts and the branches of  $f$  are *not* uniquely determined. On the other hand, the branching points are always fixed.

**Remark 2.1.1** It is in many cases possible, by using the fundamental domains, the branch cuts, some sheets of paper, a pair of scissors, some glue or tape, and a lot of patience, to construct reasonable models of the Riemann surface of a of a given analytic function.  $\diamond$

We shall not always describe the Riemann surface of a given analytic function, but we shall at least always try to find an expression of the *many-valued function*  $f^{-1}(z)$ .

We shall in the following sections demonstrate how the above loosely described theory more explicitly is applied in some classical examples. First we mention, however, the following *geometric criterion of stability* in which we also include Riemann surfaces.

Assume that  $f(z)$  is a polynomial or a rational function. Then there exists a *Riemann surface*  $\mathcal{R}$  of  $f(z)$ , such that  $f$  maps the  $z$ -plane bijectively onto  $\mathcal{R}$ . The question of *stability*, cf. Section 1.2, is then reduced to the statement that the image  $f(\overline{\Omega})$  of the closed right half plane  $\overline{\Omega} := \{z \in \mathbb{C} \mid \Re z \geq 0\}$  does not contain points above the point  $w = 0$ . If the image  $f(\overline{\Omega})$  of the right half plane in  $\mathcal{R}$  is bounded by a curve  $\Gamma$ , in which case  $\Gamma$  must be the image of the imaginary axis, then we get the following geometrical stability criterion.

**Theorem 2.1.2** *Assume that the domain  $f(\Omega)$  to the right of the curve  $\Gamma$  on the Riemann surface  $\mathcal{R}$  does not contain any point above  $w = 0$  (and  $\Gamma$  does not pass through any such point), then  $f(z)$  is the transition function of a stable system.*

*If the condition above is not fulfilled, then the system is unstable.*

It is usually in practice difficult to investigate  $\Gamma$  on the Riemann surface  $\mathcal{R}$  itself. One considers instead the *projection*  $\gamma$  of  $\Gamma$  onto the  $w$ -plane. We get a parametric description by putting  $z = iy$ ,  $y \in \mathbb{R}$ , into the expression of  $f(z)$  and then split into real and imaginary parts. Thus we get the parametric description of  $\gamma$ ,

$$u = u(y), \quad v = v(y), \quad y \in \mathbb{R}.$$

The condition that  $w = 0$  must not lie to the right of the curve is here only *necessary*. That it is not sufficient is seen by the example  $f(z) = z^3 - z^2 + 2z - 3$ , in which case  $w = 0$  does not lie to the right of the curve  $\gamma$ . On the other hand, since the coefficients do not have the same sign, it follows from Section 1.2 that this polynomial indeed has a root in the right half plane, so the corresponding dynamic system is unstable.

## 2.2 The two-valued function $\sqrt{z}$ .

The simplest example of a many-valued function is the *square root*  $w = \sqrt{z}$ . In this case we consider, cf. Section 2.1,

$$f : \mathbb{C} \rightarrow \mathbb{C}, \quad z = f(w) = w^2, \quad \Omega = \mathbb{C}.$$

It follows from  $f'(w) = 2w$  that  $w_0 = 0$  is the only *singular point*, corresponding to the *branching point*  $z_0 = 0$ .

Write  $z$  and  $w$  in *polar coordinates*, i.e.  $z = r e^{i\Theta}$  and  $w = R e^{i\varphi}$ . Then

$$r e^{i\Theta} = z = w^2 = R^2 e^{i2\varphi},$$

from which

$$(28) \quad r = R^2 \quad \text{and} \quad \Theta \equiv 2\pi \pmod{2\pi}.$$

It follows from (28) that  $f(\mathbb{C}) = \mathbb{C}$ , so  $\mathbb{C} \setminus f(\mathbb{C}) = \emptyset$ . Hence, we do not have a *logarithmic branching point* in this case, and every *branch cut* must be a simple curve from 0 to  $\infty$ .

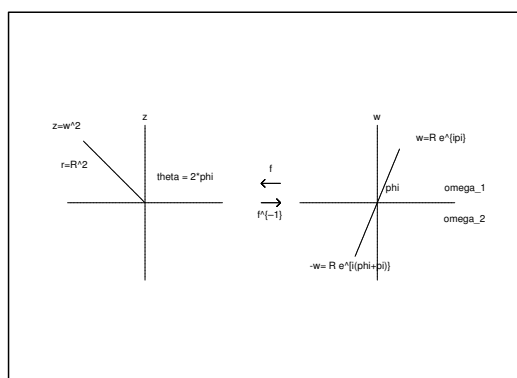


Figure 8: The  $z$ -plane and the  $w$ -plane in case of the map  $z = w^2$ .

It follows from (28) that  $z = f(w)$  has an inverse map, if for every fixed  $R > 0$  the angle  $2\varphi$  lies in an interval of *at most* the length  $2\pi$ , i.e. if the angle  $\varphi$  of  $w$  itself lies in an open interval of at most the length  $\pi$ . In fact, if both  $w$  and  $-w$  belongs to  $\omega \setminus \{0\}$ , then  $f(w) = w^2 = f(-w)$ , and  $f$  is not injective.

Define

$$\omega_1 := \{w \in \mathbb{C} \mid \Im w > 0\} = \{w = R e^{i\Theta} \mid R > 0, 0 < \Theta < \pi\},$$

and

$$\omega_2 := \{w \in \mathbb{C} \mid \Im w < 0\} = \{w = R e^{i\Theta} \mid R > 0, \pi < \Theta < 2\pi\}.$$

Then  $\omega_1$  and  $\omega_2$  are *fundamental domains*, and  $\omega_1 \cup \omega_2 = \mathbb{C} \setminus \mathbb{R}$ , where  $\mathbb{R}$  clearly is mapped into  $\mathbb{R}_+ \cup \{0\}$ . Here 0 is the *branching point*, and  $\mathbb{R}_+$  is the *branch cut* from 0 to  $\infty$ . Using the above we can then construct a model of the *Riemann surface* of the square root.

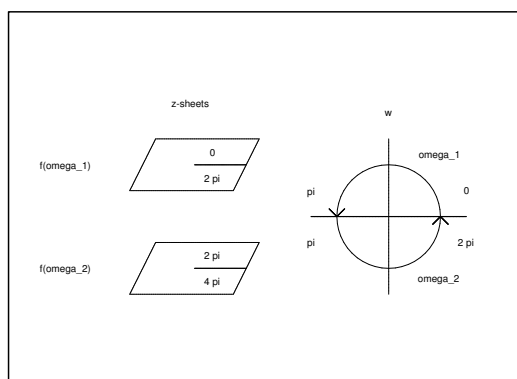


Figure 9: Construction of the Riemann surface for  $w = \sqrt{z}$ .

We take three pieces of paper, one representing the  $w$ -plane, and the remaining ones, two copies of the  $z$ -plane, lying above each other, though at this stage not touching each other. Sketch on the  $w$ -plane the singular point  $w_0 = 0$  and the real axis and the two fundamental domains  $\omega_1$  and  $\omega_2$  defined above. Cut both  $z$ -planes with a pair of scissors along the branch cut, i.e. along the real positive axis  $\mathbb{R}_+$ . The upper slit  $z$ -plane is then in a unique correspondence with  $\omega_1$ , and similarly the lower  $z$ -plane is in a unique correspondence with  $\omega_2$ . When we include the corresponding angle variations, we see that the two slit  $z$ -planes together represent an angle variation of  $4\pi$ . Then glue together the two half lines of the two  $z$ -planes, which represent the same angle, namely  $2\pi$ , and we obtain a connected surface *which allow an angle variation of  $4\pi$* . This gluing process of the two  $z$ -planes corresponds to that we unite the two fundamental domains  $\omega_1$  and  $\omega_2$  across the negative real axis in the  $w$ -plane, because it corresponds to the seam of angle  $2\pi$  in the two  $z$ -planes.

If we also add the real positive axis in the  $w$ -plane to this model, we see that this corresponds to gluing together the remaining two edges of the model of the Riemann surface. In the ordinary three dimensional space this is of course not possible without physically cutting the surface, but mentally we can easily abstract from this false cutting by just measuring the angle on the Riemann surface modulo  $4\pi$  instead, so we identify the two remaining edges.

We have above created a bijective correspondence  $z = w^2$  between the  $w$ -plane and the two slit  $z$ -planes by changing the domain of the angle in the lower one to  $[2\pi, 4\pi]$  and then we identify the angle  $2\pi$  in the two planes, and then more sophisticated, the angle 0 in the upper  $z$ -plane with the angle  $4\pi$  in the lower  $z$ -plane. Therefore, we can define  $w = \sqrt{z}$  as the inverse of this map.

As long as we consider both  $z$ -planes as defined above, the map  $w = \sqrt{z}$  is *single valued*. However, one usually does not solve mathematical problems by using a pair of scissors and some glue, so instead one *identifies the two slit  $z$ -planes  $f(\omega_1)$  and  $f(\omega_2)$* . The “image”  $w = \sqrt{z}$  by this identification must

be one point  $w_1 \in \omega_1$ , and at the same time another one  $w_2 \in \omega_2$ , so we usually obtain a point from each of the two chosen *fundamental domains*  $\omega_1$  and  $\omega_2$ , and  $w = \sqrt{z}$  has become *two-valued*. Only  $z = 0$  is uniquely mapped into  $w = 0$ . This corresponds to that  $w = 0$  is the *singular point*, and that  $z = 0$  is the *branching point*.

Finally,  $w = \sqrt{z}$  is already found in *Ventus, Complex Functions Theory a-1*. If  $w^2 = z = r e^{i\Theta}$ , then its two solutions are given by

$$(29) \quad w = \pm \sqrt{r} \cdot \exp\left(i \cdot \frac{\Theta}{2}\right) = \sqrt{r} \cdot \exp\left(i \cdot \frac{\Theta + 2p\pi}{2}\right), \quad p = 0, 1.$$

If in particular  $z = -1$ , then  $\sqrt{-1} = \pm i$ , which explains why we in general do not write  $\sqrt{-1}$  but use  $i$  instead for the imaginary unit.

### 2.3 The $n$ -valued function $\sqrt[n]{z}$

After this thorough investigation of the square root  $\sqrt{z}$  in Section 2.2 it is not hard to generalize the model to  $\sqrt[n]{z}$ ,  $n \in \mathbb{N} \setminus \{1\}$ , so we only sketch the main points and leave the details to the reader. It is well-known that the square root is always difficult in Mathematics. Roughly speaking, “if one has understood the square root, then the rest is easy”, as it once was put to the author by a professor.

Choosing  $\Omega = \mathbb{C}$  we define in general  $f_n : \mathbb{C} \rightarrow \mathbb{C}$  for  $n \in \mathbb{N} \setminus \{1\}$  by

$$z = f_n(w) := w^n.$$

From  $f'_n(w) = n w^{n-1}$  follows that  $w = 0$  is the only *singular point*, corresponding to the *branching point*  $z = 0$ . This is not a logarithmic branching point, and every *branch cut* must be a simple curve from 0 to  $\infty$ .

We choose the following *fundamental domains*,

$$\begin{aligned} \omega_1 &= \left\{ w = R e^{i\varphi} \mid R > 0 \text{ and } 0 < \varphi < \frac{2\pi}{n} \right\}, \\ \omega_2 &= \left\{ w = R e^{i\varphi} \mid R > 0 \text{ and } \frac{2\pi}{n} < \varphi < \frac{4\pi}{n} \right\}, \\ &\vdots \\ \omega_n &= \left\{ w = R e^{i\varphi} \mid R > 0 \text{ and } 2\pi \cdot \frac{n-1}{n} < \varphi < 2\pi \right\}. \end{aligned}$$

The *boundary curves*, i.e. the half lines given by the angles  $\varphi = j \cdot \frac{2\pi}{n}$  for each fixed  $j = 0, 1, \dots, n-1$ , are all mapped into the *branch cut*  $\mathbb{R}_+$ . In this case we get  $n$  stacked slit  $z$ -planes, each in a bijective correspondence with precisely one of the fundamental domains. When we glue together edges (branch cuts) of the  $z$ -planes representing the same angle  $2\pi j$ ,  $j = 1, \dots, n-1$  we get an  $n$ -leafed Riemann surface with 0 as its *branch point*. Finally, the remaining two edges are identified (cannot be done physically in this model without falsely cutting through the other sheets of paper), i.e. the angle is counted modulo  $2n\pi$ . In other words, we allow an angular variation of  $2n\pi$  on this Riemann surface with respect to 0. The uniquely determined inverse map from the Riemann surface to the  $w$ -plane is denoted by  $w = \sqrt[n]{z}$ .

If we identify all the  $n$  slit  $z$ -planes, we get instead an  $n$ -valued map  $\sqrt[n]{\cdot} : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \setminus \{0\}$  with its natural extension  $\sqrt[n]{0} = 0$  (notice, just one value, contrary to any other point in  $\mathbb{C} \setminus \{0\}$ ). It follows from *Ventus, Complex Functions Theory a-1*, Chapter 3, that if  $z \neq 0$ , then

$$(30) \quad w = \sqrt[n]{z} = \sqrt[n]{r} \cdot \exp\left(i \cdot \frac{\Theta + 2p\pi}{n}\right), \quad p = 0, 1, \dots, n-1,$$

where we *define* in the *positive real case*  $\sqrt[n]{r} \in \mathbb{R}_+$ , since otherwise this notation is ambiguous. We have furthermore used polar coordinates in the form  $z = r e^{i\Theta}$ .

When  $n > 2$ , there is no natural convention of sign, so one should always recall that  $\sqrt[n]{\cdot}$  in general is an  $n$ -valued function.

## 2.4 The infinite-valued function $\log z$

For  $\Omega = \mathbb{C}$  we define a function  $f : \Omega \rightarrow \mathbb{C}$  by

$$f(w) = e^w \quad (= \exp w).$$

It follows from  $f'(w) = e^w \neq 0$  for all  $w \in \mathbb{C}$  that this map has *no singular points*. It was proved in *Ventus, Complex Functions Theory a-1*, Chapter 4, that  $f(\mathbb{C}) = \mathbb{C} \setminus \{0\}$ . Clearly,  $z_0 = 0$  is an isolated point of  $\mathbb{C} \setminus f(\mathbb{C}) = \{0\}$ , hence  $z_0 = 0$  is a *logarithmic branch point*.

According to *Ventus, Complex Functions Theory a-1*, Chapter 4, its many-valued inverse map  $\log \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}$  is given by

$$(31) \quad w = \log z := \text{Log } z + 2ip\pi, \quad p \in \mathbb{Z}.$$

By definition, the *principal logarithm* is defined by

$$(32) \quad \text{Log } z = \ln |z| + i \cdot \text{Arg } z \quad \text{for } z \in \mathbb{C} \setminus \{0\},$$

where the *principal argument* is uniquely determined by  $\text{Arg } z \in ]-\pi, \pi]$ . By varying  $p \in \mathbb{Z}$  in (31) it therefore follows that we may choose the *fundamental domains*  $\omega_p$ ,  $p \in \mathbb{Z}$ , as the infinitely many strips

$$(33) \quad \omega_p = \{w \in \mathbb{C} \mid \Im w \in ](2p - 1)\pi, (2p + 1)\pi[ \}, \quad p \in \mathbb{Z}.$$

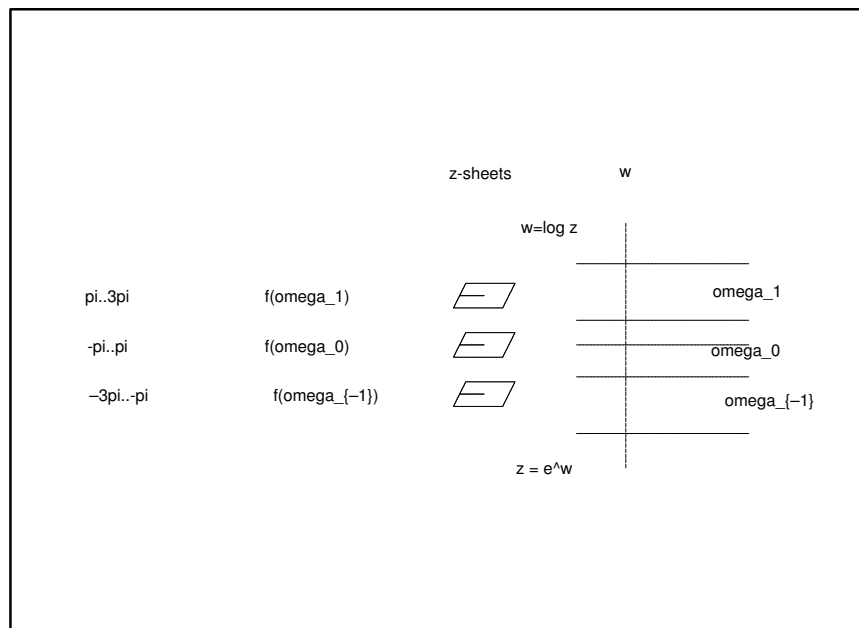


Figure 10: Construction of the Riemann surface of  $w = \log z$  with a logarithmic branch point.

The lines  $\Im w = (2p + 1)\pi$  for  $p \in \mathbb{Z}$  are all mapped into the *branch cut* by the map

$$f(w) = e^w = e^{\Re w} \cdot e^{i(2p+1)\pi} = -e^{\Re w},$$

i.e. into the real negative half axis  $\mathbb{R}_-$ . Then each of the fundamental domains  $\omega_j$  is in a bijective correspondence with the slit  $z$ -plane  $f(\omega_j)$  of number  $j \in \mathbb{Z}$ , so the inverse map,  $\log$ , maps the infinite-leaved Riemann surface bijectively onto the  $w$ -plane. Notice that the *logarithmic branch point*  $z_0 = 0$  does *not* belong to the Riemann surface.

If we identify all the slit  $z$ -planes, we get instead the *infinite-valued function*  $\log$  given by (31). We introduced in *Ventus, Complex Functions Theory a-1*, Chapter 4, the many-valued argument function by

$$\arg z = \{\text{Arg } z + 2p\pi \mid p \in \mathbb{Z}\} \quad \text{for } z \in \mathbb{C} \setminus \{0\}$$

(and not defined for  $z = 0$ ). Then it follows from (31) and (32) that

$$(34) \quad \log z = \ln |z| + i \cdot \arg z, \quad z \in \mathbb{C} \setminus \{0\}.$$

**Remark 2.4.1** Although there are many similarities between the real logarithm,  $\ln : \mathbb{R}_+ \rightarrow \mathbb{R}$ , and the many-valued logarithm  $\log$  above, defined on  $\mathbb{C} \setminus \{0\}$ , not all rules of computations for  $\ln$  can be generalized to  $\log$ . As an example of what may go wrong, choose  $z = 1 + i$ . Then

$$|z| = \sqrt{2} \quad \text{and} \quad \arg z = \left\{ \frac{\pi}{4} + 2p\pi \mid p \in \mathbb{Z} \right\},$$

thus

$$2 \log(1 + i) = 2 \left\{ \frac{1}{2} \ln 2 + i \left( \frac{\pi}{4} + 2p\pi \right) \right\} = \ln 2 + i \frac{\pi}{2} + 4ip\pi, \quad \text{for } p \in \mathbb{Z},$$

so  $2 \log(1 + i)$  is determined *modulo*  $4\pi i$ .

For comparison,  $z^2 = (1 + i)^2 = 2i$ , where

$$|z^2| = 2 \quad \text{and} \quad \arg(z^2) = \left\{ \frac{\pi}{2} + 2ip\pi \mid p \in \mathbb{Z} \right\},$$

thus

$$\log(\{1 + i\}^2) = \left\{ \ln 2 + i \frac{\pi}{2} + 2ip \mid p \in \mathbb{Z} \right\},$$

and the two sets  $\log(\{1 + i\}^2)$  and  $2 \log(1 + i)$  are *not* identical. Cf. also an example in *Ventus, Complex Functions Theory a-1*, Chapter 1.  $\diamond$

Finally, we shall compare the present three types of logarithms,

$$\ln : \mathbb{R}_+ \rightarrow \mathbb{R}, \quad \text{Log} : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \quad \text{and} \quad \log : \mathbb{C} \setminus \{0\} \rightarrow \mathcal{R},$$

where  $\mathcal{R}$  denotes the corresponding Riemann surface.

Here,  $\ln$  is the usual *real* logarithm, known from real Calculus, and the *principal logarithm*  $\text{Log}$  is a natural extension of  $\ln$ , i.e.

$$\text{Log } z := \ln |z| + i \cdot \text{Arg } z, \quad z \in \mathbb{C} \setminus \{0\}, \quad \text{Arg } z \in ] -\pi, \pi[$$

while  $\log$  is interpreted as an *infinite-valued function*, when all the sheets of the Riemann surface  $\mathcal{R}$  are identified.

In the presentation above, neither of the three logarithms is analytic. The real logarithm is not, because  $\mathbb{R}_+ \subset \mathbb{C}$  does not contain interior points from  $\mathbb{C}$ . The principal logarithm  $\text{Log}$  is discontinuous on  $\mathbb{R}_-$ , so it cannot be analytic either. However, if we remove  $\mathbb{R}_-$  from the domain of  $\text{Log}$ , then this restriction of  $\text{Log}$  becomes analytic. Finally, it does not make sense to call analytic a many-valued function, though it is indeed true that *each of its branches is analytic*. One obvious example is

$$\text{Log}_p z := \text{Log } z + 2ip\pi, \quad z \in \mathbb{C} \setminus (\mathbb{R}_- \cup \{0\}), \quad p \in \mathbb{Z},$$

where in particular  $\text{Log}_0 = \text{Log}$ .

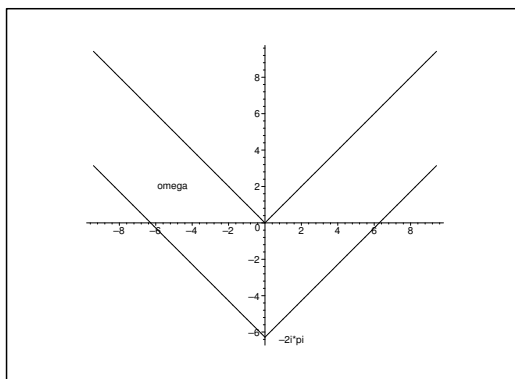


Figure 11: Example of a simple, though unusual fundamental domain  $\omega$  of the logarithm.



The *fundamental domains* need not be as nice as indicated above. For instance, the domain  $\omega$  on Figure 11 is a fundamental domain, because it is mapped bijectively by the map  $f(w) = e^w$  onto a slit  $z$ -plane, in which the *branch cut* is composed of the two differentiable curves

$$e^{(-1+i)t}, \quad t \in \mathbb{R}_+, \quad \text{and} \quad e^{(1+i)t}, \quad t \in \mathbb{R}_+ \cup \{0\},$$

which can also be written

$$e^t \left\{ \cos t + i \cdot \frac{t}{|t|} \cdot \sin t \right\}, \quad t \in \mathbb{R} \setminus \{0\},$$

supplied with 1 for  $t = 0$ . The latter example shows that the fundamental domains should be chosen with care, no matter that we have a large freedom in defining them.

## 2.5 The many-valued functions $a^z$ and $z^a$

Using the infinite-valued logarithm we can define the *general exponentials*  $w = a^z$  and the *general power functions*  $w = z^a$ . Their inverse functions  $z = f(w)$  are usually not analytic functions, so there does not exist a natural Riemann surface describing the general exponentials or the general power functions. It can be proved that they are usually described by a bijective correspondence between *two* Riemann surfaces, one above the  $z$ -plane, and another one above the  $w$ -plane. This implies of course that we may expect some very strange phenomena, which go against one's intuition. In other words: *Trust only the definitions, when these multi-valued functions are considered!*

**Remark 2.5.1** According to unsubstantiated folklore the first person ever to consider these functions (in the seventeenth century) became a lunatic, because he spent many years in vain to interpret these strange results, and he ended his days in a mental hospital.  $\diamond$

**Definition 2.5.1** Let  $a \in \mathbb{C} \setminus \{0, e\}$  be a fixed number. We define the general exponential  $a^z$  as the many-valued function

$$(35) \quad a^z := \exp(z \log a).$$

It follows from this definition that it is strictly speaking not correct to write  $e^z$  for  $\exp z$ . However, the notation  $e^z$  is often more convenient to use than  $\exp z$ , so it has been convention to write  $e^z$  in this sense. For that reason only,  $a = e$  has been excluded from Definition 2.5.1. One may of course also use  $e^z$  in the sense of Definition 2.5.1, but then this should be explicitly noted in order to avoid confusion.

On the other hand, the exclusion of  $a = 0$  is very natural, because  $\log 0$  does not make any sense at all in (35).

We notice that  $a^z$  actually is a symbol containing infinitely many unique functions. This is seen in the following way: Since  $\log a$  is given by the set

$$\log a = \{\text{Log } a + 2ip\pi \mid p \in \mathbb{Z}\}, \quad a \in \mathbb{C} \setminus \{0, e\},$$

we may for each fixed  $p \in \mathbb{Z}$  define the uniquely determined value

$$\text{Log}_p a := \text{Log } a + 2ip\pi,$$

and we get the uniquely determined analytic function

$$(a^z)_p := \exp(z \operatorname{Log} a + 2ip\pi z), \quad z \in \mathbb{C},$$

the derivative of which is given by

$$\frac{d}{dz} (a^z)_p = \operatorname{Log}_p \cdot (a^z)_p, \quad \text{for fixed } p \in \mathbb{Z}.$$

We define the *principal value* of  $a^z$  by choosing  $p = 0$ , i.e.

$$(a^z)_0 := \exp(z \operatorname{Log} a),$$

corresponding to the *principal logarithm*  $\operatorname{Log}$ .

If we let  $z = 0$  in (35), we get  $a^0 = 1$ , so the infinite many unique functions  $(a^z)_p$ ,  $p \in \mathbb{Z}$ , making up the set  $a^z$ , all coincide for  $z = 0$ .

For practical computations we note the following: If we put  $\Theta = \operatorname{Arg} z$ , then

$$\log a = \ln a = \ln |a| + i(\Theta + 2p\pi), \quad p \in \mathbb{Z}.$$

Hence, for  $z = x + iy \neq 0$ ,

$$\begin{aligned} a^z &:= \exp(z \log a) = \exp(\{x + iy\} \cdot \{\ln |a| + i(\Theta + 2p\pi)\}) \\ (36) \quad &= \exp(\{x \ln |a| - y(\Theta + 2p\pi)\} + i\{x(\Theta + 2p\pi) + y \ln |a|\}), \quad p \in \mathbb{Z}. \end{aligned}$$

**Example 2.5.1** Choose  $a = 1$ . Then  $\log 1 = 2ip\pi$ ,  $p \in \mathbb{Z}$ , so  $\ln |a| = 0$  and  $\Theta = 0$ . It follows from (36) that

$$\begin{aligned} 1^z &= \exp(\{x \cdot 0 - y \cdot 2p\pi\} + i\{x \cdot 2p\pi + y \cdot 0\}) \\ &= e^{-2py} \cdot \{\cos(2xp\pi) + i \sin(2xp\pi)\}, \quad p \in \mathbb{Z}. \end{aligned}$$

Choose  $y = 0$ , thus  $z = x \in \mathbb{R}$ . Then

$$1^x = \cos(2xp\pi) + i \cdot \sin(2xp\pi) = e^{2ixp\pi}, \quad p \in \mathbb{Z}.$$

If  $x \in \mathbb{R} \setminus \mathbb{Q}$  is irrational, it can be proved that  $1^x = \{e^{2ixp\pi} \mid p \in \mathbb{Z}\}$  is a *complex* point set, which is dense in the unit circle.

If  $x = \frac{1}{n}$  for some  $n \in \mathbb{N}$ , we get the usual  $n$  roots of the binomial equation  $w^n = 1$ , symbolized by

$$(37) \quad \sqrt[n]{1} = \left\{ \exp\left(\frac{2ip\pi}{n}\right) \mid p = 0, \dots, n-1 \right\}.$$

This shows that the previous theory given in *Ventus, Complex Functions Theory a-1* is in agreement with the present extension.  $\diamond$

**Example 2.5.2** It is not so difficult to accept the result of Example 2.5.1, because formula (37) is in agreement with a previously known formula. However, the result becomes less understandable, if we choose  $a = z = i$ , which we shall do now. Then we get by the *definition*,

$$i^i := \exp(i \cdot \log i) = \exp(i \cdot \{0 + 2ip\pi\}) = \exp(-2p\pi), \quad p \in \mathbb{Z},$$

so the symbol  $i^i$  described only by the imaginary number  $i$  is actually a set of *real positive* numbers,

$$i^i = \{e^{2n\pi} \mid n \in \mathbb{Z}\},$$

where we have put  $n = -p \in \mathbb{Z}$  for convenience.  $\diamond$

Analogously to Definition 2.5.1 we define

**Definition 2.5.2** Fix  $a \in \mathbb{C}$ . The general power function  $z^a$  is defined on the set  $\mathbb{C} \setminus \{0\}$  by

$$(38) \quad z^a := \exp(a \log z), \quad z \in \mathbb{C} \setminus \{0\}.$$

To every branch  $\text{Log}_p z$  of  $\log z$  then corresponds a branch  $(z^a)_p$  of  $z^a$ , and a small computation gives

$$\frac{d}{dz} (z^a)_p = \frac{a}{z} \cdot (z^a)_p.$$

The branch corresponding to the principal logarithm  $\text{Log } z$  and the slit domain  $\Omega := \mathbb{C} \setminus (\mathbb{R}_- \cup \{0\})$  is called the *principal value* of  $z^a$ .

When  $a = n \in \mathbb{Z}$ , there is nothing new in (38).

When  $a \in \mathbb{Q}$  we are still in agreement with the previous theory, cf. also Example 2.5.1.

Finally, when  $z \in \mathbb{C} \setminus \mathbb{Q}$ , then every reasonable geometric interpretation of  $z$  totally disappears.

**Example 2.5.3** If we choose  $a = i$  and  $z \neq 0$ , then it follows from (38) that

$$\begin{aligned} z^i &:= \exp(i\{\ln |z| + i(\text{Arg } z - 2p\pi)\}) \\ &= e^{-\text{Arg } z + 2p\pi} \cdot \{\cos(\ln |z|) + i \sin(\ln |z|)\}, \quad p \in \mathbb{Z}, \end{aligned}$$

where we have used that  $\arg z = \{\text{Arg } z - 2p\pi \mid p \in \mathbb{Z}\}$ .  $\diamond$

## 2.6 The many-valued function $\arcsin z$

We shall leave the unpleasant many-valued functions  $a^z$  and  $z^a$  and instead proceed with the inverses of the trigonometric and hyperbolic functions. We shall in this section give a full treatment of the inverse of  $f(w) = \sin w$ . The other inverses are only sketched in the following section, because the reader may easily find them herself/himself by copying the method given here.

We consider  $z = f(w) = \sin w$ , where  $\Omega = \mathbb{C}$ . Since

$$f'(w) = \cos w = 0 \quad \text{for } w = \frac{\pi}{2} + p\pi, \quad p \in \mathbb{Z},$$

cf. *Ventus, Complex Functions Theory a-1*, it follows that the *singular points* form the set

$$\left\{ \frac{\pi}{2} + p\pi \mid p \in \mathbb{Z} \right\}, \quad \text{where } \left\{ \sin \left( \frac{\pi}{2} + p\pi \right) \mid p \in \mathbb{Z} \right\} = \{-1, 1\},$$

showing that every singular point by sine is mapped into one of the two *branching points*  $\{-1, 1\}$ .

We shall prove that in this case we do not have any (finite) *logarithmic branching point*. It suffices to prove that  $f(\mathbb{C}) = \sin(\mathbb{C}) = \mathbb{C}$ , or, put in another way, that the equation  $f(w) = \sin w = z$  has at least one solution for every  $z \in \mathbb{C}$ . It follows from the definition of  $\sin w$  in *Ventus, Complex Functions Theory a-1* that we shall only solve the equation

$$\sin w = \frac{1}{2i} (e^{iw} - e^{-iw}) = z.$$

Since  $2ie^{iw} \neq 0$  for every  $w \in \mathbb{C}$ , this equation is equivalent to

$$(39) \quad (e^{iw})^2 - 2ize^{iw} - 1 = 0,$$

which is an equation of second order in the unknown variable  $e^{iw}$ , where we already have derived a solution formula in *Ventus, Complex Functions Theory a-1*. Hence it is straightforward to write down its solutions,

$$(40) \quad e^{iw} = iz \pm \sqrt{(iz)^2 + 1} = iz \pm \sqrt{1 - z^2}.$$

It is well-known that the product of the solutions of a second order equation is equal to the constant term, hence the products of the solutions (40) of the equation (39) is  $-1$ . In particular, none of the solutions (40) is 0 for any  $z \in \mathbb{C}$ . Therefore, we can take the (complex) logarithm of (40), from which we derive the solution formula

$$(41) \quad \boxed{w := \arcsin z = \frac{1}{i} \log \left( iz \pm \sqrt{1 - z^2} \right) \quad \text{for every } z \in \mathbb{C}.}$$

It follows in particular from (41) that arcsin does not have a logarithmic branching point (in spite of the fact that log enters (41)). Furthermore, (41) is an exact formula, in which the square root for  $z \neq \pm 1$  has two values — here indicated by  $\pm$  though this sign in principal is superfluous — and log is the well-known infinite-valued complex logarithm.

The most obvious choice of *fundamental domains* of the logarithm are as shown in Section 2.4 horizontal strips of height  $2\pi$ . Due to the factor  $\frac{1}{i}$  in (41) we first *guess* that the fundamental domains of  $\arcsin z$  may be chosen as *vertical strips* of width  $2\pi$ . We immediately see, however, that this cannot be the whole story, because the square root  $\pm\sqrt{1 - z^2}$  for  $z \neq \pm 1$  is 2-valued.

We compensate for the square root being 2-valued by choosing in our next guess of the fundamental domains the vertical strips to be of width  $\pi$ . Since the ordinary *real* function  $\text{Arcsin}: [-1, 1] \rightarrow \mathbb{R}$  has the range  $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ , we are led to consider the domains

$$\omega_p = \left\{ z \in \mathbb{C} \mid \Re z \in \left] -\frac{\pi}{2} + p\pi, \frac{\pi}{2} + p\pi \right[ \right\}, \quad p \in \mathbb{Z}.$$

First consider the case  $p = 0$ , i.e. the vertical strip

$$\omega_0 = \left\{ z \in \mathbb{C} \mid \Re z \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ \right\}.$$

We shall prove that  $\sin \omega_0$  is the complex plane  $\mathbb{C}$  with the exception of the branch cuts, which are the images by sine of the vertical lines  $\Re z = \pm \frac{\pi}{2}$ .

It follows from

$$z = x + iy = \sin w = \sin u \cdot \cosh v + i \cos u \cdot \sinh v$$

that

$$x = \sin u \cdot \cosh v \quad \text{and} \quad y = \cos u \cdot \sinh v.$$

The curve (actually an interval on the real axis)

$$(u(t), v(t)) = (t, 0), \quad \text{for } t \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ ,$$

is then mapped onto

$$(x(t), y(t)) = (\sin t, 0) \quad \text{for } t \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ ,$$

i.e. into the interval  $] -1, 1[$  on the real axis in  $\mathbb{C}$ , traversed from  $-1$  to  $+1$ .

Let  $a \in \mathbb{R} \setminus \{0\}$ . Then the curve (a horizontal line segment)

$$(42) \quad (u(t), v(t)) = (t, a) \quad \text{for } t \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ ,$$

is mapped onto the curve

$$(43) \quad (x(t), y(t)) = (\cosh a \cdot \sin t, \sinh a \cdot \cos t).$$

Since  $\cosh a \neq 0$  and  $\sinh a \neq 0$ , we get

$$\sin t = \frac{x(t)}{\cosh a} \quad \text{and} \quad \cos t = \frac{y(t)}{\sinh a},$$

hence

$$(44) \quad \left( \frac{x}{\cosh a} \right)^2 + \left( \frac{y}{\sinh a} \right)^2 = 1.$$

Now,  $y(t) = \sinh a \cdot \cos t$ ,  $t \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ , has always the same sign as  $a$ , hence the line segment (42) is mapped into a half ellipse of equation (44), and every half ellipse starts from the *branch cut*  $]-\infty, -1[$  in the  $z$ -plane and ends at the *branch cut*  $]1, +\infty[$ , cf. the curves  $A$  and  $C$  on Figure 12.

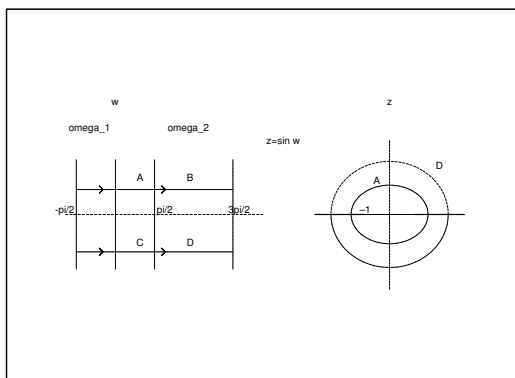


Figure 12: Geometric analysis of a fundamental domain of  $\arcsin z$ .

If we put  $\alpha = \sinh^2 a > 0$ , then  $\cosh^2 a = 1 + \sinh^2 a = 1 + \alpha$ , and (44) is written

$$(45) \quad \frac{x^2}{1 + \alpha} + \frac{y^2}{\alpha} = 1 \quad \text{for } a \neq 0.$$

Therefore, if  $(x, y)$  is any point in the  $z$ -plane for which  $y \neq 0$ , then we shall only prove that  $(x, y)$  lies on precisely one of these ellipsoidal arcs.

We define for given  $(x, y) \in \mathbb{R}^2$ , where  $y \neq 0$ , a function  $\varphi$  on  $\mathbb{R}_+$  by

$$\varphi(\alpha) := \frac{x^2}{1 + \alpha} + \frac{y^2}{\alpha}, \quad \text{for } \alpha \in \mathbb{R}_+.$$

Then  $\varphi'(\alpha) < 0$ , so  $\varphi$  is strictly decreasing, and since  $\lim_{\alpha \rightarrow 0^+} \varphi(\alpha) = +\infty$  and  $\lim_{\alpha \rightarrow +\infty} \varphi(\alpha) = 0$ , there exists precisely one  $\alpha > 0$ , such that (45) is fulfilled.

Then we use that  $\sinh^2 a = \alpha$  and that  $\sinh a$  and  $y$  always have the same sign, so we conclude that

$$\sinh a = \frac{y\sqrt{\alpha}}{|y|}, \quad \alpha > 0 \text{ and } y \neq 0,$$

and it follows that the point  $(x, y)$  lies on a uniquely determined ellipsoidal arc from our system. We therefore conclude that  $\omega_0$  is a *fundamental domain*, because two different points on the same ellipsoidal arc are mapped into two different points on the corresponding line segment in  $\omega_0$ .

The domain  $\omega_{2n} = \omega_0 + 2n\pi$  is characterized by  $\Re z \in \left] -\frac{\pi}{2} + 2n\pi, \frac{\pi}{2} + 2n\pi \right[$ , so by using the periodicity of the trigonometric functions occurring in (43) we conclude that we by the same analysis as above for every  $\omega_{2m}$ ,  $n \in \mathbb{Z}$ , obtain the same ellipsoidal arcs, so every  $\omega_{2n}$ ,  $n \in \mathbb{Z}$ , is also a fundamental domain.

Then consider  $\omega_1 = \left\{ z \in \mathbb{C} \mid \left] \frac{\pi}{2}, \frac{3\pi}{2} \right[ \right\}$ . In this case the horizontal curves given by the parametric description

$$(u(t), v(t)) = (t, a), \quad t \in \left] \frac{\pi}{2}, \frac{3\pi}{2} \right[ ,$$

are mapped onto

$$\begin{aligned} (x(t), y(t)) &= (\cosh a \cdot \sin t, \sinh a \cdot \cos t), & t \in \left] \frac{\pi}{2}, \frac{3\pi}{2} \right[ , \\ &= -(\cosh a \cdot \sin \tau, \sinh a \cdot \cos \tau), & \tau \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ , \end{aligned}$$

where  $t = \tau + \pi$ , cf. also (42) and (43). The change of sign implies that the curves (43) are reflected with respect to the point  $(0, 0)$ . This is indicated on Figure 12 by the curves called *B* and *D*.

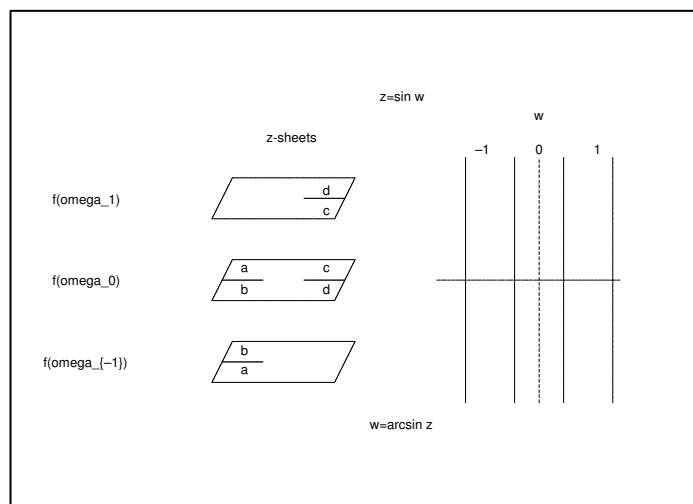


Figure 13: Construction of the Riemann surface of  $\arcsin z$ .

Then stack all the fundamental domains above each other and piece them together as indicated on Figure 13, i.e. the *branch cuts* called  $a$  and  $c$  are glued together, and the same is done *mentally* for the branch cuts called  $b$  and  $d$  without cutting the previously glued branch cuts  $a$  and  $c$ . This is of course physically impossible in the three dimensional space, but we deduce at least that we *locally* must have a square root structure in the neighbourhood of each of the two *branching points*  $-1$  and  $1$ . It follows by the periodicity that  $f(\omega_{2n-1})$  and  $f(\omega_{2n})$ ,  $n \in \mathbb{Z}$ , are glued together as a square root along the branch cut  $] -\infty, -1[$  on the real axis (in the complex plane), and  $f(\omega_{2n})$  and  $f(\omega_{2n+1})$ ,  $n \in \mathbb{Z}$ , are glued together along the other branch cut  $]1, +\infty[$  on the real axis.

Summing up we obtain the Riemann surface of  $w = \arcsin z$  as an infinite stack of  $z$ -planes  $f(\omega_n)$ ,  $n \in \mathbb{Z}$ , where each of these is connected with the previous one  $f(\omega_{n-1})$  by a square root structure around one of the branching points, and is connected with the following one  $f(\omega_{n+1})$  by a square root structure around the other branching point, so the Riemann surface has infinitely many sheets.

If all the slit  $z$ -planes are identified, i.e. if we project the Riemann surface onto the underlying  $z$ -plane  $C$ , then we get the infinite-valued function given by (41).



Let us choose in the slit  $z$ -plane that branch of the square root in (41) which is defined by

$$\sqrt{1-z^2} := \exp\left(\frac{1}{2} \operatorname{Log}(1-z^2)\right) \quad \text{for } z \in \mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[),$$

where  $\operatorname{Log}$  denotes the *principal logarithm*. Then we define the *principal value*  $\operatorname{Arcsin} z$  of  $\arcsin z$  by

$$(46) \quad \operatorname{Arcsin} z := \frac{1}{i} \operatorname{Log}\left(iz + \exp\left(\frac{1}{2} \operatorname{Log}(1-z^2)\right)\right), \quad \text{for } z \in \mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[).$$

Clearly,  $\operatorname{Arcsin} z$  is an analytic branch of  $\arcsin z$ , i.e.  $\operatorname{Arcsin} z$  is an analytic function in the open domain  $\mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[)$ . In practice one extends  $\operatorname{Arcsin} z$  to the points  $\{-1, 1\}$ , using that  $\sqrt{1-(\pm 1)^2} = 0$ . This extension is continuous, though not analytic at  $\{-1, 1\}$ , because  $\{-1, 1\}$  are not interior points. We check by the following computation

$$\operatorname{Arcsin}(\pm 1) = \frac{1}{i} \operatorname{Log}(\pm i + 0) = \frac{1}{i} \left\{ \pm i \frac{\pi}{2} \right\} = \pm \frac{\pi}{2},$$

that this continuous extension agrees with the familiar definition of  $\operatorname{Arcsin}(\pm 1)$  known from *Real Calculus*.

More generally we show that if  $z = t \in ]-1, 1[$ , then (46) agrees with the definition of  $\operatorname{Arcsin}_{\mathbb{R}} t$  given in *Real Calculus*. We write for convenience  $\operatorname{Arcsin}_{\mathbb{C}} z$  for the function defined by (46) in the following proof.

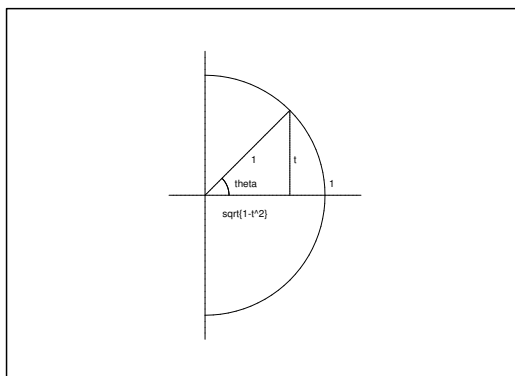


Figure 14: Geometry of  $\operatorname{Arcsin} t$  for  $t \in ]-1, 1[$ .

By the chosen branch above of the square root we first see that  $\sqrt{1-t^2} > 0$  for  $t \in ]-1, 1[$ . Furthermore,

$$\left| it + \sqrt{1-t^2} \right| = \sqrt{t^2 + (1-t^2)} = 1,$$

cf. Figure 14, where  $\Theta = \operatorname{Arcsin}_{\mathbb{R}} t$  in the *real* sense. Then we get the following computation using the definition of  $\operatorname{Log} z = \ln |z| + i \operatorname{Arg} z$ ,

$$\operatorname{Arcsin}_{\mathbb{C}} t = \frac{1}{i} \operatorname{Log}\left(it + \sqrt{1-t^2}\right) = \frac{1}{i} \left\{ \ln 1 + i \operatorname{Arg}\left(\sqrt{1-t^2} + it\right) \right\} = \Theta = \operatorname{Arcsin}_{\mathbb{R}} t,$$

and we have proved that  $\operatorname{Arcsin}_{\mathbb{C}} t = \operatorname{Arcsin}_{\mathbb{R}} t$  for  $t \in ]-1, 1[$ .

## 2.7 The many-valued arcus functions and area functions

In this section we shall shortly without proofs, which are left to the reader, sketch the most important results on the arcus functions and the area functions. The name “arcus function” is due to the fact that they are geometrically connected with arcs, so they all have the prefix “arc” or “Arc”. The area functions are similarly geometrically connected with areas, so they have the prefix “ar” or “Ar”, supplied with the suffix “h” instead, and *not* “arc” or “Arc”, which is sometimes seen in texts.

For completeness we start with  $\arcsin z$ , which has already been treated in more details in the previous Section 2.6.

### 2.7.1 The infinite-valued arcus function $\arcsin z$

We derived in Section 2.6 that the inverse of  $z = \sin w$  is given by the infinite-valued function

$$(47) \quad \arcsin z = \frac{1}{i} \log \left( iz \pm \sqrt{1 - z^2} \right), \quad z \in \mathbb{C},$$

with the two *branching points*  $z_1 = -1$  and  $z_2 = +1$ .

The *principal value*  $\operatorname{Arcsin} z$  is the analytic function

$$(48) \quad \operatorname{Arcsin} z = \frac{1}{i} \operatorname{Log} \left( iz + \exp \left( \frac{1}{2} \operatorname{Log} (1 - z^2) \right) \right), \quad z \in \Omega = \mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[).$$

If  $z = t \in ]-1, 1[$ , then the complex  $\operatorname{Arcsin} t$  is equal to  $\operatorname{Arcsin}_{\mathbb{R}} t$  known from *Real Calculus*, and it can be extended by continuity to the branching points  $z_1 = -1$  and  $z_2 = +1$ .

### 2.7.2 The infinite-valued arcus function $\arccos z$

The inverse function of  $z = \cos w$  is given by the infinite-valued function

$$(49) \quad \arccos z = \frac{1}{i} \log \left( z \pm i \sqrt{1 - z^2} \right), \quad z \in \mathbb{C},$$

with the two *branching points*  $z_1 = -1$  and  $z_2 = +1$ .

The *principal value*  $\operatorname{Arccos} z$  is the analytic function

$$(50) \quad \operatorname{Arccos} z = \frac{1}{i} \operatorname{Log} \left( z + i \exp \left( \frac{1}{2} \operatorname{Log} (1 - z^2) \right) \right), \quad z \in \Omega = \mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[).$$

If  $z = t \in ]-1, 1[$ , then the complex  $\operatorname{Arccos} t$  is equal to  $\operatorname{Arccos}_{\mathbb{R}} t$  known from *Real Calculus*, and it can be extended by continuity to the branching points  $z_1 = -1$  and  $z_2 = +1$ .

### 2.7.3 The infinite-valued arcus function $\arctan z$

The inverse function of  $z = \tan w$  for  $w \neq \frac{\pi}{2} + p\pi$ ,  $p \in \mathbb{Z}$ , is given by the infinite-valued function

$$(51) \quad \arctan z = \frac{1}{2i} \log \left( \frac{1 + iz}{1 - iz} \right) = \frac{1}{2i} \log \left( \frac{i - z}{i + z} \right), \quad z \in \mathbb{C} \setminus \{-i, i\},$$

which has the two *logarithmic branching points*  $z_1 = -i$  and  $z_2 = i$ .

The *principal value*  $\operatorname{Arctan} z$  is given by

$$(52) \quad \boxed{\operatorname{Arctan} z = \frac{1}{2i} \operatorname{Log} \left( \frac{1+iz}{1-iz} \right) = \frac{1}{2i} \operatorname{Log} \left( \frac{i-z}{i+z} \right), \quad z \in \Omega = \mathbb{C} \setminus \{iy \mid y \in \mathbb{R}, |y| \geq 1\},}$$

and when  $z = t \in \mathbb{R}$ , this complex defined function is identical with the familiar  $\operatorname{Arctan}_{\mathbb{R}} t$ , known from *Real Calculus*.

The *fundamental domain* of (52) is a slit complex plane with the *branch cuts* along the imaginary axes, i.e. along  $\{iy \mid y \geq 1\}$  and  $\{iy \mid y \leq -1\}$ . Since  $\pm i$  are *logarithmic branching points*, the Riemann surface becomes easier to construct than the Riemann surfaces of  $\arcsin z$  or  $\arccos z$ , because there will only be one way to connect all the infinitely many fundamental domains.

#### 2.7.4 The infinite-valued arcus function $\operatorname{arccot} z$

The inverse function of  $z = \cot w$  for  $w \neq p\pi$ ,  $p \in \mathbb{Z}$ , is given by the infinite-valued function

$$(53) \quad \boxed{\operatorname{arccot} z = \frac{1}{2i} \log \left( \frac{z+i}{z-i} \right), \quad z \in \mathbb{C} \setminus \{-i, i\},}$$

which has the two *logarithmic branching points*  $z_1 = -i$  and  $z_2 = i$ .

The *principal value*  $\operatorname{Arccot} z$  is given by

$$(54) \quad \boxed{\operatorname{Arccot} z = \frac{1}{2i} \operatorname{Log} \left( \frac{z+i}{z-i} \right), \quad z \in \Omega = \mathbb{C} \setminus \{iy \mid y \in \mathbb{R}, |y| \leq 1\}.}$$

It does *not* agree with the real  $\operatorname{Arccot}_{\mathbb{R}} t$  for  $t \in \mathbb{R}$ , because the branch cut of (54) is the interval “ $[-i, i]$ ” on the imaginary axis, which cuts through the real axis.

It is, however, possible to find another branch of  $\operatorname{arccot} z$  which is equal to  $\operatorname{Arccot}_{\mathbb{R}} t$  for  $t \in \mathbb{R}$ . It is defined by

$$(55) \quad \boxed{\operatorname{Arccot}_0 z = \frac{\pi}{2} - \operatorname{Arctan} z = \frac{\pi}{2} - \frac{1}{2i} \operatorname{Log} \left( \frac{i-z}{i+z} \right), \quad z \in \Omega = \mathbb{C} \setminus \{iy \mid y \in \mathbb{R}, |y| \geq 1\}.}$$

#### 2.7.5 The infinite-valued area function $\operatorname{arsinh} z$

The inverse function of  $z = \sinh w$  is given by the infinite-valued function

$$(56) \quad \boxed{\operatorname{arsinh} z = \log \left( z \pm \sqrt{z^2 + 1} \right), \quad z \in \mathbb{C}.}$$

Its two *branching points* are  $z_1 = -i$  and  $z_2 = i$ .

Its *principal value* is given by

$$(57) \quad \boxed{\operatorname{Arsinh} z = \operatorname{Log} \left( z + \exp \left( \frac{1}{2} \operatorname{Log}(z^2 + 1) \right) \right)}, \quad z \in \Omega = \mathbb{C} \setminus \{iy \mid y \in \mathbb{R}, |y| \geq 1\}.$$

When  $z = t \in \mathbb{R}$ , then (57) is equal to the usual real  $\operatorname{Arsinh}_{\mathbb{R}} t$ .

### 2.7.6 The infinite-valued area function $\operatorname{arcosh} z$

The inverse function of  $z = \cosh w$  is given by the infinite-valued function

$$(58) \quad \boxed{\operatorname{arcosh} z = \log \left( z \pm \sqrt{z^2 - 1} \right)}, \quad z \in \mathbb{C}.$$

Its *branching points* are  $z_1 = -1$  and  $z_2 = 1$ .

The *principal value*  $\operatorname{Arcosh} z$  is tricky, because it is defined by

$$(59) \quad \boxed{\operatorname{Arcosh} z = \operatorname{Log} \left( z + \exp \left( \frac{1}{2} \operatorname{Log}(z^2 - 1) \right) \right)}.$$

This definition requires that  $z$  must not

- lie on the imaginary axis, in which case  $z^2 - 1$  does not lie in the domain of Log,
- or in the real interval  $] - 1, 1[$ , in which case  $z^2 - 1$  again becomes negative and hence not in the domain of Log,
- or  $z = x$  must not lie in the real interval  $] - \infty, -1[$ , in which case  $x + \exp\left(\frac{1}{2} \operatorname{Log}(x^2 - 1)\right) < 0$ , and therefore not in the domain of Log,
- or  $z \neq \pm 1$ , because  $\operatorname{Log} 0$  is not defined.

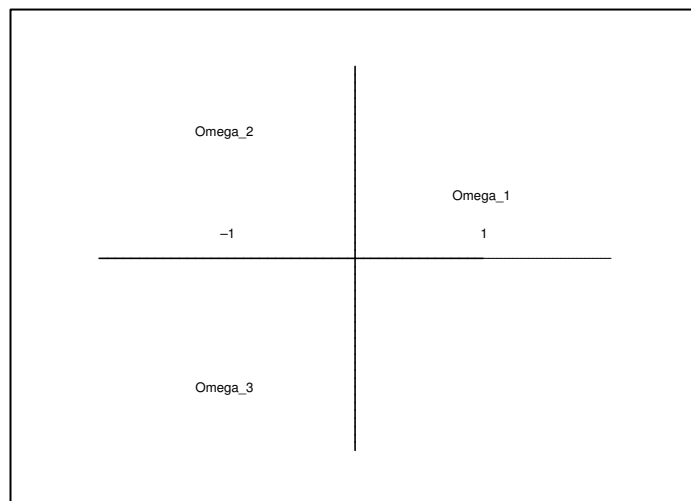


Figure 15: Domain of the principal value  $\operatorname{Arcosh} z$ .

According to these exclusions we conclude that the domain of  $\operatorname{Arcosh} z$  is the union of the three open connected domains  $\Omega_1$ ,  $\Omega_2$  and  $\Omega_3$  on Figure 15.

If we choose  $\Omega_1$  as our domain of (59), then it can be proved that for  $z = t \in ]1, +\infty[$  the function given by (59) is identical with the *real*  $\operatorname{Arcosh}_{\mathbb{R}} t$ .

Of course this is not true for  $\Omega_2$  and  $\Omega_3$  for the simple reason that the interval  $]1, +\infty$  does not lie in these two connected domains.

The above may seem strange, because  $\Omega_1$  does not correspond to a *fundamental domain* in the  $w$ -plane. The reason of this unexpected analysis is that we use (58) as our definition and then apply the somewhat rigid definition of the *principal logarithm*. Let us look closer at the process. The method is to solve the equation

$$z = \cosh w = \frac{1}{2} \{e^w + e^{-w}\}$$

with respect to  $w$ , i.e. first solve the equation of second order in  $e^w$ ,

$$\{e^w\}^2 - 2z e^w + 1 = 0,$$

and then take the logarithm of the solutions. We end of course with (58) as in the real case, and apparently no complex analysis is involved. So far, so good, and nothing can be changed here.

Then let us try to analyze what could be done instead. Since the square root is 2-valued, formula (58) is also equivalent to

$$(60) \quad \boxed{\operatorname{arcosh} z = \log\left(z \pm \sqrt{1 - z^2}\right) = i \arccos z, \quad z \in \mathbb{C},}$$

and this version can then be used to define another branch of  $\operatorname{arcosh} z$  by instead using  $i \cdot \operatorname{Arccos} z$  in the slit domain  $\Omega = \mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[)$ . This version may look better than the previous one. Unfortunately there is of course no hope that this branch is equal to the real  $\operatorname{Arcosh}_{\mathbb{R}} t$  for  $t > 0$ , because  $]1, +\infty[$  does not lie in the domain of the alternative branch sketched above. So only  $\Omega_1$  meets our demands for the time being, although  $\Omega_1$  clearly does not correspond to a *fundamental domain*.

The analysis above shows again that the many-valued functions may be very difficult to treat, when we want to choose a convenient branch for a given problem. In particular, the reader is warned against only to trust her/his intuition.

### 2.7.7 The infinite-valued area function $\operatorname{artanh} z$

The inverse of the function  $z = \tanh w$  for  $w \neq i\left\{\frac{\pi}{2} + p\pi\right\}$ ,  $p \in \mathbb{Z}$ , is given by the infinite-valued function

$$(61) \quad \boxed{\operatorname{artanh} z = \frac{1}{2} \log\left(\frac{1+z}{1-z}\right), \quad z \in \mathbb{C} \setminus \{-1, 1\},}$$

with the two *logarithmic branching points*  $z_1 = -1$  and  $z_2 = 1$ .

Its *principal value* is given by

$$(62) \quad \boxed{\operatorname{Artanh} z = \frac{1}{2} \operatorname{Log}\left(\frac{1+z}{1-z}\right), \quad z \in \Omega = \mathbb{C} \setminus (]-\infty, -1] \cup [1, +\infty[),}$$

and it is equal to the well-known *real*  $\operatorname{Artanh}_{\mathbb{R}} t$  for  $t \in ]-1, 1[$ .

### 2.7.8 The infinite-valued area function $\operatorname{arcoth} z$

The inverse of the function  $z = \coth w$  for  $w \neq ip\pi$ ,  $p \in \mathbb{Z}$ , is given by the infinite-valued function

$$(63) \quad \boxed{\operatorname{arcoth} z = \frac{1}{2} \log\left(\frac{z+1}{z-1}\right), \quad z \in \mathbb{C} \setminus \{-1, 1\},}$$

with the two *logarithmic branching points*  $z_1 = -1$  and  $z_2 = 1$ .

Its *principal value* is given by

$$(64) \quad \boxed{\operatorname{Arcoth} z = \frac{1}{2} \operatorname{Log}\left(\frac{z+1}{z-1}\right), \quad z \in \Omega = \mathbb{C} \setminus [-1, 1].}$$

It is equal to the usual *real*  $\operatorname{Arcoth}_{\mathbb{R}} t$ , when  $z = t \in ]-\infty, -1[ \cup ]1, +\infty[$ .

### 2.7.9 Some further remarks on the arcus and area functions

We notice in particular the principal values

$$(65) \quad \operatorname{Arcsin} z, \quad \operatorname{Arccos} z, \quad \operatorname{Arctan} z, \quad \operatorname{Arsinh} z \quad \text{and} \quad \operatorname{Artanh} z,$$

because they are all defined and analytic in the open disc  $B(0, 1)$  of centre 0 and radius 1. This is *not* the case for remaining  $\operatorname{Arccot} z$ ,  $\operatorname{Arcosh} z$  and  $\operatorname{Arcoth} z$ . We see that the five functions of (65) all have convergent series expansions from  $z_0 = 0$  in the set  $B(0, 1)$ , and these series are obtained from the *real* series expansions known from *Real Calculus* by replacing  $x \in ]-1, 1[$  by the complex variable  $z \in B(0, 1)$ .

The functions  $\operatorname{Arccot} z$  and  $\operatorname{Arcoth} z$  allow instead convergent *Laurent series* expanded from  $z_0 = 0$  in the open domain  $\mathbb{C} \setminus B[0, 1]$ .

Finally, it follows from the discussion of the principal value  $\operatorname{Arcosh} z$  in Section 2.7.6 that the definition of  $\operatorname{Arcosh} z$  was chosen so unfortunate that we can neither get a reasonable power series expansion nor a reasonable Laurent series expansion when expanded from  $z_0 = 0$ .

### 2.8 Example: The inverse of $z = 3w^4 - 4w^3 - 6w^2 + 12w$

In this section we analyze a simple example of an algebraic equation in two complex variables. It has been constructed such that it is indeed possible to describe the corresponding Riemann surface without using too complicated computations. In general, such an analysis is far more difficult, when we try to describe the bijective correspondence between a  $z$ -Riemann surface and a  $w$ -Riemann surface. The present example can give some guidelines of how to handle the general case.

We choose the function  $f : \Omega = \mathbb{C} \rightarrow \mathbb{C}$  as the polynomial function

$$(66) \quad f(w) = 3w^4 - 4w^3 - 6w^2 + 12w,$$

and we want to describe the  $w$ -Riemann surface of the equation

$$(67) \quad f(w) = 3w^4 - 4w^3 - 6w^2 + 12w = z.$$

Notice that since  $z$  is uniquely determined by given  $w$ , the  $z$ -Riemann surface is just the usual  $z$ -plane  $\mathbb{C}$ .

By the *Fundamental Theorem of Algebra* the equation (67) has for given  $z \in \mathbb{C}$  four solutions  $w_1, w_2, w_3, w_4$ . For most  $z \in \mathbb{C}$  they are mutually different, so  $w = f^{-1}(z)$  is usually 4-valued, which implies that the corresponding Riemann surface is composed of four branches, i.e. we shall glue four slit complex planes together in a way which we shall describe in the following.

We shall find the *singular points* in the  $w$ -plane, where the function  $w = f^{-1}(z)$  is not 4-valued. These are the solutions of the equation  $f'(w) = 0$ , i.e.

$$f'(w) = 12(w^3 - w^2 - w + 1) = 12(w^2 - 1)(w - 1) = 12(w - 1)^2(w + 1) = 0.$$

We conclude that the two singular points are  $w_1 = 1$  and  $w_2 = -1$ . They correspond to the two *branching points* in the  $z$ -plane,

$$z_1 = f(w_1) = 5 \quad \text{and} \quad z_2 = f(w_2) = -11.$$

It follows implicitly from Section 2.3 that if  $w_0 \in \Omega$  is a zero of order  $n - 1$  of  $f'(w)$ , i.e.

$$f^{(j)}(w_0) = 0 \quad \text{for } j = 1, \dots, n - 1, \quad \text{and} \quad f^{(n)}(w_0) = f^{(1+n-1)}(w_0) \neq 0,$$

then the Riemann surface is qualitatively of the same structure as the  $n$ -th root  $\sqrt[n]{z - z_0}$  in a neighbourhood of  $z_0 = f(w_0)$ , cf. Section 2.3.

Since  $w_1 = 1$  is a double root of  $f'(w)$ , i.e.  $n - 1 = 2$ , the Riemann surface must have the structure of a cubic root in some neighbourhood of  $z_1 = f(w_1) = 5$ . Similarly we obtain the structure of a square root in some neighbourhood of  $z_2 = f(w_2) = -11$ .

**Remark 2.8.1** We could actually have used this principle already in Section 2.6, where we discussed the Riemann surface of  $\arcsin z$  and thus saved a lot of work. It was not done, because then the focus of the text would have been shifted away from the main issue.  $\diamond$



The reader who does not believe in the qualitative topological argument above may instead use the following *alternative* direct proof, which is also of independent interest.

When we factorize  $f(w) - f(w_1) = f(w) - 5$ , we get

$$f(w) - 5 = 3w^4 - 4w^3 - 6w^2 + 12w - 5 = (3w + 5)(w - 1)^3.$$

In a small neighbourhood of  $w_0 = 1$  the factor  $3w + 5$  only deviates very little from  $3 \cdot 1 + 5 = 8$ , so in this neighbourhood  $(w - 1)^3$  determines the structure of the Riemann surface as locally the Riemann surface of a cubic root, because we shall solve the equation

$$8(w - 1)^3 \approx (3w + 5)(w - 1)^3 = z - f(w_1) = z - 5$$

in a neighbourhood of  $z_1 = 5$ .

Analogously, it follows from

$$\begin{aligned} f(w) - f(w_2) &= f(w) + 11 = 3w^4 - 4w^3 - 6w^2 + 12w + 11 \\ &= (3w^2 - 10w + 11)(w + 1)^2 = z - f(w_2) = z + 11 = z - z_2, \end{aligned}$$

that  $(w + 1)^2$  locally determines the structure of a square root, because the factor  $3w^2 - 10w + 11$  in some small neighbourhood of  $w_2 = -1$  deviates very little from 24, so

$$w \approx -1 + \sqrt{\frac{z - z_2}{24}}, \quad \text{in a neighbourhood of } z_2 = -11,$$

and we have produced an alternative proof of the claim.

Usually one acts from the *fundamental domains* and then finds the *branch cuts*, but since the root structures are well-known, we can here instead first fix the *branch cuts* and then find the *fundamental domains*.

We choose the *branch cut*  $] -\infty, -11]$  for the square root, and the *branch cut*  $[5, +\infty[$  for the cubic root.

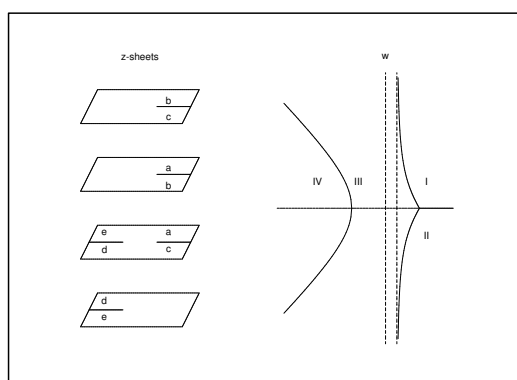


Figure 16: Construction of the Riemann surface of  $3w^4 - 4w^3 - 6w^2 + 12w = z$ .

When we construct the Riemann surface we must start with four  $z$ -planes, all containing at least one *branch cut*. We have in total five branch cuts, two from the square root and three from the cubic root. Therefore, just one plane must contain *two* branch cuts, and the remaining ones only one branch cut. Let plane III on Figure 16 contain both branch cuts.

We have above plane III the two planes I and II, which together with III describe the cubic root. Therefore, their branch cuts are given by  $]5, +\infty[$ .

Similarly, planes III and IV describe the square root, so their branch cuts are given by  $] - \infty, -1]$ .

Due to the previous qualitative analysis we immediately see that the planes are glued together along the edges called  $a$ ,  $b$  and  $d$ , respectively, and we *imagine* that we can also identify the edges called  $c$  and  $e$ . (Of course, this cannot be physically realized in the usual three dimensional space without cutting through an existing surface.)

It remains to find the *fundamental domains* in the  $w$ -plane. They must be bounded by curves which by  $z = f(w)$  are mapped either onto  $] - \infty, -1]$ , which must hold for two curves from  $w = -1$  to  $\infty$ , or onto  $]5, +\infty[$ , which must hold for three curves from  $w = +1$  to  $\infty$ . The half tangents at  $w = +1$  must mutually form the angle  $\frac{2\pi}{3}$ , because we in a small neighbourhood approximately have the structure of a cubic root.

Assume that

$$C : w(t) = w = u + iv, \quad t \in \mathcal{I},$$

is the parametric description of such a curve  $C$ . Then  $f(C)$  is either  $] - \infty, -1]$  or  $]5, +\infty[$ . In particular,  $f(C)$  is contained in  $\mathbb{R}$ , so we conclude that  $\Im f(w) = 0$ , thus

$$0 = \Im \{3w^4 - 4w^3 - 6w^2 + 12w\} = 4v \{3u^3 - 3uv^2 - 3u^+v^2 - 3u + 3\}.$$

This implies that either  $v = 0$ , or

$$3u^3 - 3uv^2 - 3u^2 + v^2 - 3u + 3 = v^2(1 - 3u) + 3(u^3 - u^2 - u + 1) = v^2(1 - 3u) + 3(u - 1)^2(u + 1),$$

so either

$$v = 0, \quad \text{or} \quad v^2 = \frac{3(u - 1)^2(u + 1)}{3u - 1} \quad \text{for } u \neq \frac{1}{3}.$$

Assume that  $v = 0$ . Obviously, the interval  $] - 11, 5]$  between the two branching points does not belong to the range, so when we look at Figure 16 it follows that only  $u \geq +1$  is possible, and the curve  $w(t) = u + iv = u$  for  $u = t \geq +1$  is mapped onto the branch cut  $]5, +\infty[$ .

Since the half tangents of the curves at  $w = +1$  form mutually the angle  $\frac{2\pi}{3}$  we also conclude that  $u \leq 1$  for the two remaining curves of the cubic root.

If  $v \neq 0$ , then

$$v^2 = \frac{3(u - 1)^2(u + 1)}{3u - 1}, \quad \text{for } u \neq \frac{1}{3}.$$

Since  $v^2 > 0$ , we must have  $u \neq 1$  and also  $\frac{u + 1}{3u - 1} > 0$ , so either  $u < -1$ , or  $\frac{1}{3} < u < 1$ , or  $u > 1$ . The latter possibility is immediately rejected, because we just proved above that  $u \leq 1$  for the remaining curves describing the cubic root.

We get from the above,

$$v = \pm(u-1)\sqrt{\frac{3(u+1)}{3u-1}} = \pm(u-1)\sqrt{1 + \frac{4}{3u-1}} \quad (\approx \pm(u-1) \quad \text{if } u < -1 \text{ large}),$$

where, using the continuity at the finite end points,

- 1) the *square root*,  $u \in ]-\infty, -1]$ , with the asymptotes  $\pm(u-1)$ ,
- 2) the remaining two curves of the *cubic root*,  $u \in \left] \frac{1}{3}, 1 \right]$ .

Clearly,  $|v| \rightarrow +\infty$  for  $u \rightarrow \frac{1}{3}+$ , so  $u = \frac{1}{3}$  is a vertical asymptote of the two curves above related to the cubic root.

Together these curves divide the  $w$ -plane into four fundamental domains. The fundamental domain, which contains the vertical axis, has both  $w = -1$  and  $w = +1$  as boundary points, so it must necessarily correspond to plane II with two branch cuts. Then IV must correspond to the other fundamental domain which is characterized by also having  $w = -1$  as a boundary point. The remaining two fundamental domains correspond to the planes I and III, depending only on how they are glued together in the model of the Riemann surface.

This example was constructed in order to show the main steps in the construction of a Riemann surface, and yet the computations should not be too complicated.

## 2.9 Simple examples of potential flows in Physics

We shall in the following present the four simplest 2-dimensional *potentials* occurring in the technical sciences. In the first three cases, *source* (and *sink*), *vortex* and *vortex source/sink*, the potential is described by a many-valued function. In the last case, the *dipoles*, we obtain a usual analytic function as its potential.

### 2.9.1 Sources and sinks

The *potential* of a *source* at  $z = 0$  is defined as a many-valued function of the form

$$(68) \quad F(z) = \frac{N}{2\pi} \log z + c, \quad z \in \mathbb{C} \setminus \{0\},$$

where  $N \in \mathbb{R} \setminus \{0\}$  and  $c \in \mathbb{C}$  are constants.

The number  $N$  is called the *strength of the source*. If  $N < 0$ , we instead call the source a *sink*.

When the potential  $F(z) = \varphi(x, y) + i \cdot \psi(x, y)$  is split into its real and imaginary parts, we get

$$(69) \quad \varphi(x, y) = \frac{N}{2\pi} \ln |z| + c_1, \quad \text{and} \quad \psi(x, y) = \frac{N}{2\pi} \arg z + c_2.$$

The *equipotential curves* are defined as the curves implicitly given by  $\varphi(x, y) = k$ , where  $k \in \mathbb{R}$  is a constant. They are in the present case circles  $|z| = C$ , where  $C = \exp\left(\frac{2\pi}{N}(k - c_1)\right)$ , cf. Figure 17.

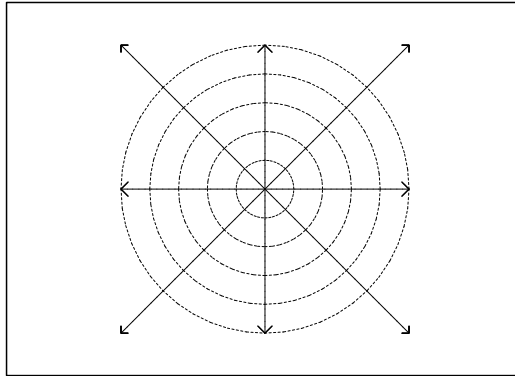


Figure 17: Streamlines and equipotential curves of a source.

The *streamlines* are defined as the curves implicitly given by  $\psi(x, y) = k$ , where  $k \in \mathbb{R}$  is a constant. In the present case they are the half lines given in polar coordinates by  $\Theta \in [0, 2\pi[$  fixed, where  $z = r \cdot e^{i\Theta}$  and  $r > 0$  varies. Cf. Figure 17.

If the *strength*  $N > 0$  is positive, we interpret this model as a point source at  $z = 0$ , from which a flow is going radially outwards. If instead  $N < 0$  is negative, then  $z = 0$  is interpreted as a point sink, where the radial flow is going into  $z = 0$ , at which point it disappears.

### 2.9.2 Vortices

A *vortex* at  $z = 0$  is defined by its complex many-valued potential

$$(70) \quad F(z) = \frac{\Gamma i}{2\pi} \log z + c, \quad z \in \mathbb{C} \setminus \{0\}.$$

The number  $\Gamma \in \mathbb{R} \setminus \{0\}$  is called the *intensity of the vortex*. When we split  $F(z)$  into its real and imaginary parts, we get

$$(71) \quad \varphi(x, y) = -\frac{\Gamma}{2\pi} \arg z + c_1, \quad \text{and} \quad \psi(x, y) = \frac{\Gamma}{2\pi} \ln |z| + c_2.$$

The *equipotential curves* are given by  $\varphi(x, y) = k_1$ , where  $k_1$  is some real constant, and the *streamlines* are determined by  $\psi(x, y) = k_2$ , where  $k_2$  is some real constant. Cf. also Figure 17. We see that the streamlines are circles of centre  $z = 0$ , so the flow whirls around  $z = 0$  in concentric circles.

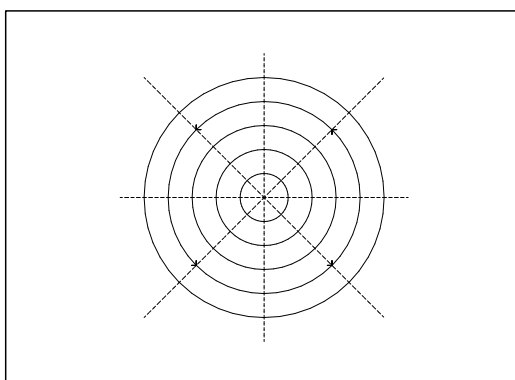


Figure 18: Streamlines and equipotential curves of a vortex.

### 2.9.3 Vortex sources/sinks

When a source and a vortex, both at  $z = 0$ , are added, we get a *vortex source/sink* of the complex potential

$$(72) \quad F(z) = \frac{N + i\Gamma}{2\pi} \log z + c, \quad z \neq 0,$$

of the *real potential*

$$(73) \quad \varphi(x, y) = \frac{N}{2\pi} \ln |z| - \frac{\Gamma}{2\pi} \arg z + c_1,$$

and the *flow function*

$$(74) \quad \psi(x, y) = \frac{\Gamma}{2\pi} \ln |z| + \frac{N}{2\pi} \arg z + c_2.$$

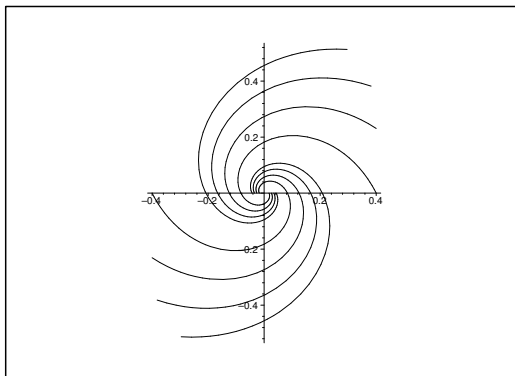


Figure 19: Streamlines a vortex sink.

Using polar coordinates,  $z = r e^{i\Theta}$ , the flow function is given by

$$\psi(x, y) = \frac{1}{2\pi} \{ \Gamma \cdot \ln r + N\Theta \} + c_2,$$

so the streamlines,  $\psi(x, y) = k$ , are in polar coordinates given by

$$(75) \quad r = C \cdot \exp\left(-\frac{N}{\Gamma} \Theta\right), \quad \Theta \in \mathbb{R},$$

which are recognized as *logarithmic spirals*. Cf. Figure 19

Similarly for the *equipotential curves*, which also become a system of logarithmic spirals, orthogonal to the system of streamlines.

#### 2.9.4 Dipoles

For given  $h > 0$ , consider a system consisting of a *source* at  $z_1 = -\frac{h}{2}$  of strength  $\frac{N}{h}$ , and a *sink* at  $z_2 = \frac{h}{2}$  of strength  $-\frac{N}{h}$ , cf. Figure 20.

We choose the complex potential of this system as the analytic function

$$F_h(z) = \frac{N}{2\pi h} \operatorname{Log}\left(z + \frac{h}{2}\right) - \frac{N}{2\pi h} \operatorname{Log}\left(z - \frac{h}{2}\right).$$

When we take the limit  $h \rightarrow 0+$  of this system we get the complex potential of a *dipole*,

$$(76) \quad F(z) = \frac{N}{2\pi} \lim_{h \rightarrow 0+} \frac{1}{h} \left\{ \operatorname{Log}\left(z + \frac{h}{2}\right) - \operatorname{Log}\left(z - \frac{h}{2}\right) \right\} = \frac{N}{2\pi} \cdot \frac{1}{z}, \quad \text{for } z \neq 0.$$

Clearly,  $F(z) = \frac{N}{2\pi} \cdot \frac{1}{z}$  for  $z \in \mathbb{C} \setminus \{0\}$  is single-valued, and it is easy to prove that the *streamlines* are circular arcs of centre  $(0, y)$  on the imaginary axis and of radius  $|y|$ , where only the point  $z = 0$  has been excluded from the circles.

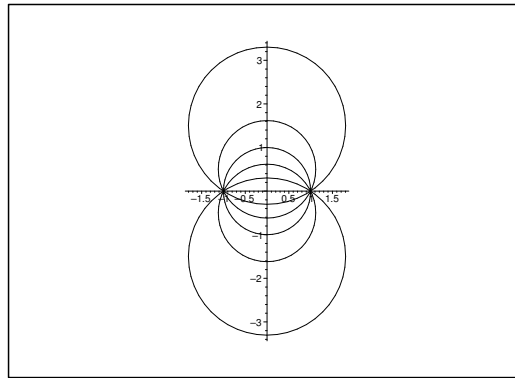


Figure 20: Streamlines of a source and a sink of numerically equal strength.

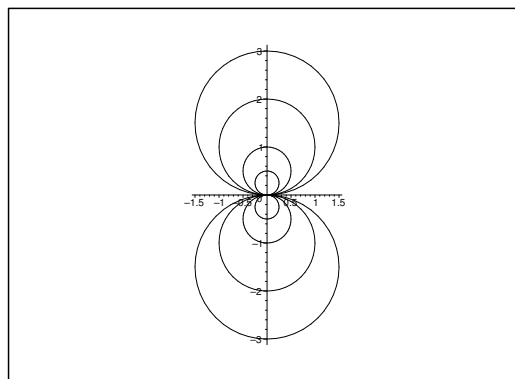


Figure 21: Streamlines of a dipole, obtained by taking the limit  $h \rightarrow 0+$  in Figure 20.

Similarly, the *equipotential curves* are circular arcs of centre  $(x, 0)$  on the real axis and of radius  $|x|$ , where only the point  $z = 0$  has been excluded from the circles.

### 3 Conformal mappings and the Dirichlet problem

#### 3.1 Conformal mappings

Let  $f : \Omega \rightarrow \mathbb{C}$  be an analytic map for which  $f'(z) \neq 0$  for every  $z \in \mathbb{C}$ . Choose any point  $z_0 \in \Omega$  and put for convenience  $f(z_0) = w_0$  and  $f'(z_0) = \rho e^{i\alpha}$ . Then according to a theorem in *Ventus, Complex Functions Theory a-1*, Chapter 3, some neighbourhood  $\omega$  of  $z_0$  is by  $f$  mapped *bijectively* onto a neighbourhood  $f(\omega)$  of  $w_0$ .

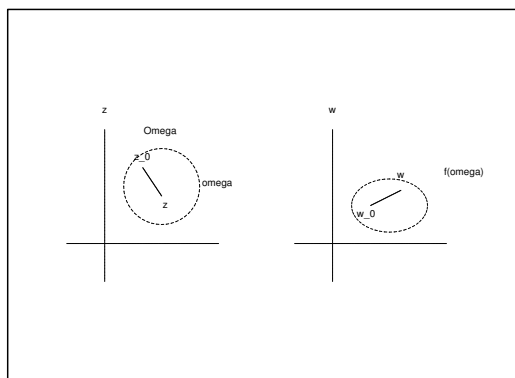


Figure 22: If  $f'(z_0) \neq 0$ , then  $f$  is locally bijective.

We notice that  $\Delta f := f(z_0 + \Delta z) - f(z_0)$  is for small values of  $|\Delta z|$  approximately equal to  $f'(z_0) \Delta z$ , hence the map is in some (small) neighbourhood of  $z_0$  approximately given by

$$w - w_0 = f'(z_0) \cdot (z - z_0) = \rho e^{i\alpha} \cdot (z - z_0).$$

The latter map is a *similarity*, by which  $z_0$  is mapped into  $w_0$ , and every vector  $z - z_0$  from  $z_0$  is mapped into a vector  $w - w_0$ , where  $w - w_0$  is obtained from the vector  $z - z_0$  by multiplying its length by  $\rho$  and rotating it by the angle  $\alpha$ .

Such a map  $f$  is called a *conformal map* (a local similarity), when also the sense of orientation of the complex plane is preserved. The number  $\rho = |f'(z_0)|$  is called the *stretching factor* of the map at  $z_0$ , and the number  $\alpha = \text{Arg } f'(z_0)$  is called the *angle of rotation* of the conformal map  $f$  at  $z_0$ .

An angle at  $z_0$  is determined by two vectors from  $z_0$ . By a conformal map both vectors are rotated the same angle  $\alpha$ , so the size of the angle is preserved. We say that a *conformal map  $f$  preserves angles*.

It follows from the above that an analytic function  $f$  is *conformal* at every point  $z$  in its domain for which also  $f'(z) \neq 0$ . We mention without proof the opposite result:

**Theorem 3.1.1** *Let  $\Omega \subseteq \mathbb{C}$  be an open domain. If a map  $f : \Omega \rightarrow \mathbb{C}$  is conformal everywhere in  $\Omega$ , then  $f$  is analytic in  $\Omega$ .*



**Remark 3.1.1** It should be emphasized that there of course exist analytic maps  $f$  which are not conformal in all of its open domain  $\Omega$ . This is not the case at points  $z \in \Omega$  where  $f'(z) = 0$ , and  $f$  must necessarily be analytic at  $z$ , because otherwise the derivative did not exist.  $\diamond$

In practical applications conformal maps are very important, because they can be applied to reduce complicated *Dirichlet problems* of the 2-dimensional *Laplace differential operator*

$$\Delta := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, \quad z = x + iy,$$

into simpler problems, where one can apply a catalogue of known solutions, or in more complicated cases, where a practical solution formula exists. One example of this has already been given in *Ventus, Complex Functions Theory a-2*, Chapter 2 on *Poisson's integral formula* for the *Dirichlet problem* of a disc. By combining Poisson's integral formula with the considerations above, it follows that all we need is to find a conformal map, which maps the simply connected open domain of the original Dirichlet problem to the open unit disc and then apply *Poisson's integral formula*.

In principle, this is easy to understand. However, when we try to apply this method, we may come across some nasty approximation problems, which cannot be solved at this step of the description of the theory. Instead we head on by creating an arsenal of well-known conformal maps, which later may be useful in solving the Dirichlet problem mentioned above.

### 3.2 The Möbius transformations

The simplest possible class of conformal maps is the class of *Möbius transformations*, also called the class of *bilinear* or *linear fractional transformations*, or *homographies*. They are defined in the following definition.

**Definition 3.2.1** A Möbius transformation is an analytic map  $w = f(z)$  of the form

$$(77) \quad w = \frac{az + b}{cz + d}, \quad \text{where } a, b, c, d \in \mathbb{C} \text{ and } \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0.$$

If  $c = 0$ , then its domain is  $\mathbb{C}$ . If  $c \neq 0$ , then its domain is  $\mathbb{C} \setminus \left\{ -\frac{d}{c} \right\}$ .

By a small computation,

$$\frac{dw}{dz} = \frac{ad - bc}{(cz + d)^2} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \cdot \frac{1}{(cz + d)^2} \neq 0, \quad \text{defined for } z \neq -\frac{d}{c},$$

thus a Möbius transformation is conformal everywhere in its domain.

Another small computation shows that (77) has its inverse given by

$$(78) \quad \boxed{z = \frac{-dw + b}{cw - a}, \quad w \neq \frac{a}{c}, \quad \begin{vmatrix} -d & b \\ c & -a \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0,}$$

which again is a Möbius transformation.

**Remark 3.2.1** It is quite natural to let the matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  correspond to the map given in (77).

Then (78) corresponds to the matrix  $\begin{pmatrix} -d & b \\ c & -a \end{pmatrix}$ , which is obtained from the former matrix by interchanging the two diagonal elements and then change the signs of the diagonal, and finally, interchange the variables  $z$  and  $w$ . Since we often shall find the inverse of a Möbius transformations, this observation will prove very useful in practice.  $\diamond$

We proved above that (77) defines a bijective map of  $\mathbb{C} \setminus \left\{ -\frac{d}{c} \right\}$  onto  $\mathbb{C} \setminus \left\{ \frac{a}{c} \right\}$  (trivial modifications, if  $c = 0$ ). If we put  $\mathbb{C}^* = \mathbb{C} \cup \{\infty\}$ , and let  $z = -\frac{d}{c}$  be mapped into  $w = \infty$ , and  $z = \infty$  into  $w = \frac{a}{c}$ , then the *extended Möbius transformation becomes a bijective map of  $\mathbb{C} \cup \{\infty\}$  onto itself*.

Conversely, it can be proved that *the only bijective conformal maps  $f : \mathbb{C} \cup \{\infty\} \rightarrow \mathbb{C} \cup \{\infty\}$  are the (extended) Möbius transformation*.

It was previously shown that a Möbius transformation corresponds to a regular  $(2 \times 2)$ -matrix of complex coefficients by

$$f(z) = \frac{az + b}{cz + d} \sim \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0.$$

A simple and straightforward, and also extremely tedious computation, which is left to the reader, shows that composition of two Möbius transformations corresponds to matrix multiplication of their corresponding matrices.

**Theorem 3.2.1** *The composition of two Möbius transformations  $f$  and  $g$  is again a Möbius transformation,  $f \circ g$ . If*

$$f \sim \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \quad \text{and} \quad g \sim \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix},$$

then

$$f \circ g \sim \begin{pmatrix} a_3 & b_3 \\ c_3 & d_3 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}.$$

Notice that the matrices are put in the same order as their corresponding Möbius transformations.

**Remark 3.2.2** It follows from Theorem 3.2.1 that the class of Möbius transformations becomes a group with the usual composition of maps as “multiplication”. This group is isomorphic with a subgroup of the matrix group

$$(79) \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{C}; \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 1 \right\}.$$

In fact, if a matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  corresponds to the Möbius transformation  $f(z) = \frac{az + b}{cz + d}$ , then all matrices corresponding to  $f(z)$  are given by

$$\lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}, \quad \text{for } \lambda \in \mathbb{C} \setminus \{0\}.$$

Since the equation in  $\lambda$

$$\begin{vmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{vmatrix} = \lambda^2 \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 1$$

has two solutions,  $\lambda_1$  and  $\lambda_2 = -\lambda_1$ , we conclude that the group of Möbius transformations is isomorphic with a subgroup of index 2 of the group described in (79).  $\diamond$

The following four simple examples will help us to describe the Möbius transformations in Theorem 3.2.2.

**Example 3.2.1** *Every translation*

$$w = z + b, \quad b \in \mathbb{C},$$

is a Möbius transformation. In fact,

$$w = z + b = \frac{1 \cdot z + b}{0 \cdot z + 1} \sim \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad \text{where } \begin{vmatrix} 1 & b \\ 0 & 1 \end{vmatrix} = 1 \neq 0. \quad \diamond$$

**Example 3.2.2** Every *similarity* of *stretching factor*  $\varrho > 0$ , and every *rotation* by an angle  $\alpha \in \mathbb{R}$  around the origo is a Möbius transformation. In fact, their composition is homothetic and

$$w = az = \frac{az + 0}{0 \cdot z + 1} \sim \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{where } \begin{vmatrix} a & 0 \\ 0 & 1 \end{vmatrix} = a = \varrho e^{i\alpha} \neq 0,$$

which is clearly a Möbius transformation. The similarity is obtained by choosing  $\alpha = 0$ , and the rotation is obtained by choosing  $\varrho = 1$ .  $\diamond$

**Example 3.2.3** A composition of the Möbius transformations of Example 3.2.1 and Example 3.2.2 gives the *general similarity* for  $a \in \mathbb{C} \setminus \{0\}$  and  $b \in \mathbb{C}$ ,

$$w = az + b = \frac{a \cdot z + b}{0 \cdot z + 1} \sim \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}, \quad \text{where } \begin{vmatrix} a & b \\ 0 & 1 \end{vmatrix} = a \neq 0. \quad \diamond$$

**Example 3.2.4** The *reciprocal transformation* is a Möbius transformation,

$$w = \frac{1}{z} = \frac{0 \cdot z + 1}{1 \cdot z + 0} \sim \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{where } \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0. \quad \diamond$$

Let

$$f(z) = \frac{az + b}{cz + d} \sim \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{where } \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0,$$

be a Möbius transformation. If  $c = 0$ , then

$$f(z) = \frac{a}{d}z + \frac{b}{d} \sim \begin{pmatrix} a/d & b \\ 0 & 1 \end{pmatrix},$$

is of course a general similarity, cf. Example 3.2.3.

If  $c \neq 0$ , then we get by a simple matrix computation (again left to the reader),

$$(80) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} (bc - ad)/c & a/c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix}.$$

We therefore conclude from the previous four examples and Theorem 3.2.1 that we have proved

**Theorem 3.2.2** *Every Möbius transformation is either a general similarity, or it is the composition of a general similarity, the reciprocal transformation, and another general similarity.*

Due to Theorem 3.2.2 and formula (80) the investigation of the properties of the Möbius transformations is reduced to the investigations of the general similarities and the reciprocal transformation. Hence, if we can find geometrical properties which are preserved by these two special Möbius transformations, then they are preserved in general by Möbius transformations. Once this was realized by mathematicians in the past, they decided that the following definition is quite useful in the theory of Möbius transformations.

**Definition 3.2.2** *A circle in the extended complex plane  $\mathbb{C} \cup \{\infty\}$  is either a circle in the usual sense in  $\mathbb{C}$ , or a straight line, which is considered as a circle of “infinite radius” and no centre and passing through  $\infty$ .*

Assume that a finite circle has radius  $r > 0$  and its centre at  $b \in \mathbb{C}$ . Then its equation is given by  $|z - b| = r$ , hence

$$(81) \quad |z - b|^2 - r^2 = (z - b)(\bar{z} - \bar{b}) - r^2 = z\bar{z} - \bar{b}z - b\bar{z} + \{|b|^2 - r^2\} = 0.$$

A straight line in the plane has the equation  $\alpha x + \beta y = c \in \mathbb{R}$ , where  $(\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ . If we put  $b := \frac{1}{2}(\alpha + i\beta)$ , an elementary computation shows that  $2\Re\{\bar{b}z\} = \alpha x + \beta y = c$ , so the equation of a straight line can also be written

$$(82) \quad -2\Re\{\bar{b}, z\} + c = -\{\bar{b}z + \overline{\bar{b}z}\} + c = -\bar{b}z - b\bar{z} + c = 0.$$

When we compare (81) and (82) it follows that a general circle in  $\mathbb{C} \cup \{\infty\}$  has the equation

$$(83) \quad az\bar{z} - \bar{b}z - b\bar{z} + c = 0, \quad a, c \in \mathbb{R}, \quad b \in \mathbb{C}, \quad |b|^2 - ac > 0.$$

We see that if  $a \neq 0$ , then we get (81), and if  $a = 0$ , then we get (82).

Using (83) we prove

**Theorem 3.2.3** *Every Möbius transformation preserves the class of general circles in  $\mathbb{C} \cup \{\infty\}$ , i.e. every general circle in  $\mathbb{C} \cup \{\infty\}$  is by a Möbius transformation mapped onto another general circle in  $\mathbb{C} \cup \{\infty\}$ .*

PROOF. It follows from Theorem 3.2.2 that we shall only prove the claim for the general similarities, in which case it is trivial, and for the reciprocal transformation  $w = \frac{1}{z}$ . When we put  $z = \frac{1}{w}$  into (83) it follows that the image curve has the equation

$$\frac{a}{w\bar{w}} - \frac{\bar{b}}{w} - \frac{b}{\bar{w}} + c = 0.$$

When this equation is multiplied by  $w\bar{w} = |w|^2$ , we get

$$cw\bar{w} - bw - \bar{b}\bar{w} + a = 0, \quad c, a \in \mathbb{R}, \quad \bar{b} \in \mathbb{C}, \quad |\bar{b}|^2 - ca > 0,$$

and the claim follows.  $\square$

It is well-known that a circle in the complex plane is uniquely determined by three of its points. Therefore, a Möbius transformation  $f$  is uniquely determined if we just know the (mutually different) images  $w_1 = f(z_1)$ ,  $w_2 = f(z_2)$  and  $w_3 = f(z_3)$  of three mutually different points  $z_1$ ,  $z_2$  and  $z_3$  in the  $z$ -plane. In fact, if

$$w = f(z) = \frac{az + b}{cz + d},$$

then we get by insertion of  $w_k = f(z_k)$ ,  $k = 1, 2, 3$ , three linear equations in the four unknowns  $a, b, c, d$ . Now,  $(c, d) \neq (0, 0)$ , so at least one of the coordinates is  $\neq 0$ , and since the Möbius transformation is unaltered, if the numerator and the denominator are multiplied by the same constant  $\lambda \neq 0$ , we may assume that either  $c$  or  $d$  is equal to 1 (this choice depends on which one is  $\neq 0$ ). By such a choice we reduce the system to three linear equations in three unknowns, and the theory from *Linear Algebra* takes over.

The above is of course not a precise proof, though the method works in practice. The details of *Linear Algebra* sketched above were not carried out in full details, because we have the following more easy method, which is highly recommended instead of the probably better known method from *Linear Algebra*. We first introduce

**Definition 3.2.3** *Let  $z_1, z_2, z_3 \in \mathbb{C}$  be mutually different complex constants, and let  $z \in \mathbb{C}$  be a complex variable. We define the cross-ratio of  $z, z_1, z_2, z_3$  (in this order) by*

$$(84) \quad \varphi(z) = [z, z_1, z_2, z_3] := \frac{z - z_2}{z - z_3} : \frac{z_1 - z_2}{z_1 - z_3}.$$

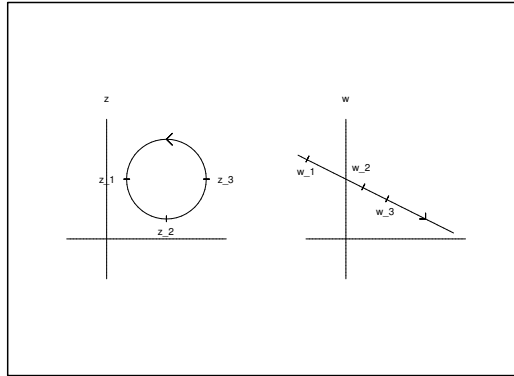


Figure 23: A Möbius transformation maps general circles into general circles in  $\mathbb{C} \cup \{\infty\}$ , and domains “to the left” onto domains “to the left”.

The cross-ratio  $\varphi(z)$  is clearly a Möbius transformation, so its domain can be extended to  $\mathbb{C} \cup \{\infty\}$  by putting

$$\varphi(z_3) = \infty \quad \text{and} \quad \varphi(\infty) = 1 : \frac{z_1 - z_2}{z_1 - z_3} = \frac{z_1 - z_3}{z_1 - z_2}.$$

Since  $\varphi(z)$  is a Möbius transformation, we also allow  $z_1, z_2, z_3$  to be  $\infty$ , as long as all three constants are mutually different. We mention that

$$(85) \quad [z, \infty, z_2, z_3] = \frac{z - z_2}{z - z_3}, \quad [z, z_1, \infty, z_3] = \frac{z_1 - z_3}{z - z_3}, \quad [z, z_1, z_2, \infty] = \frac{z - z_2}{z_1 - z_2},$$

which formally are obtained by putting  $\frac{\infty}{\infty} := 1$  into (84). Notice that *this is the only place in Complex Functions Theory that we allow this rule of computation.*

The importance of the cross-ratios follows from the next theorem.

**Theorem 3.2.4** *Every Möbius transformation  $w = f(z)$  preserves cross-ratios, i.e.*

$$[z, z_1, z_2, z_3] = [f(z), f(z_1), f(z_2), f(z_3)] = [w, w_1, w_2, w_3].$$

PROOF. If  $w = f(z) = az + b$  is a *general similarity*,  $a \neq 0$ , then we put  $w : k = az_k + b$ ,  $k = 1, 2, 3$ , and we get

$$\begin{aligned} [w, w_1, w_2, w_3] &= \frac{w - w_2}{w - w_3} : \frac{w_1 - w_2}{w_1 - w_3} = \frac{(az + b) - (az_2 + b)}{(az + b) - (az_3 + b)} : \frac{(az_1 + b) - (az_2 + b)}{(az_1 + b) - (az_3 + b)} \\ &= \frac{a(z - z_2)}{a(z - z_3)} : \frac{a(z_1 - z_2)}{a(z_1 - z_3)} = [z, z_1, z_2, z_3]. \end{aligned}$$

If instead  $w = \frac{1}{z}$  is the *reciprocal transformation*, we get similarly,

$$[w, w_1, w_2, w_3] = \frac{\frac{1}{z} - \frac{1}{z_2}}{\frac{1}{z} - \frac{1}{z_3}} : \frac{\frac{1}{z_1} - \frac{1}{z_2}}{\frac{1}{z_1} - \frac{1}{z_3}} = \frac{z_2 - z}{z_3 - z} : \frac{z_2 - z_1}{z_3 - z_1} = \frac{z - z_2}{z - z_3} : \frac{z_1 - z_2}{z_1 - z_3} = [z, z_1, z_2, z_3].$$

Since the claim holds for the *general similarities* and the *reciprocal transformation*, it follows from Theorem 3.2.2 that it holds in general.  $\square$

We immediately get by using Theorem 3.2.4 and the previous remarks that a general circle in  $\mathbb{C} \cup \{\infty\}$  is determined by three mutually different points the following result

**Corollary 3.2.1** *Let  $z_1, z_2, z_3 \in \mathbb{C} \cup \{\infty\}$  be three mutually different points in the  $z$ -plane, and  $w_1, w_2, w_3 \in \mathbb{C} \cup \{\infty\}$  three mutually different points in the  $w$ -plane. Let  $f$  denote the Möbius transformation, for which*

$$f(z_1) = w_1, \quad f(z_2) = w_2, \quad f(z_3) = w_3.$$

*Then  $w = f(z)$  is uniquely determined by the cross-ratio*

$$(86) \quad [z, z_1, z_2, z_3] = [f(z), w_1, w_2, w_3].$$



The advantage of (86) is that according to (84) the right hand side can be written in the form

$$(87) \quad \frac{af(z) + b}{cf(z) + d} = [z, z_1, z_2, z_3] = \varphi(z),$$

where, apart from a common factor  $\neq 0$ , the coefficients  $a, b, c, d$  are determined by  $w_1, w_2, w_3$ . Then by using the inversion formula (78) we get

$$(88) \quad f(z) = \frac{-d\varphi(z) + b}{c\varphi(z) - a}.$$

It is in principle possible to express (88) alone in the given complex constants  $z_1, z_2, z_3, w_1, w_2, w_3$  and the variable  $z$ , but it is not worth the effort, because the expression becomes very complicated. A far better strategy in practice is first to compute  $\varphi(z)$  by (87) and then insert the result into (88) and reduce.

**Remark 3.2.3** It should here be mentioned that the cross-ratio defined above also occurs in the theory of *Riccati differential equations*, which otherwise has nothing to do with *Complex Functions Theory*.  $\diamond$

**Example 3.2.5** As an illustration of the theory and method described above we want to find the Möbius transformation which maps  $z = 1, i, 0$  into  $w = 0, i, 1$  in the given order.

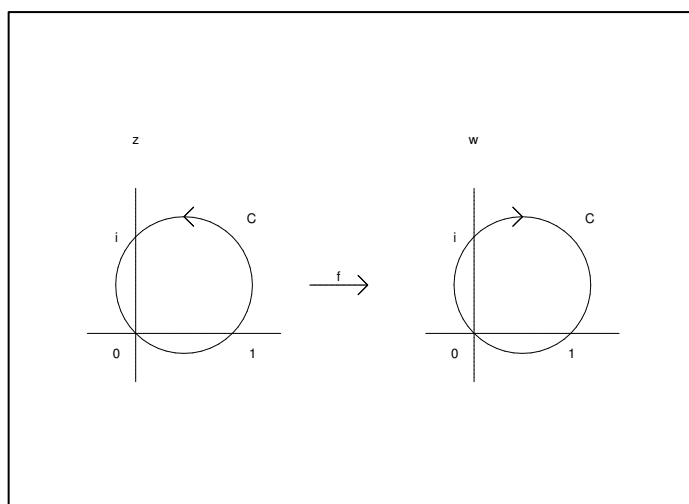


Figure 24: The Möbius transformation of Example 3.2.5 reverses the direction on the circle, so the open disc is mapped onto the interior of its complementary set.

It is convenient, though not necessary, first to make a geometrical analysis, cf. Figure 24. The point sets  $z = 1, i, 0$  and  $w = 0, i, 1$  both determine the same circle  $C$  of the equation

$$C: \quad \left| z - \frac{1+i}{2} \right| = \frac{1}{\sqrt{2}}.$$

The difference is that  $z = 1, i, 0$  on the circle are ordered in the *positive* sense of the plane, while  $w = 0, i, 1$  on the circle are ordered in the *negative* sense of the plane. Since domains to the left of closed curves are always mapped onto domains to the left of the image curves by Möbius transformations (or, in general, by analytic functions) seen in the defined direction of the curves, we conclude that the inner disc is mapped onto the unbounded set outside the circle  $C$ , i.e.  $\left|z - \frac{1+i}{2}\right| < \frac{1}{\sqrt{2}}$  is by  $w = f(z)$  mapped onto  $\left|w - \frac{1+i}{2}\right| > \frac{1}{\sqrt{2}}$  in  $\mathbb{C} \cup \{\infty\}$ .

After this geometrical digression we turn to the suggested method, which is the main topic here. It follows from Corollary 3.2.1 that  $w = f(z)$  is given by

$$[z, 1, i, 0] = [w, 0, i, 1],$$

i.e.

$$\frac{z-i}{z-0} : \frac{1-i}{1-0} = \frac{w-i}{w-1} : \frac{0-i}{0-1},$$

which is reduced to

$$\frac{w-i}{w-1} = z_1 := \frac{i}{1-i} \cdot \frac{z-i}{z}.$$

Then by the inverse transformation (78),

$$w = \frac{z_1 - i}{z_1 - 1} = \frac{\frac{i}{1-i} \cdot \frac{z-i}{z} - i}{\frac{i}{1-i} \cdot \frac{z-i}{z} - 1} = \frac{i(z-i) - i(1-i)z}{i(z-i) - (1-i)z} = \frac{-z+1}{(2i-1)z+1},$$

i.e.

$$w = \frac{-z+1}{(2i-1)z+1}.$$

CHECK:

$$\frac{-z_1+1}{(2i-1)z_1+1} = \frac{-1+1}{(2i-1)+1} = 0 = w_1,$$

$$\frac{-z_2+1}{(2i-1)z_2+1} = \frac{-i+1}{(2i-1)i+1} = \frac{1-i}{-2-i+1} = \frac{1-i}{-1-i} = \frac{(1-i)(-1+i)}{(-1-i)(-1+i)} = \frac{2i}{2} = i = w_2,$$

$$\frac{-z_3+1}{(2i-1)z_3+1} = \frac{1}{1} = 1 = w_3.$$

It follows that we have found our wanted Möbius transformation.  $\diamond$

### 3.3 Some special Möbius transformations

We shall in this section shortly list all possible Möbius transformations between the upper half plane and the unit disc. All proofs are very simple applications of Corollary 3.2.1, so they are left to the reader as an exercise.

#### 3.3.1 Unit disc onto unit disc

The Möbius transformations which map the unit disc onto itself are all of the structure

$$(89) \quad w = e^{i\alpha} \cdot \frac{z - z_0}{1 - \bar{z}_0 z}, \quad \text{where } \alpha \in \mathbb{R} \text{ and } |z_0| < 1.$$

#### 3.3.2 Unit disc onto complementary set of the unit disc

The Möbius transformations which map the open unit disc onto the complementary set of the closed unit disc are either of the structure

$$(90) \quad w = e^{i\alpha} \cdot \frac{z - z_0}{1 - \bar{z}_0 z}, \quad \text{where } \alpha \in \mathbb{R} \text{ and } |z_0| > 1,$$

or of the structure (more or less corresponding to  $z_0 = \infty$ )

$$(91) \quad w = e^{i\alpha} \cdot \frac{1}{z}, \quad \text{where } \alpha \in \mathbb{R}.$$

#### 3.3.3 Upper half plane onto upper half plane

The Möbius transformations which map the upper half plane  $\Im z \geq 0$  onto itself all have the structure

$$(92) \quad w = \frac{az + b}{cz + d}, \quad \text{where } a, b, c, d \in \mathbb{R} \text{ and } \begin{vmatrix} a & b \\ c & d \end{vmatrix} > 0.$$

#### 3.3.4 Upper half plane onto unit disc

The Möbius transformations which map the upper half plane  $\Im z \geq 0$  onto the unit disc  $|w| \leq 1$ , have all the structure

$$(93) \quad w = e^{i\alpha} \cdot \frac{z - z_0}{z - \bar{z}_0}, \quad \text{where } \alpha \in \mathbb{R} \text{ and } \Im z_0 > 0.$$

#### 3.3.5 Unit disc onto upper half plane

The Möbius transformations which map the unit disc  $|z| \leq 1$  onto the upper half plane  $\Im w \geq 0$ , have all the structure

$$(94) \quad w = \frac{\bar{w}_0 z - w_0 e^{i\alpha}}{z - e^{i\alpha}}, \quad \text{where } \alpha \in \mathbb{R} \text{ and } \Im w_0 > 0.$$

### 3.4 The transformation $w = z^2$

We now leave the Möbius transformations in order to discuss other types of conformal maps. The treatment of these is closely connected with the theory of *Riemann surfaces*. The difference is that we here work in the opposite direction. Instead of defining a topological surface in the  $w$ -space which is in bijective correspondence with (almost) all of the  $z$ -plane, we here choose to find fundamental domains in the  $z$ -plane, which (almost) is mapped bijectively and conformally onto the  $w$ -plane.

**Remark 3.4.1** Notice that compared with Chapter 2, where we dealt with the Riemann surfaces, we have here for convenience tacitly interchanged the variables  $z$  and  $w$ . In Chapter 2 we always created the Riemann surfaces in the  $z$ -space. By this interchange of variables we now have to consider topological surfaces in the  $w$ -space instead.  $\diamond$

In the applications in the subsequent sections the reader should always notice the domains in the  $z$ -space which by some given conformal map under consideration is mapped onto (almost) all of the  $w$ -plane, or even better for the *Dirichlet problem* considered later on, onto the upper half plane.

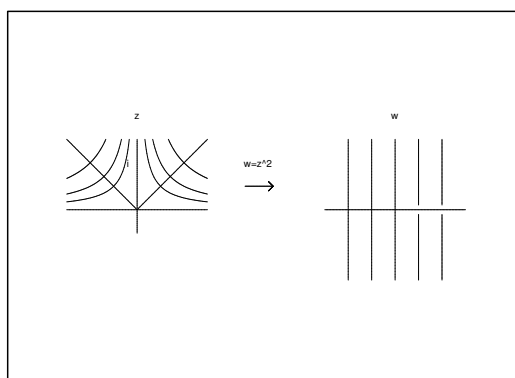


Figure 25: The open upper half plane (a fundamental domain) in the  $z$ -plane is by  $w = z^2$  mapped onto the slit  $w$ -plane.

We shall in this section treat the transformation  $w = z^2$ , which is not bijective in all of the  $z$ -plane. The map is in *polar coordinates* given by

$$\varrho = r^2, \quad \varphi = 2\Theta, \quad \text{where } z = r e^{i\Theta} \text{ and } w = \varrho e^{i\varphi}.$$

This description shows that every angular space in the  $z$ -plane of vertex 0 and angle  $\alpha > 0$  is mapped into another angular space in the  $w$ -plane of vertex 0 and angle  $2\alpha$ .

If  $0 < 2\alpha \leq 2\pi$ , i.e.  $0 < \alpha \leq \pi$ , then we can find a branch of the 2-valued square root function, such that the inverse map is uniquely defined by this branch, also denoted  $z = \sqrt{w}$ . In particular, the open upper half plane  $\Im z > 0$  is by  $w = f(z) = z^2$  mapped bijectively and conformally onto the slit plane  $\mathbb{C} \setminus (\mathbb{R}_+ \cup \{0\})$ , and the open first quadrant in the  $z$ -plane is by the same mapping mapped bijectively onto the open upper half plane in the  $w$ -plane.

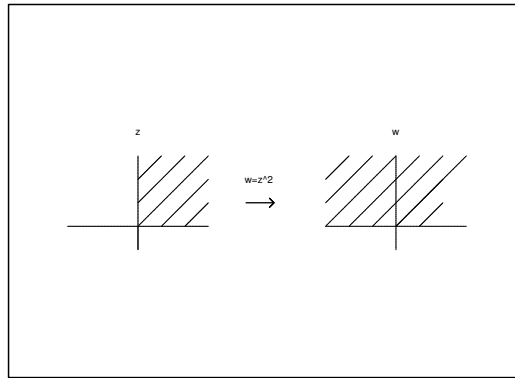


Figure 26: The open first quadrant is mapped bijectively onto the open upper half plane by the map  $w = z^2$ .

Since  $f'(z) = 2z \neq 0$  for  $\Im z > 0$ , this map is conformal in the open upper half plane.

We get from

$$w = u + iv = z^2 = (x + iy)^2 = x^2 - y^2 + 2ixy$$

that

$$u = x^2 - y^2 \quad \text{and} \quad v = 2xy.$$

Hence the curves  $u = c$  in the slit  $w$ -plane correspond to the *level curves*  $x^2 - y^2 = c$  in the open upper half plane, and the curves  $v = c$  in the slit  $w$ -plane correspond to the *level curves*  $2xy = c$  in the open upper half plane, cf. Figure 25. The systems of curves  $u = \text{const.}$  and  $v = \text{const.}$  are orthogonal. Since the map is conformal, the angles are preserved, so the level curves in the upper half plane in the  $z$ -plane must also form orthogonal systems of curves.

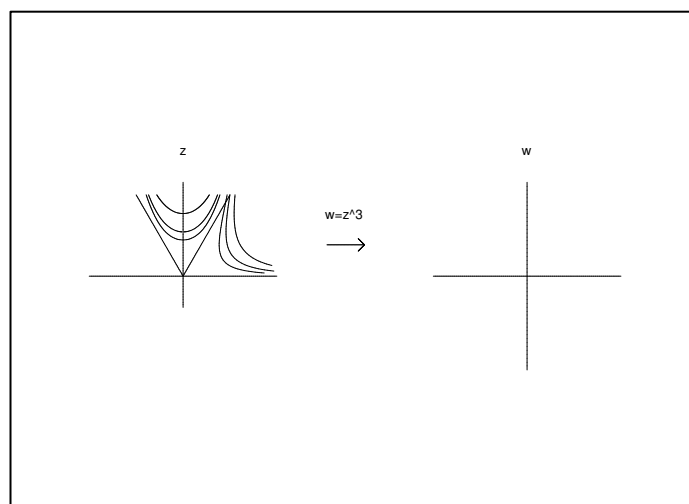


Figure 27: The principal domain (an angular space with vertex at 0) is by  $w = z^3$  mapped bijectively and conformally onto the slit  $w$ -plane.

It is easy to generalize the above to transformations of the form

$$w = f(z) = z^n, \quad n \in \mathbb{N}.$$

Using polar coordinates we see that the angles are multiplied by  $n$ . In particular, the angular space  $0 < \text{Arg } z < \frac{2\pi}{n}$  is mapped bijectively onto the slit  $w$ -plane  $\mathbb{C} \setminus (\mathbb{R}_+ \cup \{0\})$ , cf. Figure 27.

The general case of the angular space  $0 < \text{Arg } z < \frac{\pi}{\alpha}$ , where  $\alpha \geq 1$  is positive real, can be treated similarly by using the principal value of the many-valued function  $w = z^\alpha$ , so this angular space is mapped bijectively and conformally onto the open upper half plane. Thus, every angular space in the upper half plane can be mapped bijectively and conformally onto the open upper half plane.

By choosing another branch of  $\log z$  we can extend this result to any angular space in the  $z$ -plane. The simple details are left to the reader.

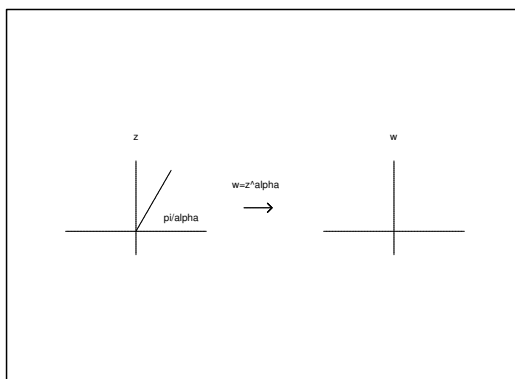


Figure 28: The angular space  $0 < \text{Arg } z < \frac{\pi}{\alpha}$ , where  $\alpha > 1$ , is by the principal value of the many-valued function  $w = z^\alpha$  mapped onto the open upper half plane in the  $w$ -plane.

### 3.5 The exponential transformation $w = \exp z$

If  $w = e^z$ , then  $\frac{dw}{dz} = e^z \neq 0$  for all  $z \in \mathbb{C}$ , so this transformation is conformal everywhere in  $\mathbb{C}$ . We apply *polar coordinates* for  $w$ , which means that we put  $w = \varrho e^{i\varphi}$ , and *rectangular coordinates* for  $z$ , i.e.  $z = x + iy$ . Then the map is described by

$$\varrho = e^x, \quad \varphi = y + 2n\pi, \quad n \in \mathbb{Z}.$$

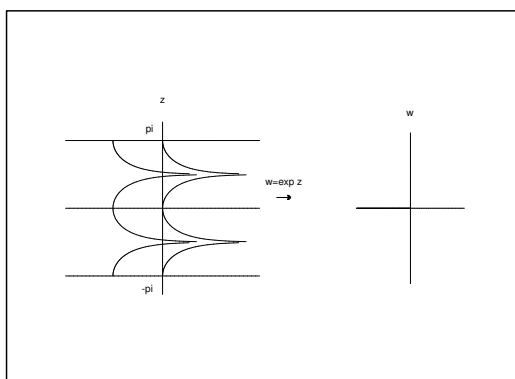


Figure 29: The exponential transform  $w = \exp z$ .

If we consider the horizontal strip in the  $z$ -plane given by  $y = \Im z \in ] - \pi, \pi[$ , cf. Figure 29, then this strip is mapped bijectively onto the slit  $w$ -plane  $\mathbb{C} \setminus (\mathbb{R}_- \cup \{0\})$  by the map

$$(95) \quad \varrho = e^x, \quad \text{for } x \in \mathbb{R}, \quad \varphi = y \in ] - \pi, \pi[,$$

which shows that the inverse map is the *principal logarithm*  $\text{Log}$ .

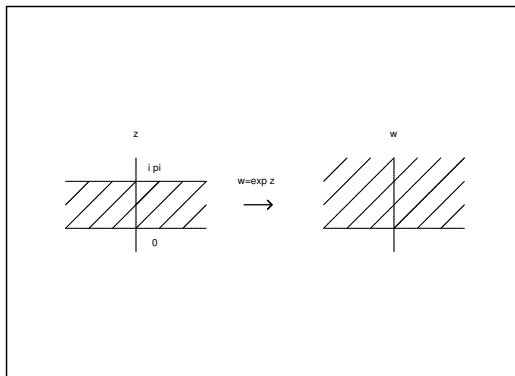


Figure 30: An horizontal strip of height  $\pi$  is mapped bijectively onto the upper half plane by the exponential transformation  $w = \exp z$ .

In particular, it follows that the infinite strip  $y = \Im z \in ]0, \pi[$  in the  $z$ -plane is mapped by  $\exp$  conformally and bijectively onto the open upper half  $w$ -plane.

Now,  $w = \rho e^{i\varphi} = u + iv$ , so the map (95) is in rectangular coordinates in the  $w$ -plane given by

$$(96) \quad u = e^x \cos y \quad \text{and} \quad v = e^x \sin by.$$



The *level curves* in the  $z$ -plane corresponding to  $u = c$  is according to (96) implicitly given by  $e^x \cos y = c$ . If  $c = 0$ , then we get the lines  $y = \pm \frac{\pi}{2}$ , and if  $c \neq 0$ , then

$$x = \ln\left(\frac{c}{\cos y}\right), \quad \text{where } \begin{cases} y \in ]-\frac{\pi}{2}, \frac{\pi}{2}[ , & \text{if } c > 0, \\ y \in ]-\pi, \frac{\pi}{2}[ \cup ]\frac{\pi}{2}, \pi[ , & \text{if } c < 0. \end{cases}$$

Similarly, the *level curves* in the  $z$ -plane corresponding to  $v = c$  is implicitly given by  $e^x \sin y = 0$ . If  $c = 0$ , then  $y = 0$ , and if  $c \neq 0$ , then

$$x = \ln\left(\frac{c}{\sin y}\right) \quad \text{for } \begin{cases} y \in ]0, \pi[ , & \text{when } c > 0, \\ y \in ]-\pi, 0[ , & \text{when } c < 0. \end{cases}$$

Since the map is conformal, the two systems of level curves are orthogonal to each other in the  $z$ -plane.

### 3.6 The sine transformation $w = \sin z$

We consider  $w = \sin z$  in the vertical strip  $x = \Re z \in ]-\frac{\pi}{2}, \frac{\pi}{2}[$ . Since  $\frac{d}{dz} \sin z = \cos z \neq 0$  in this strip, the map is conformal.

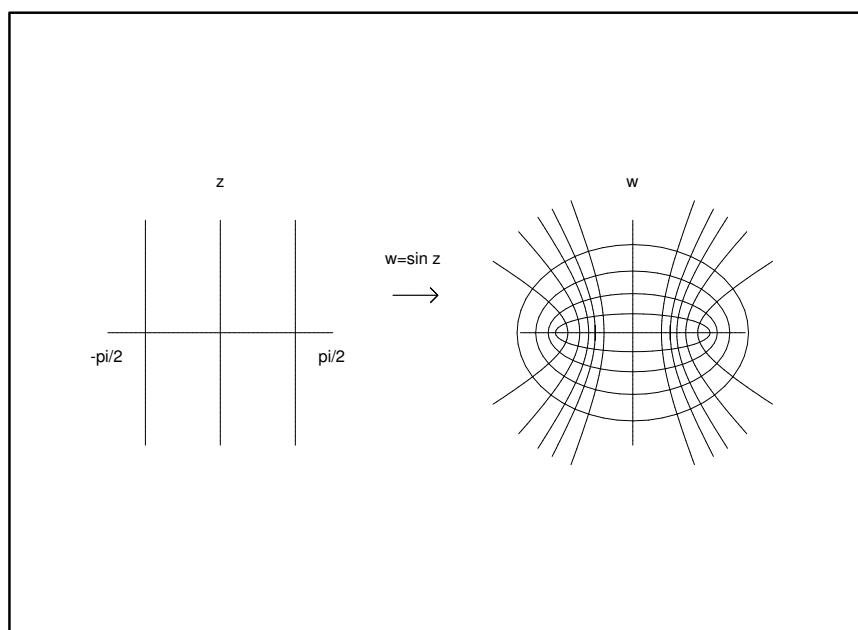


Figure 31: The vertical strip of breadth  $\pi$  is mapped onto a slit  $w$ -plane.

We shall show that this strip is mapped bijectively onto a slit  $w$ -plane, where the cuts are the real  $u$ -intervals  $]-\infty, -1[$  and  $[1, +\infty[$ . It suffices to prove that every curve of the form  $z = t + iy_0$ , where

$y_0 \in \mathbb{R}$  is fixed, and  $t \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$  is the variable, is mapped bijectively onto a curve in the slit  $w$ -plane, and that every point in the slit  $w$ -plane lies on precisely one of these curves.

We then split the sine transformation into its real and imaginary parts,

$$w = \sin z = \sin x \cosh y + i \cos x \sinh y = u + iv.$$

If  $y = y_0$  is kept fixed, i.e. we consider the horizontal line segment in Figure 31, the the image curve in the  $w$ -plane has the parametric description

$$(97) \quad u(t) = \cosh y_0 \cdot \sin t, \quad v(t) = \sinh y_0 \cdot \cos t, \quad \text{for } t \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[.$$

If  $y_0 = 0$ , then  $v = 0$ , and the interval  $] -1, 1[$  on the  $u$ -axis in the  $w$ -plane is traversed precisely once.

Then notice that  $v$  and  $y_0$  always have the same sign, so it is enough to assume that  $y_0 > 0$ , because we can just reflect these curves in the real axis in order to get the remaining curves. Then eliminate  $t$  of (97) to get that the line segment at height  $y_0 > 0$  is mapped into an half ellipse in the upper  $w$ -plane of the equation

$$(98) \quad \frac{u^2}{\cosh^2 y_0} + \frac{v^2}{\sinh^2 y_0} = 1,$$

which has the half axes  $\cosh y_0$  and  $\sinh y_0$ .

Let  $(u_0, v_0)$  be any point in the upper half  $w$ -plane. We shall prove that there is precisely one  $y_0 \in \mathbb{R}_+$ , such that  $(u_0, v_0)$  lies on the curve given by (98).

Since  $v_0 > 0$ , we define a continuous and strictly decreasing function by

$$\Phi(y) = \frac{u_0^2}{\cosh^2 y} + \frac{v_0^2}{\sinh^2 y}, \quad y > 0.$$

It follows from

$$\lim_{y \rightarrow 0^+} \Phi(y) = +\infty \quad \text{and} \quad \lim_{y \rightarrow +\infty} \Phi(y) = 0,$$

that there exists precisely one  $y_0 \in \mathbb{R}_+$ , such that  $\Phi(y_0) = 1$ , and the claim is proved.

The open half strip, given by  $x \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ ,  $y > 0$ , is by  $w = \sin z$  mapped bijectively and conformally onto the open upper half  $w$ -plane.

Finally, we shall find the *level curves* in the  $z$ -plane of the map  $w = \sin z$ , corresponding to  $u = c$ , or  $v = c$ . It follows from (97) that these are implicitly given by

$$u = c = \sin x \cosh y \quad \text{and} \quad v = c = \cos x \sinh y,$$

where  $x \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$  and  $y \in \mathbb{R}$ .

If  $u = 0$ , then  $x = 0$ , i.e. the imaginary axis.

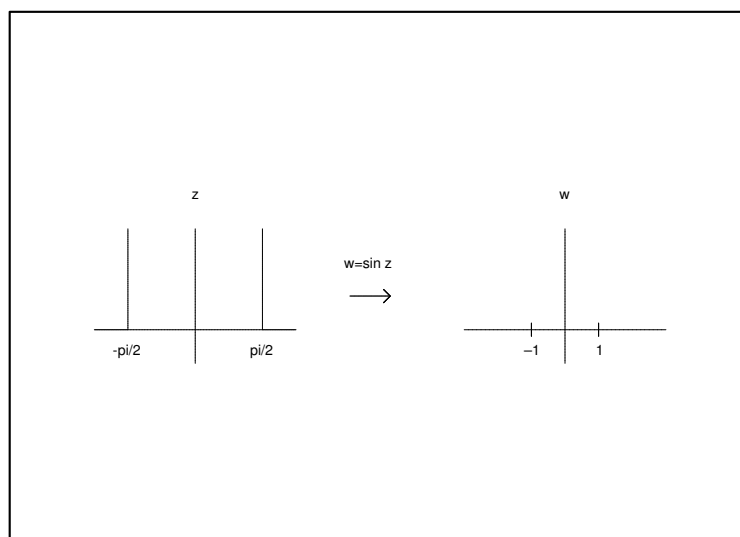


Figure 32: An half strip is mapped bijectively onto the open upper half plane by the sine transformation.

If  $u = c \neq 0$ , then

$$x = \operatorname{Arccos} \left( \frac{c}{\cosh y} \right), \quad \text{where } \begin{cases} |y| > \ln(|c| + \sqrt{c^2 - 1}) & \text{for } |c| > 1, \\ y \in \mathbb{R} & \text{for } |c| < 1. \end{cases}$$

If  $v = 0$ , then  $y = 0$ , i.e. the line segment  $\left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$  on the  $x$ -axis.

If  $v = c \neq 0$ , then

$$y = \ln \left( \frac{c}{\cos x} + \sqrt{1 + \frac{c^2}{\cos^2 x}} \right) \quad \text{for } x \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[.$$

### 3.7 Zhukovski's function

*Zhukovski's function* (in some books spelled *Joukowski's function* in a mixture of French and German transcription of the Russian name) is defined by

$$(99) \quad w = f(z) := \frac{1}{2} \left\{ z + \frac{1}{z} \right\}, \quad \text{for } z \in \mathbb{C} \setminus \{0\}.$$

It follows from  $f'(z) = \frac{1}{2} \{1 - z^{-2}\}$  that *Zhukovski's function* is conformal in  $\mathbb{C} \setminus \{-1, 0, 1\}$ .

The inverse of *Zhukovski's function* is found by solving the following equation of second degree,

$$(100) \quad z^2 - 2wz + 1 = 0$$

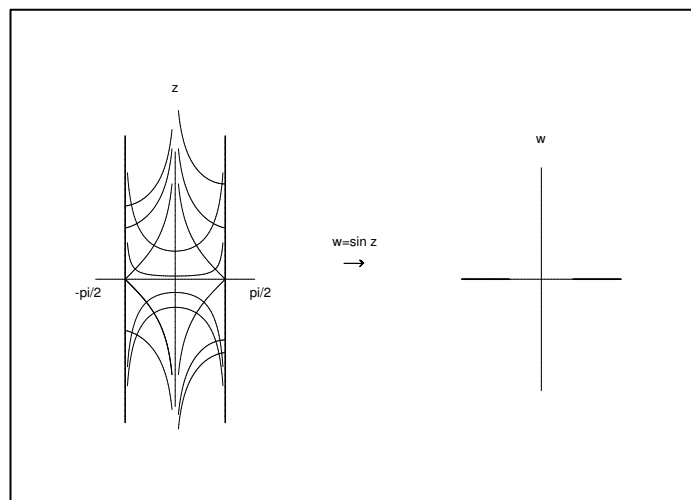


Figure 33: The level curves of the sine transformation.

with respect to  $z$ . There will for every  $w$  value usually, in fact, for every  $w \neq \pm 1$ , correspond two  $z$ -values,  $z_1$  and  $z_2$ .

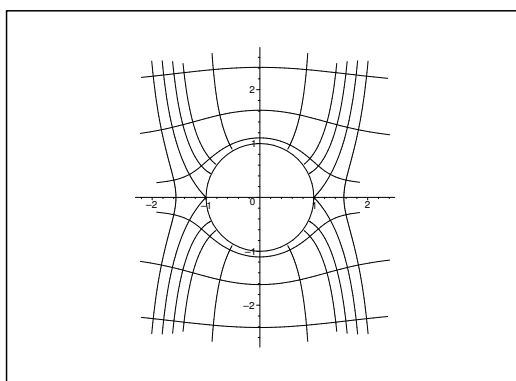


Figure 34: Level curves of Zhukovskii's function defined on the complementary set of the unit disc.

Using that the product  $z_1 \cdot z_2 = 1$ , cf. (100), it follows that if  $|z_1| > 1$ , then  $|z_2| < 1$ , and *vice versa*. Therefore, these two domains are both mapped bijectively onto a slit  $w$ -plane.

Then notice that (100) can be solved for *every*  $w \in \mathbb{C}$ , so the exceptional set in the  $w$ -plane must be given by the image of the unit circle in the  $z$ -plane. This unit circle is in polar coordinates described by  $z = e^{i\Theta}$ ,  $\Theta \in [0, 2\pi]$ , thus

$$w = \frac{1}{2} \{ e^{i\Theta} + e^{-i\Theta} \} = \cos \Theta, \quad \text{for } \Theta \in [0, 2\pi],$$

so the boundary curve is mapped (twice) onto the interval  $[-1, 1]$  on the real  $w$ -axis.

The image of the set  $|z| > 1$  (as well of the set  $|z| < 1$ ) is by *Zhukovski's function* the slit complex  $w$ -plane with the exception of the cut along the segment  $[-1, 1]$  on the real  $u$ -axis.

When we split into real and imaginary parts, we get

$$u = \frac{1}{2} \left\{ x + \frac{x}{x^2 + y^2} \right\} \quad \text{and} \quad v = \frac{1}{2} \left\{ y - \frac{y}{x^2 + y^2} \right\},$$

from which we get the level curves in  $|z| > 1$ . These are in polar coordinates given by

$$\frac{r^2 + 1}{2r} \cdot \cos \Theta = u_0 \quad \text{and} \quad \frac{r^2 - 1}{2r} \cdot \sin \Theta = v_0,$$

where  $u_0$ , resp.  $v_0$ , are constants. The *equipotential curves*  $u = u_0 = \text{constant}$  are given by the equation

$$r = \frac{u_0 \pm \sqrt{u_0^2 - \cos^2 \Theta}}{\cos \Theta} \quad \text{for } |\cos \Theta| < |u_0|,$$

where the sign in  $\pm$  is always chosen such that  $r > 1$ .

Similarly we find for the *streamlines*  $v = v_0$  the equation in polar coordinates,

$$(101) \quad r = \frac{v_0}{\sin \Theta} + \sqrt{1 + \frac{v_0^2}{\sin^2 \Theta}}, \quad \text{where either } \Theta \in ]0, \pi[ \quad \text{or} \quad \Theta \in ]\pi, 2\pi[.$$

Notice that we have only  $+$  in front of the square root which is taking with its positive value. The interpretation of (101) is that it describes the streamlines of a *stationary potential flow* around an obstacle, which here is the unit disc. Such patterns of flow are actually found in practice, where one only requires that the *Reynold's number* is sufficiently small. When *Reynold's number* becomes too big, then the flow becomes unstable, and we get some unwanted eddies behind the disc seen in the direction of the flow.

Zhukovski used as early as in 1911 the inverse map of  $w = \frac{1}{2} \left\{ z + \frac{1}{z} \right\}$  as discussed above together with another complex potential to describe the flow past the wing of an aeroplane.

First notice that if the friction of the air is very small, then we can replace any streamline by the boundary of a rigid body without disturbing the flow, which will just glide along this body. In the present case the body is represented by the unit disc.

Zhukovski's clever idea was first to consider the many-valued function

$$(102) \quad w := \frac{k}{2\pi i} \log z = \frac{k}{2\pi} \arg z + i \frac{k}{2\pi} \ln r = u + iv,$$

where  $k$  is a real constant. Its derivative is single-valued,

$$(103) \quad \frac{dw}{dz} = \frac{k}{2\pi i} \frac{1}{z},$$

so we have uniquely defined a *velocity field* of some flow. The *streamlines* are given by  $v = v_0$  constant, thus by (102),

$$\frac{k}{2\pi} \ln r = v_0, \quad \text{or} \quad r = \exp\left(\frac{2\pi v_0}{k}\right) = \text{a positive constant},$$

so the streamlines of (102) are circles of centre 0. In particular, the unit circle is one of its streamlines, so we can combine (102) with *Zhukovski's function* to get the (many-valued) complex potential

$$(104) \quad w = \frac{1}{2} \left\{ z + \frac{1}{z} \right\} - \frac{k}{2\pi i} \log z, \quad |z| > 0.$$

This potential function also represents a flow around the unit disc. More precisely, its streamlines represents a distorted translation flow past the unit disc. In polar coordinates the stream function is given by

$$v = \Im \left\{ \frac{1}{2} \left\{ r e^{i\Theta} + \frac{1}{r e^{i\Theta}} \right\} - \frac{k}{2\pi i} \log z \right\} = \frac{1}{2} \left\{ r - \frac{1}{r} \right\} \sin \Theta + \frac{k}{2\pi} \ln r,$$

so the streamlines given by  $v = v_0$  constant must fulfil the equation

$$\sin \Theta = \frac{2r}{r^2 - 1} \left\{ v_0 - \frac{k}{2\pi} \ln r \right\}.$$

We shall never obtain  $r = 1$ , if  $v_0 \neq 0$ . However, if  $v_0 = 0$ , then

$$\sin \Theta = -\frac{k}{2\pi} \ln r \frac{2r}{r^2 - 1} = -\frac{k}{\pi} \cdot \frac{r}{r + 1} \cdot \frac{\ln r}{r - 1} \rightarrow -\frac{k}{2\pi} \quad \text{for } r \rightarrow 1 + .$$

Therefore, if  $0 < k < 2\pi$ , then precisely two streamlines will reach the unit circle, namely at the two angles  $\Theta$  for which  $\sin \Theta = -\frac{k}{2\pi} \in ]-1, 0[$ . Both these points on the unit circle lie in the lower half plane, so the flow will give a thrust upwards on the unit disc, cf. Figure 35.

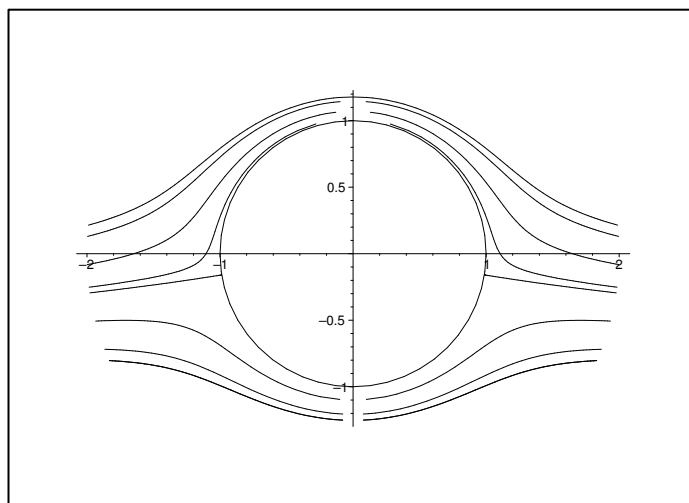


Figure 35: Streamlines for  $k = 1$  of the many-valued potential

$$w = \frac{1}{2} \left\{ z + \frac{1}{z} \right\} - \frac{1}{2\pi i} \log z.$$

Of course the unit disc does not look like the profile of a wing of an aeroplane. We still need to apply *Riemann's mapping theorem*, which is too difficult to be given here in all details. The considerations above may already seem rather difficult.

### 3.8 The Schwarz-Christoffel transformation

Conformal maps are often used when we want to transform a Dirichlet problem of the Laplace differential equation from Physics into a simpler Dirichlet problem in the open half plane, or the unit disc, where in both cases solution formulæ are known. Here we mention that conformal maps leave the Laplace differential operator  $\Delta$  invariant, which will be proved in Section 3.9. This observation reduces the problem to that we shall “only” consider the problem of mapping bijectively a simply connected domain into e.g. the upper half plane by a conformal map. The *existence* of such a conformal map follows from

**Theorem 3.8.1** The Riemann mapping theorem. *Let  $\Omega \subset \mathbb{C}$  be a simply connected open domain which is not all of  $\mathbb{C}$ ,  $\Omega \neq \mathbb{C}$ . Then there exists a bijective and conformal map  $f : \Omega \rightarrow \{w \in \mathbb{C} \mid \Im w > 0\}$  of  $\Omega$  onto the upper half plane.*

This theoretical result is very deep, and its proof is about three pages long. As it does not give us any clue of how in practice to *construct* such a bijective and conformal map, we shall skip the proof here.

It was hinted in the paragraph above that it may be exceedingly difficult to find such a map, and it surely is. Most of the known cases have already been treated in Sections 3.3–3.6 above. There is still another class of simply connected open domains, for which we can explicitly construct such a conformal map using an explicit solution formula, namely domains bounded by a polygon having a finite number of vertices, one or more of which may lie at infinity.



**Theorem 3.8.2** The Schwarz-Christoffel formula (1867-1869). Given  $n$  real points on the  $x$ -axis,  $x_1 < x_2 < \dots < x_n$ , and corresponding  $n$  real numbers  $a_1, a_2, \dots, a_n$ , fulfilling

$$|a_j| < 1 \quad \text{for all } j = 1, \dots, n, \quad \text{and} \quad a_1 + a_2 + \dots + a_n = 2.$$

The Schwarz-Christoffel transformation  $F$  is given by

(105)

$$F(z) := A \int_{z_0}^z \frac{d\zeta}{(\zeta - x_1)^{a_1} (\zeta - x_2)^{a_2} \dots (\zeta - x_n)^{a_n}} + B, \quad \text{for } A \in \mathbb{C} \setminus \{0\} \text{ and } B \in \mathbb{C} \text{ constants,}$$

where we define  $(\zeta - x_j)^{a_j} := \exp(a_j \operatorname{Log}^*(\zeta - x_j))$  for  $\zeta$  lying in the closed upper half plane with the exception of  $\zeta = x_j$ , and where  $\operatorname{Log}^*(z)$  is any branch of the logarithm, for which  $\operatorname{Log}^*(x) = \ln x$  for  $x > 0$  and which has its branch cut from 0 to  $\infty$  in the (open) lower half plane.

- The Schwarz-Christoffel transformation (105) maps the real axis in the  $z$ -plane onto a closed curve consisting of line segments in the  $w$ -plane.
- The vertices in the  $w$ -plane are the points

$$w_1 = F(x_1), w_2 = F(x_2), \dots, w_n = F(x_n),$$

and at a vertex  $w_j = F(x_j)$  the tangent (seen in the sense of direction of the image curve) is rotated the angle  $a_j\pi$ .

- If the closed curve does not cut itself, it is a polygon  $P$  in the  $w$ -plane, and (105) is a bijective and conformal map of the upper half  $z$ -plane  $\Im z > 0$  onto the domain  $\Omega$  which has the polygon  $P$  as its boundary.

SKETCH OF PROOF. The proof of this theorem is not an easy task either, and we shall here not give a correct one. However, a couple of comments may make the result plausible. First notice the roles of the constants  $A \in \mathbb{C} \setminus \{0\}$  and  $B \in \mathbb{C}$  in (105). The constant  $A$  defines a stretching and a rotation, and the constant  $B$  a translation. Neither of these operations will affect the size of the rotation of the tangents at the vertices. The core of (105) is therefore (of course) the simpler map

$$(106) \quad w = F_0(z) := \int_{z_0}^z \frac{d\zeta}{(\zeta - x_1)^{a_1} \dots (\zeta - x_n)^{a_n}}.$$

When (106) is differentiated, we get

$$(107) \quad \frac{dw}{dz} = \frac{1}{(\zeta - x_1)^{a_1} \dots (\zeta - x_n)^{a_n}}, \quad |a_j| < 1 \text{ and } a_1 + \dots + a_n = 2.$$

When  $z = x \in \mathbb{R}$  and  $x < x_1$  ( $x_1 < x_2 < \dots < x_n$ ), then

$$a_j \operatorname{Arg}^*(x - x_j) = a_j\pi \quad \text{for all } j = 1, \dots, n,$$

hence

$$\arg \left( \frac{1}{(\zeta - x_1)^{a_1} \dots (\zeta - x_n)^{a_n}} \right) = -a_1\pi - \dots - a_n\pi + 2p\pi = 2(p-1)\pi, \quad p \in \mathbb{Z}, \quad \text{for } x < x_1,$$

so it is real and positive,

$$\frac{1}{(\zeta - x_1)^{a_1} \cdots (\zeta - x_n)^{a_n}} > 0 \quad \text{for } x < x_1.$$

Then we let  $x$  pass through  $x_1$ , so we consider  $x_1 < x < x_2 < \cdots < x_n$ , in which case

$$a_1 \operatorname{Arg}^*(x - x_1) = 0 \quad \text{and} \quad a_j \operatorname{Arg}^*(x - x_j) = a_j \pi \quad \text{for all } j = 2, \dots, n,$$

hence

$$\arg \left( \frac{1}{(\zeta - x_1)^{a_1} \cdots (\zeta - x_n)^{a_n}} \right) = -a_2 \pi - \cdots - a_n \pi + 2p\pi = 2(p-1)\pi + a_1 \pi, \quad p \in \mathbb{Z}, \quad \text{for } x_1 < x < x_2,$$

which shows that the tangent has been rotated the angle  $a_1 \pi$ , when we pass through  $x = x_1$  on the real line.

In the next step we consider  $x_1 < x_2 < x < x_3 < \cdots < x_n$ , and conclude likewise that the tangent is rotated the angle  $a_2 \pi$  when  $x$  passes through the singular point  $x_2$  of integration. In this way we proceed and find that the rotation of the tangent is  $a_j \pi$  at the image point  $w_j$  of the singular point  $x_j$ .

Choose for convenience  $z_0 = x_0 \in \mathbb{R}$ , such that  $x_0 < x_1$  in (106). The assumption that all  $|a_j| < 1$  assures that the line integral along the real axis is convergent. The fixed direction of the image curve segment for  $x_j < x < x_{j+1}$  must indeed consist of line segments, mutually forming angles of the size  $a_j \pi$  seen in the sense of orientation, and the condition  $a_1 + \cdots + a_n = 2$  assures that the total rotation of the tangents along the curve adds up to  $a_1 \pi + \cdots + a_n \pi = 2\pi$ . Since the real axis supplemented with the infinite point in this case is considered as a closed curve, the same is true for the image curve, so if none of the points of the image curve is  $\infty$ , then the image curve is indeed a closed curve. If instead the image curve somewhere tends to  $\infty$ , then it must necessarily also leave  $\infty$ , so we get an extended closed image curve by adding  $\infty$  in this case.

If the curve does not cut itself, then it is a polygon  $P$ , and (106) must necessarily map the upper half plane onto the interior of  $P$ . The map of  $\mathbb{R} \cup \{\infty\}$  onto  $P$  is clearly bijective, forcing the map of the open upper half plan also to be bijective onto the interior of  $P$  (this claim requires a larger proof), and it is an analytic function with  $\frac{dw}{dz}$  given by (107)  $\neq 0$ , so it is conformal. ( $\square$ )

The above is not a fully correct proof, but it gives some guidelines and also shows that the results are reasonable.

We still have a problem in the construction. All proofs focus on the rotation of the tangents, thus fixing the constants  $a_1, \dots, a_n$ , which define the rotation of the tangents at the points

$$w_1 = F(x_1), \dots, w_n = F(x_n).$$

The problem is:

*Given  $w_1, \dots, w_n \in \mathbb{C}$ , which are the vertices in the given order of a polygon in the  $w$ -plane. They surely define the rotations  $a_j \pi$  of all tangents, which can be found by analyzing the figure. However, we still have to find the real constants  $x_1 < x_2 < \cdots < x_n$ , such that  $w_j = F(x_j)$  for  $j = 1, \dots, n$ .*

The literature does not give many hints in this direction, so usual one must use trial and error on a computer.

Another drawback of the formula (106) is that the integral may be very difficult to compute explicitly.

Theorem 3.8.2 can be extended to also allow  $a_j = \pm 1$ . If  $a_j = 1$ , then the image point is put equal to  $\infty$ , because the integral is divergent. We shall show this in Example 3.8.2 below, but first we consider a simpler (infinite) example in Example 3.8.1.

Notice also that the *streamlines* are trivial in the upper  $z$ -plane, namely the lines which are parallel to the  $x$ -axis, i.e. given by  $y = y_0 > 0$  constant. The *streamlines* in the  $w$ -space are then given by the curve with the parametric description,

$$(u(x, y_0), v(x, y_0)), \quad \text{for } x \in \mathbb{R}, \text{ and } y = y_0 > 0 \text{ a constant.}$$

**Example 3.8.1** Given the semi-infinite strip in the  $w$ -plane shown in Figure 36. By definition, the strip lies always to the left of the boundary curve, thus we have uniquely defines a sense of orientation of the boundary curve. It follows immediately from this analysis that the rotation of the tangents are  $a_1\pi = \frac{\pi}{2}$  at  $w_1 = i\pi$ , and  $a_2\pi = \frac{\pi}{2}$  at  $w = 0$ , thus  $a_1 = a_2 = \frac{1}{2}$ .

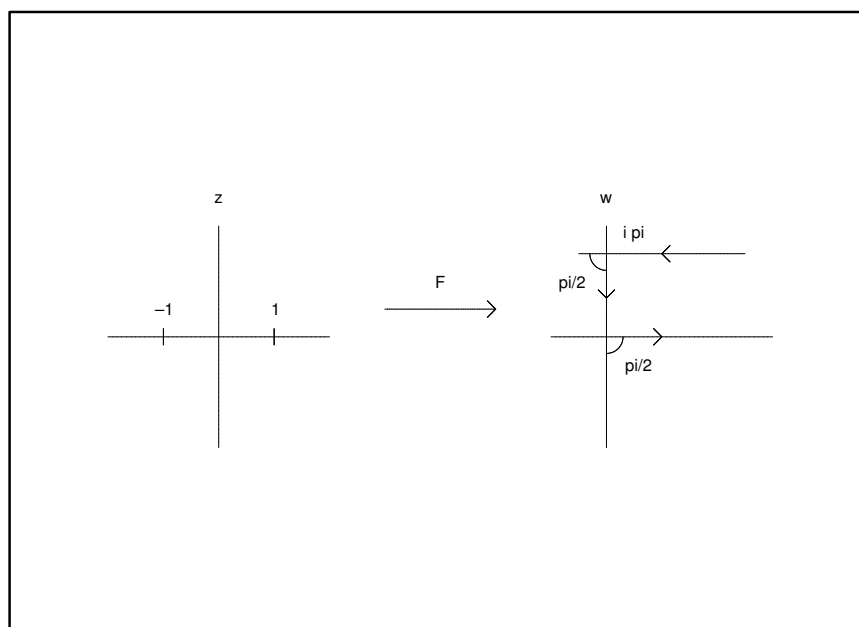


Figure 36: Analysis of the Schwarz-Christoffel transformation of Example 3.8.1.

Choose  $x_1 = -1$  and  $x_2 = 1$ . Then the required transformation is defined by

$$\frac{dw}{dz} = A \cdot \frac{1}{(z+1)^{\frac{1}{2}}(z-1)^{\frac{1}{2}}} = \frac{A}{\sqrt{z^2-1}},$$

thus by an integration,

$$w = A \int^z \frac{d\zeta}{\sqrt{\zeta^2 - 1}} = A \cdot \operatorname{Arcosh} z + B.$$

Since  $w = 0$  corresponds to  $z = 1$ , we get

$$0 = A \cdot \operatorname{Arcosh} 1 + B = B, \quad \text{thus } B = 0.$$

Furthermore,  $w = i\pi$  corresponds to  $z = -1$ , so

$$i\pi = A \cdot \operatorname{Arcosh}(-1) = A \cdot i\pi, \quad \text{thus } A = 1,$$

and the required transformation is then

$$w = \operatorname{Arcosh} z \quad \text{or} \quad z = \cosh w.$$

From the latter follows that

$$z = x + iy = \cosh u \cdot \cos v + i \cdot \sinh u \cdot \sin v,$$

so

$$x = \cosh u \cdot \cos v \quad \text{and} \quad y = \sinh u \cdot \sin v.$$

The *streamlines* are given by  $y = y_0$  constant, i.e. implicitly by

$$\sinh u \cdot \sin v = y_0 > 0,$$

or explicitly, where  $u$  is given as a function of  $v$ ,

$$u = \operatorname{Arsinh} \left( \frac{y_0}{\sin v} \right) = \ln \left( \frac{y_0}{\sin v} + \sqrt{1 + \frac{y_0^2}{\sin^2 v}} \right), \quad v \in ]0, \pi[.$$

It is easily seen that

$$u(v) \geq u \left( \frac{\pi}{2} \right) = \ln \left( y_0 + \sqrt{1 + y_0^2} \right). \quad \diamond$$

**Example 3.8.2** We consider the slit upper half  $w$ -plane of Figure 37. First find the orientation of the boundary curve such that the domain always lies to the left of it. It follows that the cut at height  $i\pi$  is run through twice, giving a rotation of the tangent at  $w = i\pi$  of the amount of  $a_1\pi = -\pi$ , i.e.  $a_1 = -1$ , and at  $B: w = \infty$ , of the amount of  $a_2\pi = +\pi$ , because it “continues to the left” on the  $u$ -axes at  $\infty$  in the reverse direction, cf. Figure 37, so  $a_2 = +1$ .

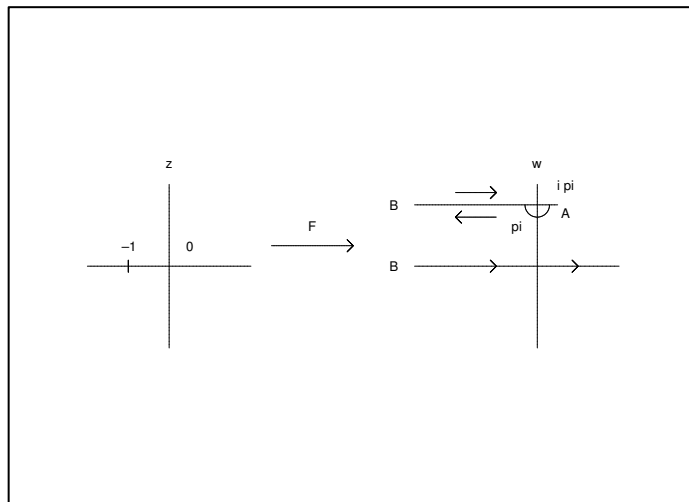


Figure 37: Analysis of the Schwarz-Christoffel transformation of Example 3.8.2.

We choose  $x_1 = -1$  and  $x_2 = 0$ , by which choice the Schwarz-Christoffel transformation must have the structure

$$\frac{dw}{dz} = A \cdot \frac{1}{(z+1)^{(-1)}z^1} = A \cdot \frac{z+1}{z} = A \cdot \left\{ 1 + \frac{1}{z} \right\},$$

from which we get by an integration,

$$(108) \quad w = A \cdot \{z + \operatorname{Log} z\} + B, \quad \Im z > 0.$$

We shall find the constants  $A = A_1 + A_2$  and  $B = B_1 + iB_2$ . We first get by taking the imaginary part of (108) that

$$\begin{aligned} v &= \Im \{(A_1 + iA_2)(x + iy + \ln|z| + i \operatorname{Arg} z) + (B_1 + B_2)\} \\ (109) \quad &= A_1 y + A_2 x + A_2 \ln|z| + A_1 \cdot \operatorname{Arg} z + B_2. \end{aligned}$$

Use that if  $w \rightarrow \infty$  on the line  $v = \pi$ , where  $u \rightarrow -\infty$ , then the image  $z = x \in \mathbb{R}_-$  is real and negative, so  $y = 0$ , and furthermore,  $x \rightarrow 0$ . Hence,  $\operatorname{Arg} z = \pi$  through this limit process, and it follows from (109), using that  $y = 0$ ,

$$\pi = \lim_{x \rightarrow 0^-} \{A_1 \cdot 0 + A_2 \cdot x + A_2 \ln|x| + A_1 \cdot \pi + B_2\} = A_1 \pi + B_2 + \lim_{x \rightarrow 0^-} A_2 \cdot \ln|x|.$$

Because  $\ln|x| \rightarrow -\infty$  for  $x \rightarrow 0^-$ , this is only possible, if  $A_2 = 0$ . We get that

$$(110) \quad \pi = A_1 \pi + B_2.$$

If instead  $w \rightarrow \infty$  along the line  $v = 0$  and  $u \rightarrow -\infty$ , then the image  $z = x \in \mathbb{R}_+$  is positive real, so  $y = 0$  and  $x \rightarrow 0+$  through this limit process, and therefore  $\operatorname{Arg} z = 0$ . Returning to (109) we have

$$0 = \lim_{x \rightarrow 0^+} \{A_1 \cdot 0 + A_1 \cdot 0 + B_2\} = B_2,$$

so  $B_2 = 0$ , and using (110) we get  $A_1 = 1$ . Thus (108) is reduced to

$$(111) \quad w = z + \operatorname{Log} z + B_1.$$

Finally, the point  $w = i\pi$  is mapped into  $z = -1$ , so it follows from (111) that

$$i\pi = -1 + i\pi + B_1,$$

hence  $B_1 = 1$ , and the required function is

$$(112) \quad w = z + \operatorname{Log} z + 1.$$

When we take the real and imaginary parts of (112) we get

$$u = x + \ln \sqrt{x^2 + y^2} + 1 \quad \text{and} \quad v = y + \operatorname{Arg} z = y + \operatorname{Arccot} \frac{x}{y}.$$

The *streamlines* in the  $z$ -plane are trivially given by  $y_0 > 0$  constant, hence a parametric description of the *streamlines* in the  $w$ -space is for given  $y = y_0 > 0$  given by

$$(113) \quad (u, v) = \left( t + \ln \sqrt{t^2 + y_0^2} + 1, y_0 + \operatorname{Arccot} \frac{t}{y_0} \right), \quad \text{for } t \in \mathbb{R} \text{ and } y_0 > 0 \text{ a constant,}$$

cf. Figure 38.

Note that even this rather simple example gives a lot of computation, when one wants to find all constants.  $\diamond$

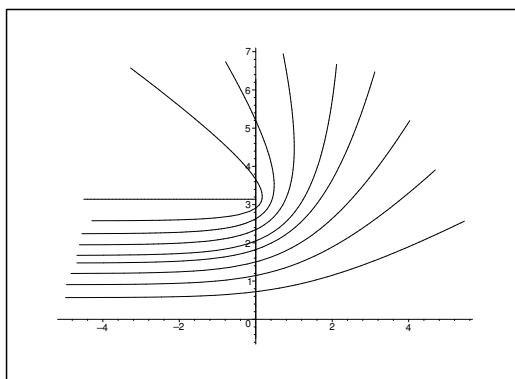


Figure 38: Some typical streamlines the Schwarz-Christoffel transformation of Example 3.8.2 based on (113). They can be interpreted as the streamlines for the flow of an ideal incompressible fluid from an infinite straight channel.

### 3.9 The Dirichlet problem

Some very important problems in Physics and the technical sciences deal with stationary fields and their equipotential surfaces (or lines) and field lines. In case of two-dimensional fields this theory was already mentioned in *Ventus, Complex Functions Theory a-2*, where it was modeled by analytic functions. In fact, if the analytic function  $w = f(z)$  is given as a model, we first split it into its real and imaginary parts,

$$(114) \quad w = f(z) = u(x, y) + i \cdot v(x, y), \quad z \sim (x, y) \in \Omega,$$

where both functions  $u$  and  $v$  are harmonic, i.e.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{and} \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0, \quad \text{for } (x, y) \in \Omega.$$

Given (114) as the model, then the *equipotential curves* are given by

$$u(x, y) = c_1, \quad \text{where } c_1 \in \mathbb{R} \text{ is an arbitrary constant,}$$

and the imaginary part  $v(x, y)$  is called the *stream function*, and the *streamlines* are given by

$$v(x, y) = c_2, \quad \text{where } c_2 \in \mathbb{R} \text{ is an arbitrary constant.}$$

Furthermore, with the exception of possible “singular points”, where  $f'(z) = 0$ , or at isolated boundary points, where  $f(z)$  is not defined, the system of all equipotential curves is always perpendicular to the system of all streamlines.

The important thing here is therefore to solve the *Laplace differential equation*

$$(115) \quad \Delta \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0, \quad \text{where for short } \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

One way of solving (115) is of course to find an analytic map  $w = f(z)$ , such that  $\psi = \Re f(z)$ , or  $\psi = \Im f(z)$ . However, without some further help this method does not look too promising to be

recommended as standard, although it can sometimes be applied. Another way is to find a direct *real* solution formula for some well-defined elementary domains. We shall in this section and in the following one show how this is done. We first make precise the problem, we are going to solve.

**Definition 3.9.1** *Given an open and simply connected domain  $\Omega$  of non-degenerated boundary  $\Gamma = \partial\Omega$  without double points. The Dirichlet problem on  $\Omega$  is defined as the following boundary value problem:*

*Given a real and piecewise continuous function  $h(x, y)$  on the boundary  $\Gamma$  of  $\Omega$ , find an harmonic function  $u(x, y)$  defined in  $\Omega$  which is equal to  $h(x, y)$  on the boundary at every point of continuity of  $h(x, y)$ , i.e.*

$$(116) \quad \left\{ \begin{array}{ll} \Delta u = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = 0 & \text{for } (x, y) \in \Omega, \\ \lim_{\substack{(x, y) \rightarrow (x_0, y_0) \\ (x, y) \in \Omega}} u(x, y) = h(x_0, y_0), & \text{if } h \text{ is continuous on } \Gamma \text{ at } (x_0, y_0) \end{array} \right.$$

In order to make the theory as simple as possible we have in Definition 3.9.1 required that  $\Omega$  is simply connected, even it is also makes sense to formulate the Dirichlet problem for domains which are not simply connected. We are of course here aiming at reasonable (though not trivial) solution formulæ.

Notice that even if  $\Omega$  is simply connected and its boundary  $\Gamma$  is “nice”, we may still find examples where  $\Gamma$  consists of several curves. One such example is given in Figure 39.



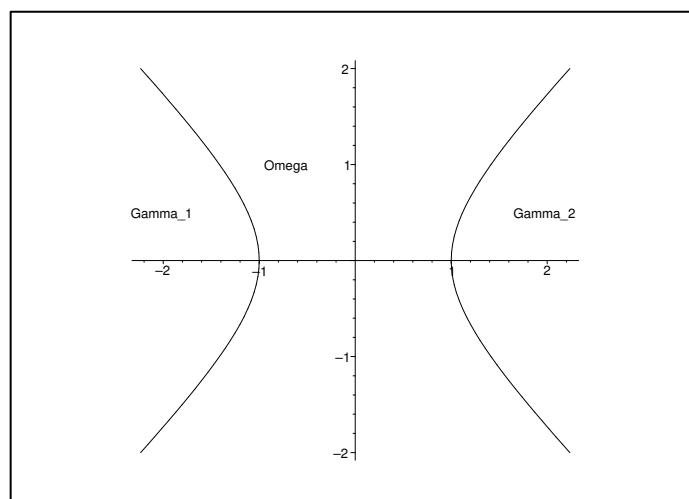


Figure 39: The domain  $\Omega$  is simply connected, and the boundary  $\Gamma$  consists of two curves  $\Gamma_1$  and  $\Gamma_2$ .

The following theorem connects the previous theory on conformal maps with the solution procedure of the Dirichlet problem. It was actually the reason why the conformal maps were treated so thoroughly in the previous sections, although they also are of independent interest.

**Theorem 3.9.1** *The composition of an harmonic function  $u(x, y)$  and a conformal map  $\Phi(\xi, \eta)$  is an harmonic function  $u \circ \Phi$  in the new variables.*

PROOF. By assumption,  $u(x, y)$  is harmonic, so to every given  $(x, y) \in \Omega$  we can find  $r > 0$ , such that  $\omega := B((x, y), r) \subseteq \Omega$ , and an analytic function  $f$  defined on  $\omega$ , such that

$$u(x, y) = \Re f(z) \quad \text{for } z \sim (x, y) \in \omega.$$

It follows from Section 3.1 that every conformal map is described by an analytic function in a simply connected domain. Thus  $(x, y) = \Phi(\xi, \eta)$  is described by

$$z = x + iy = g(\zeta) = g(\xi + i\eta) \in \Omega$$

for some analytic function  $g(\zeta)$ . Then the composition  $(f \circ g)(\zeta)$  is an analytic function, and the real part is  $u \circ \Phi(\xi, \eta)$ , which therefore is harmonic in the variables  $(\xi, \eta)$ . Since this holds for every  $(x, y) \in \Omega$ , the theorem is proved.  $\square$

Theorem 3.9.1 can also be phrased in the following way that the *Laplace differential operator* in the variables  $(x, y)$  is transformed by a conformal map into the *Laplace differential operator* in the new variables  $(\xi, \eta)$ . This observation implies that we can reduce the task of solving the Dirichlet problem to transforming the open and simply connected domain  $\Omega$  applying a conformal map into some standard domain, like either the unit disc or the open upper half plane. The boundary conditions are then carried over by the given conformal map to this standard domain.

The existence of such a conformal map follows from *Riemann's mapping theorem*. We have furthermore in the previous sections developed explicit formulæ for special domains for this process.

The solution formula for the *Dirichlet problem* on the *unit disc* for given boundary condition  $h(x, y)$ , i.e.

$$\begin{cases} \Delta u = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, & x^2 + y^2 < 1, \\ u(x, y) = h(x, y), & x^2 + y^2 = 1, \end{cases}$$

has already been given in *Ventus, Complex Functions Theory a-2* by *Poisson's integral formula*. This solution formula is for obvious reasons formulated in polar coordinates, so we use in (117) below that  $z = x + iy = r e^{i\Theta}$ ,

$$(117) \quad u(z) = u(r, \Theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{(1 - r^2) h(e^{i\varphi})}{1 + r^2 - 2r \cos(\Theta - \varphi)} d\varphi, \quad |z| = r < 1.$$

It is possible to prove, using some results from *Functional Analysis*, that if  $h(e^{i\varphi})$  is bounded and piecewise continuous, then (117) is the only *bounded* solution of (117).

**Remark 3.9.1** We mention for completeness that if  $h(e^{i\varphi})$  is not continuous, then there in general exist *unbounded* solutions, but these will not be of any relevance in the applied sciences. However, if  $h(e^{i\varphi})$  is continuous everywhere on the unit circle  $|z| = 1$ , then (117) is the only solution.  $\diamond$

Due to the term  $\cos(\Theta - \varphi)$  in the denominator of the integrand, (117) is in general difficult to apply. Instead we notice that we can use Section 3.3.5 to map the unit disc conformally and bijectively by e.g.  $w = -i \frac{z+1}{z-1}$  onto the upper half plane. Then it can be proved (the straightforward and tedious details are left to the interested reader) that we also have the following solution formula.

**Theorem 3.9.2** (The Dirichlet problem for the upper half plane). *Let  $h(x)$  be a real and bounded and piecewise continuous function on  $\mathbb{R}$ . Then the Dirichlet problem*

$$(118) \quad \begin{cases} \Delta u(x, y) = 0 & \text{for } x \in \mathbb{R} \text{ and } y \in \mathbb{R}_+, \\ u(x, 0) = h(x) & \text{for } x \in \mathbb{R} \text{ a point of continuity of } h(x), \end{cases}$$

has the following unique bounded solution in the upper half plane,

$$(119) \quad \boxed{u(x, y) = \frac{y}{\pi} \int_{-\infty}^{+\infty} \frac{h(t)}{(x-t)^2 + y^2} dt, \quad \text{for } x \in \mathbb{R} \text{ and } y > 0.}$$

**Example 3.9.1** If  $h$  is continuous and bounded and  $h(-\infty) = h(+\infty)$ , then we transform (119) back to the case of the unit disc (117) to conclude that (119) is the only bounded solution. However, if we choose the harmonic function  $u(x, y) = 2xy$ , then  $u(x, y) = \Im \{z^2\}$  is harmonic in the upper half plane, and  $u(x, 0) = 0$ , so both  $u_1(x, y) = 2xy$  and  $u_2(x, y) = 0$  are solutions of the Dirichlet problem

in the upper half plane with boundary condition 0 on the real axis. Clearly,  $u_1(x, y)$  is unbounded in the upper half plane, while  $u_2(x, y)$  is bounded. This example shows that we cannot exclude the assumption that the harmonic function should also be *bounded* in the upper half plane, if we want it to be unique.  $\diamond$

**Remark 3.9.2** The solution formula (119) has the structure of a so-called *convolution*. Given two real functions  $f$  and  $g$ , their convolution is defined by

$$(f \star g)(x) := \int_{-\infty}^{+\infty} f(t)g(x-t) dt = \int_{-\infty}^{+\infty} f(x-t)g(t) dt,$$

provided that the integrals above make sense.

In the present case we define

$$\varphi_y(x) := \frac{1}{\pi} \frac{y}{x^2 + y^2} = -\Im \left\{ \frac{1}{\pi z} \right\} \quad \text{for } y > 0 \text{ fixed.}$$

By assumption,  $h(x)$  is real, so we get from (119) that

$$(120) \quad u(x, y) = (h \star \varphi_y)(x) = -\Im \left\{ h \star \frac{1}{\pi z} \right\}(x, y), \quad y > 0,$$

where  $h \star \frac{1}{\pi z}$  denotes the convolution with respect to the real variable  $x$ .

We see by comparison that (120) is closely connected with the *Hilbert transform*  $\mathcal{H}$ , which is defined by

$$(121) \quad \mathcal{H}f(x) := \frac{1}{\pi} \text{vp} \int_{-\infty}^{+\infty} \frac{f(t)}{x-t} dt = \frac{1}{\pi} \text{vp} \left\{ f \star \frac{1}{t} \right\}(x),$$

where “vp” (“*valeur principale*”) denotes *Cauchy’s principal value*, which is here defined by

$$(122) \quad \frac{1}{\pi} \text{vp} \int_{-\infty}^{+\infty} \frac{f(t)}{x-t} dt := \frac{1}{\pi} \lim_{\varepsilon \rightarrow 0^+} \int_{-\infty}^{x-\varepsilon} + \int_{x+\varepsilon}^{+\infty} \frac{f(t)}{x-t} dt,$$

i.e. we remove a symmetric interval around the singular point  $t = x$  of the integrand, integrate over the remainder, and finally let the length of this symmetric interval tend to zero.

A direct computation of (122) is difficult and may even be impossible, but using the correspondence between (121) and (119) and then the interpretation of (118) we may have a different approach of (121) which may be successful.  $\diamond$

### 3.10 Some special Dirichlet problems for the upper half plane

Given a Dirichlet problem of one of the simply connected domains considered in the Sections 3.3–3.6, we now know how to transform them into a Dirichlet problem in the upper half plane. The solution integral (119) may still be difficult to compute explicitly, but one could try to approximate the given boundary value function  $h(x)$ ,  $x \in \mathbb{R}$ , by a piecewise constant function and then hope for a simple solution formula, which can easily be used in practice. That this indeed is the case, follows from the next theorem.

**Theorem 3.10.1** *Given the Dirichlet problem*

$$(123) \quad \begin{cases} \Delta u(x, y) = 0 & \text{for } y > 0, \\ h(x, 0) = h(x) & \text{for } x \in \mathbb{R}, \end{cases}$$

*in the upper half plane, where the given boundary function  $h(x)$  is piecewise constant,*

$$(124) \quad h(x) = \begin{cases} A_0, & \text{for } x < x_0, \\ A_j, & \text{for } x_{j-1} < x < x_j, \quad j = 1, \dots, n, \\ A_{n+1}, & \text{for } x_n < x. \end{cases}$$

*Then its unique bounded solution is given by*

$$(125) \quad u(x, y) = \frac{1}{\pi} \left\{ A_0 \cdot \text{Arg}(z - z_0) + \sum_{j=1}^n A_j \cdot \text{Arg}\left(\frac{z - x_j}{z - x_{j-1}}\right) + A_{n+1} \cdot \{\pi - \text{Arg}(z - x_n)\} \right\}.$$

PROOF. It was previously mentioned that there is only one bounded solution of (123), when  $h(x)$  given by (124) is bounded. It is therefore possible to solve the problem by only using inspection, where we use that the solution is linear in the boundary function.

First consider the simple boundary problem, where

$$h_{n+1}(x) = \begin{cases} A_{n+1} & \text{for } x > x_n, \\ 0 & \text{for } x < x_n. \end{cases}$$

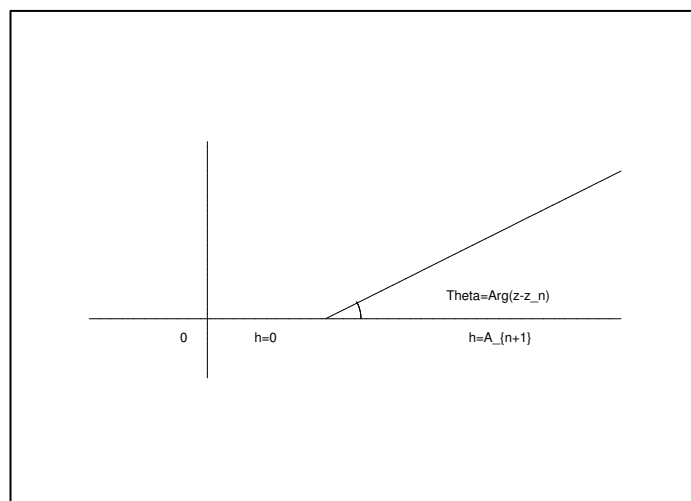


Figure 40: Geometric analysis of the boundary value problem, where  $h_{n+1}(x) = 0$  for  $x < x_n$ , and  $h_{n+1}(x) = A_{n+1}$  for  $x > x_n$ .

It follows by inspection, cf. Figure 40, that the function

$$(126) \quad u_{n+1}(x, y) := A_{n+1} \cdot \left\{ 1 - \frac{1}{\pi} \operatorname{Arg}(z - x_n) \right\}, \quad \Im z > 0,$$

is bounded, and it has the right boundary values.

Finally,  $\operatorname{Arg}(z - x_n) = \Im \operatorname{Log}(z - x_n)$  is harmonic in the open half plane, because it is the imaginary part of an analytic function, and so is  $u(x, y)$ , hence (126) is the unique bounded solution of the simplified Dirichlet problem.

The same method can be used for another simplified Dirichlet problem, where now

$$h_j(x) = \begin{cases} A_j & x_{j-1} < x < x_j, \\ 0 & \text{otherwise.} \end{cases}$$

In this case,

$$0 < \Theta_{j-1} = \operatorname{Arg}(z - x_{j-1}) < \Theta_j = \operatorname{Arg}(z - x_j) < \pi,$$

because  $x_{j-1} < x_j$ , so

$$(127) \quad \operatorname{Arg}(z - x_j) - \operatorname{Arg}(z - x_{j-1}) = \operatorname{Arg}\left(\frac{z - x_j}{z - x_{j-1}}\right),$$

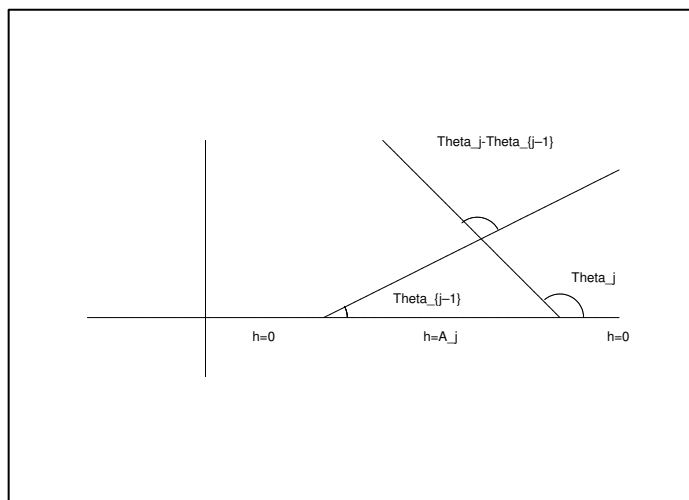


Figure 41: Geometric analysis of the boundary value problem, where  $h_j(x) = A_j$  for  $x_{j-1} < x < x_j$ , and  $h_j(x) = 0$  otherwise.

because the left hand side also lies in  $]0, \pi[$ , cf. Figure 41.

The boundary value of the left hand side of (127) is  $\pi$  for  $z = x \in ]x_{j-1}, x_j[$ , and 0 otherwise. It follows that this boundary value problem has the unique bounded solution

$$(128) \quad u_j(x, y) = \frac{A_j}{\pi} \operatorname{Arg} \left( \frac{z - x_j}{z - x_{j-1}} \right).$$

Finally, the same method proves that if

$$h_0(x) = \begin{cases} A_0 & \text{for } x < x_0, \\ 0 & \text{for } x > x_0, \end{cases}$$

then the unique bounded solution is given by

$$(129) \quad u(x, y) = \frac{A_0}{\pi} \operatorname{Arg}(z - x_0).$$

When we add (129), (128) and (126), we get the boundary function (124) as well as the solution (125), and the theorem is proved.  $\square$

**Example 3.10.1** We give here an example of a combination of the theories above. We shall find the idealized stationary distribution of temperature in an half infinite strip with fixed temperature  $T = 0$  on the two infinite boundary lines and  $T = 2$  on the bounded boundary line segment.

We choose the domain  $\Omega = \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[ \times \mathbb{R}_+$ , cf. Figure 42, which easily can be transformed into any other such half infinite strip by a rotation, a stretching and a translation. It therefore suffices to consider  $\Omega$  given above.

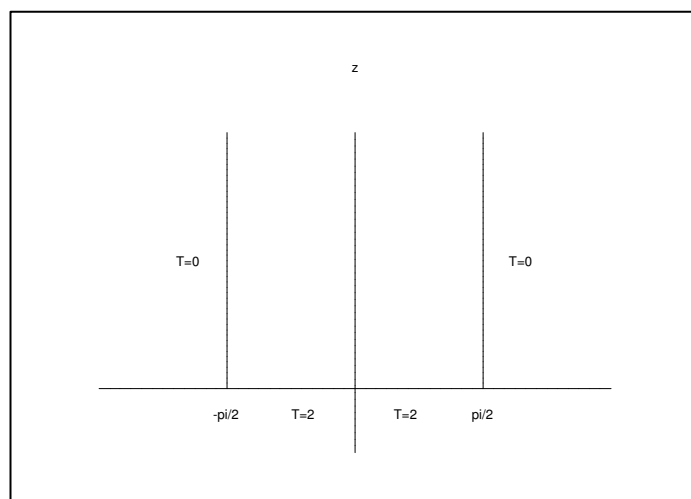


Figure 42: The chosen domain in Example 3.10.1 with the boundary values of the Dirichlet problem.

Using the sine transform of Section 3.6 the map  $z_1 = \sin z$  maps  $\Omega$  above conformally and bijectively onto the upper half  $z_1$ -plane with the boundary values

$$u_1(x_1, 0) = h_1(x_1) = \begin{cases} 0 & \text{for } x_1 < -1, \\ 2 & \text{for } -1 < x_1 < 1, \\ 0 & \text{for } x_1 > 1, \end{cases}$$

cf. Figure 43.

Then we apply (128) to get the *bounded* solution

$$u(x, y) = u_1(x_1, y_1) = \frac{1}{\pi} \arg\left(\frac{\sin z - 1}{\sin z + 1}\right).$$

We shall in the following prove that this function can be written

$$(130) \quad u(x, y) = \frac{2}{\pi} \operatorname{Arccot}\left(\frac{\sinh^2 y - \cos^2 x}{2 \sinh y \cdot \cos x}\right).$$

In fact, we use that  $\operatorname{Arg}(\lambda z_1) = \operatorname{Arg} z_1$  for  $\lambda = |\sin z + 1|^{-2} > 0$  to get

$$\begin{aligned} u(x, y) &= \frac{2}{\pi} \operatorname{Arg}\left(\frac{\sin z - 1}{\sin z + 1}\right) = \frac{2}{\pi} \operatorname{Arg}\left(\frac{(\sin z - 1)(\overline{\sin z + 1})}{|\sin z + 1|^2}\right) \\ &= \frac{2}{\pi} \operatorname{Arg}((\sin z - 1)(\overline{\sin z + 1})) = \frac{2}{\pi} \operatorname{Arg}(|\sin z|^2 + \sin z - \overline{\sin z} - 1) \\ &= \frac{2}{\pi} \operatorname{Arg}(|\sin z|^2 - 1 + 2i \Im\{\sin z\}) = \frac{2}{\pi} \operatorname{Arg}(\sin^2 x + \sinh^2 y - 1 + 2i \cdot \cos x \cdot \sinh y) \\ &= \frac{2}{\pi} \operatorname{Arg}(\sinh^2 y - \cos^2 x + i \cdot 2 \cos x \cdot \sinh y). \end{aligned}$$

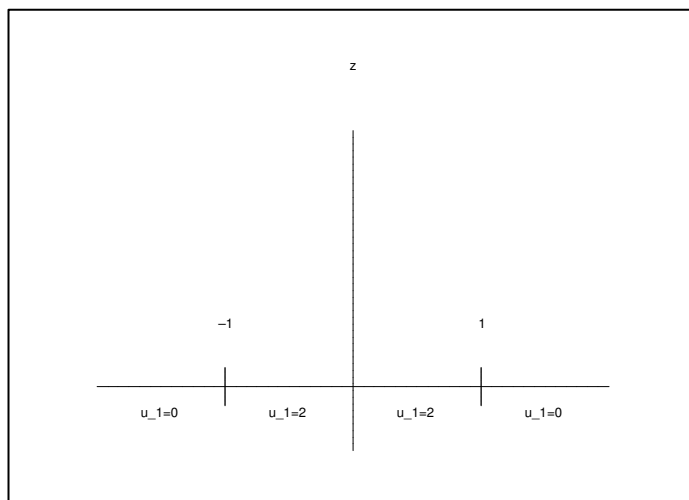


Figure 43: The Dirichlet problem of Figure 42 is by the map  $z_1 = \sin z$  mapped bijectively and conformally onto a simpler Dirichlet problem in the upper half  $z_1$ -plane.

Since  $2 \cdot \cos x \cdot \sinh y > 0$  for all  $(x, y) \in \Omega$ , the principal argument  $\text{Arg}$  is described by the usual real function  $\text{Arccot}$ , thus

$$u(x, y) = \frac{2}{\pi} \text{Arg}(\sinh^2 y - \cos^2 x + i \cdot 2 \cos x \cdot \sinh y) = \frac{2}{\pi} \text{Arccot}\left(\frac{\sinh^2 y - \cos^2 x}{2 \sinh y \cdot \cos x}\right),$$

and (130) is proved.

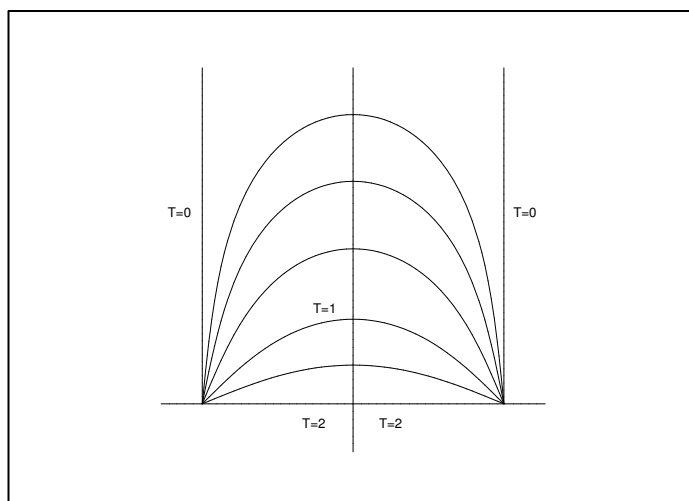


Figure 44: The level curves of the stationary distribution of the temperature field of Example 3.10.1.



Using (130) we see that the level curves  $u(x, y) = c \in ]0, 2[$  (curves of equal temperature) are given by

$$\frac{\sinh^2 y - \cos^2 x}{2 \cdot \sinh y \cdot \cos x} = \cot\left(\frac{c\pi}{2}\right), \quad 0 < c < 2 \text{ arbitrary,}$$

whence

$$\sinh^2 y - 2 \cdot \cos x \cdot \coth\left(\frac{c\pi}{2}\right) \cdot \sinh y - \cos^2 x = 0.$$

Using that  $y > 0$ , hence also  $\sinh y > 0$ , we get the following single valued solution of this equation,

$$\begin{aligned} \sinh y &= \cos x \cdot \cot\left(\frac{c\pi}{2}\right) + \sqrt{\left\{1 + \cot^2\left(\frac{c\pi}{2}\right)\right\} \cos^2 x} \\ &= \cot\left(\frac{c\pi}{2}\right) \cos x + \frac{\cos x}{\sin\left(\frac{c\pi}{2}\right)} = \frac{1 + \cos\left(\frac{c\pi}{2}\right)}{\sin\left(\frac{c\pi}{2}\right)} = \cot\left(\frac{c\pi}{4}\right) \cos x, \end{aligned}$$

where  $\sinh y > 0$  requires  $+$  in front of the square root, and where we have used that  $\cos x > 0$  for  $x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ . The level curves are therefore given by

$$y = \operatorname{Arsinh}\left(\cot\left(\frac{c\pi}{4}\right) \cos x\right) = \ln\left(\cot\left(\frac{c\pi}{4}\right) \cos x + \sqrt{1 + \cot^2\left(\frac{c\pi}{4}\right) \cos^2 x}\right) \quad \text{for } x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[ ,$$

where  $c \in ]0, 2[$  is an arbitrary constant.  $\diamond$

## Index

- amplification factor, 37  
amplitude, 25  
angle of rotation, 72  
argument function, 9  
argument principle, 7, 9, 12, 14, 18, 36  
argument variation, 11
- bilinear transformation, 74  
binomial equation, 18  
branch, 41  
branch cut, 41-43, 45, 46, 49, 53, 54, 56, 59, 65, 66  
branch point, 44-45, 52, 56, 58-60, 63
- Cauchy Integral formula, 7  
Cauchy Integral Theorem, 7  
Cauchy's principal value, 107  
Cauchy-Riemann equations, 7  
characteristic polynomial, 25, 26  
complex logarithm, 46, 47  
conformal map, 7, 8, 72  
convolution, 107  
convolution equation, 25  
criteria of stability, 7, 25  
cross-ratio, 78  
Cybernetics, 7, 26
- dipole, 67, 70  
Dirichlet problem, 7, 8, 72, 73, 84, 104, 106
- equipotential curve, 67, 69-71, 93, 103  
exponential transformation, 87
- feedback, 35  
flow function, 69  
fractional function, 26  
fundamental domain, 41, 43, 45, 46, 49, 53, 55, 59, 61, 65, 66  
Fundamental Theorem of Algebra, 21, 27, 63
- general exponential, 49  
general power function, 49, 51  
general similarity, 76, 79  
Geometry, 7
- harmonic functions, 7
- Hilbert transform, 107  
homographic, 74  
Hurwitz polynomial, 27, 30, 32, 33
- improper integrals, 7  
intensity of vortex, 69  
inverse Laplace transform, 26  
inverse mapping, 38
- Joukowski's function, 91
- Lagrange, 7  
Laplace differential equation, 103  
Laplace differential operator, 73, 105  
Laplace transform, 25  
Laurent series, 7, 63  
Laurent series expansion, 23  
level curve, 84, 89, 90  
linear differential equation, 25  
linear fractional transformation, 74  
local inverse map, 38  
logarithmic branching point, 40, 41, 46, 47, 52, 59, 60  
logarithmic derivative, 14  
logarithmic spiral, 70
- Möbius transformation, 74  
main theorem, 10  
many-valued function, 7, 8, 38, 39, 41
- n-valued map, 45  
Nyquist's criterion, 35
- Poisson's integral formula, 73, 106  
polar coordinates, 42, 84, 87  
pole, 14, 26  
potential, 67  
potential flow, 67  
principal argument, 13, 46  
principal logarithm, 46, 47, 50, 57, 61, 87  
principal value, 50, 51, 57, 58, 60, 62
- real logarithm, 47  
real potential, 69  
reciprocal transformation, 77, 79

- 
- rectangular coordinates, 87  
removable singularity, 14  
residue calculus, 7  
residuum, 24  
residuum formula, 23  
residuum theorem, 16  
Reynold's number, 94  
Riccati differential equation, 81  
Riemann mapping theorem, 95  
Riemann surface, 7, 38, 39, 41, 43, 63  
Riemann's mapping theorem, 95, 106  
rotation, 76  
Rouché's theorem, 17-21  
Routh-Hurwitz stability criterion, 33
- Schur's criterion, 32  
Schwarz-Christoffel formula, 97  
Schwarz-Christoffel transformation, 95  
series expansion, 63  
similarity, 72, 76  
sine transformation, 89  
singular point, 40-46, 63  
sink, 67  
source, 67  
square root, 42  
stability, 7, 25, 27, 41  
stability criterion, 41  
stationary potential flow, 94  
stream function, 103  
streamline, 68, 69, 94, 99, 103  
strength of source, 67  
stretching factor, 72, 76
- Taylor series, 7  
transformation by a power function, 84  
transition function, 25-27, 35  
translation, 75  
trigonometric integrals, 7  
two-valued function, 44
- uniform continuity, 10
- velocity field, 94  
vortex, 67, 69  
vortex source/sink, 67, 69
- winding number, 9, 11, 12, 14, 15, 37
-