Methods of Theoretical Physics 614

ABSTRACT

Second part of the Fall 2009 course 614, on Mathematical Methods in Theoretical Physics. This partial set of notes begins with complex analysis.

Contents

1 Functions of a Complex Variable

1.1 Complex Numbers, Quaternions and Octonions

The extension from the real number system to complex numbers is an important one both within mathematics itself, and also in physics. The most obvious area of physics where they are indispensable is quantum mechanics, where the wave function is an intrinsically complex object. In mathematics their use is very widespread. One very important point is that by generalising from the real to the complex numbers, it becomes possible to treat the solution of polynomial equations in a uniform manner, since now not only equations like $x^2 - 1 = 0$ but also $x^2 + 1 = 0$ can be solved.

The complex numbers can be defined in terms of ordered pairs of real numbers. Thus we may *define* the complex number z to be the ordered pair $z = (x, y)$, where x and y are real. Of course this doesn't tell us much until we give some rules for how these quantities behave. If $z = (x, y)$, and $z' = (x', y')$ are two complex numbers, and a is any real number, then the rules can be stated as

$$z + z' = (x + x', y + y')\,, \tag{1.1}$$

$$a\,z = (a\,x, a\,y)\,, \tag{1.2}$$

$$z\,z' = (x\,x' - y\,y', x\,y' + x'\,y)\,. \tag{1.3}$$

We also define the *complex conjugate* of $z = (x, y)$, denoted by $\bar{z}$, as

$$\bar{z} = (x, -y)\,. \tag{1.4}$$

Note that a complex number of the form $z = (x, 0)$ is therefore *real*, according to the rule (1.4) for complex conjugation; $\bar{z} = z$. We can write such a real number in the time-honoured way, simply as x. Thus we may define

$$(x, 0) = x\,. \tag{1.5}$$

The *modulus* of z, denoted by $|z|$, is defined as the positive square root of $|z|^2 \equiv \bar{z}\,z$, which, from (1.3), (1.4) and (1.5), is given by

$$|z|^2 = \bar{z}\,z = (x^2 + y^2, 0) = x^2 + y^2\,, \tag{1.6}$$

It is manifest that $|z| \geq 0$, with $|z| = 0$ if and only if $z = 0$.

It is now straightforward to verify that the following fundamental laws of algebra are satisfied:

1. *Commutative and Associative Laws of Addition*:

$$
\begin{aligned}
z_1 + z_2 &= z_2 + z_1 \,, \\
z_1 + (z_2 + z_3) &= (z_1 + z_2) + z_3 = z_1 + z_2 + z_3 \,,
\end{aligned}
\tag{1.7}
$$

2. *Commutative and Associative Laws of Multiplication*:

$$
\begin{aligned}
z_1\, z_2 &= z_2\, z_1 \,, \\
z_1\,(z_2\, z_3) &= (z_1\, z_2)\, z_3 = z_1\, z_2\, z_3 \,,
\end{aligned}
\tag{1.8}
$$

3. *Distributive Law*:

$$
(z_1 + z_2)\, z_3 = z_1\, z_3 + z_2\, z_3 \,. \tag{1.9}
$$

We can also define the operation of division. If $z_1\, z_2 = z_3$, then we see from the previous rules that, multiplying by $\bar z_1$, we have

$$
\bar z_1\,(z_1\, z_2) = (\bar z_1\, z_1)\, z_2 = |z_1|^2\, z_2 = \bar z_1\, z_3 \,, \tag{1.10}
$$

and so, provided that $|z_1| \ne 0$, we can write the *quotient*

$$
z_2 = \frac{z_3\, \bar z_1}{|z_1|^2} = \frac{z_3}{z_1} \,. \tag{1.11}
$$

(The expression $z_3\, \bar z_1/|z_1|^2$ here defines what we mean by z_3/z_1.) The fact that we can solve $z_1\, z_2 = z_3$ for z_2 when z_1 is non-zero, effectively by dividing out by z_1, is a slightly non-trivial property of the complex numbers, which we needed to check before we could assume it to be true. The analogous feature for real numbers is, of course, totally familiar, and we use it every day without a moment's thought. This property of the real number system and the complex number system is described by the statement that the real numbers and the complex numbers both form *Division Algebras*.

We can, of course, recognise that from the previous rules that the square of the complex number $(0,1)$ is $(-1,0)$, which from (1.5) is simply -1. Thus we can view $(0,1)$ as being the square root of -1:

$$
(0,1) = \mathrm{i} = \sqrt{-1} \,. \tag{1.12}
$$

As can be seen from (1.4), it has the following property under complex conjugation:

$$
\overline{(0,1)} = (0,-1) = -(0,1) \,. \tag{1.13}
$$

In other words, $\bar{\mathrm{i}} = -\mathrm{i}$.

Note that in our description of the complex numbers as ordered pairs of real numbers, this is the first time that the symbol i, meaning the square root of minus 1, has appeared. We can, of course, re-express everything we did so far in the more familiar notation, in which we write the complex number $z = (x, y)$ as

$$z = x + \mathrm{i}\, y\,. \tag{1.14}$$

The symbol i is called the *imaginary unit*.

One might be wondering at this stage what all the fuss is about; we appear to be making rather a meal out of saying some things that are pretty obvious. Well, one reason for this is that one can also go on to consider more general types of "number fields," in which some of the previous properties cease to hold. It then becomes very important to formalise things properly, so that there is a clear set of statements of what is true and what is not. Below, we will give a more extended discussion of more general fields of numbers, but it is useful first to give a summary. A reader who is not interested in this digression into more abstruse number systems can skip the detailed discussion, and pass on directly to section 1.1.2.

The property of being a division algebra, i.e. that the equation $AB = C$ can be solved for B, whenever A is non-zero, is a very important one, and it is a property one does not lightly give up. Thus the dumber systems of principal interest are those that are division algebras. There are two such number systems in addition to the real and the complex numbers, namely the quaternions and the octonions. The defining properties of these number systems can be stated very easily, in an iterative fashion in which the one builds up from reals to complex numbers to quaternions to octonions. Just as a complex number can be defined as an ordered pair of real numbers, we can define a quaternion as an ordered pair of complex numbers, and an octonion as an ordered pair of quaternions. Thus we may universally write

$$A = (a, b)\,, \tag{1.15}$$

where if a and b are real, then A is complex; if a and b are complex, then A is a quaternion; and if a andf b are quaternions, then A is an octonion. We need to specify two rules, namely the rule for multiplication, and the rule for conjugation. These rules can be stated completely generally, applying to all four of the division algebras: If $A = (a, b)$ and $B = (c, d)$, then

$$AB = (a\,c - \bar{d}\,b, d\,a + b\,\bar{c})\,, \tag{1.16}$$

$$\bar{A} = (\bar{a}, -b)\,. \tag{1.17}$$

We also have the rules for multiplication by a real number λ, and for addition:

$$\lambda\, A = (\lambda\, a, \lambda\, b)\,, \qquad A + B = (a + c, b + d)\,. \tag{1.18}$$

The rules (1.16) and (1.17), together with (1.18), are all that one needs in order to define the four divsion algebras. Note that they reduce to the rules (1.3) and (1.4), together with (1.1) and (1.2), in the case of the complex number system, for which a, b, c and d are just real numbers. Note in particular that in this case, the conjugations (the bars) on the quantities appearing on the right-hand sides of (1.16) and (1.17) serve no purpose, since a, b, c and d are real. But they do no harm, either, and the nice thing is that (1.16) and (1.17) are completely *universal* rules, applying to all the division algebras.

It is straightforward to check from (1.16) that the multiplication of quaternions is non-commutative; i.e. in general

$$AB \ne BA\,. \tag{1.19}$$

This happens because of the conjugations on the right-hand side of (1.16). It can also be checked that the multiplication rule for quaternions is still associative:

$$A(BC) = (AB)C\,. \tag{1.20}$$

Note, however, that one cannot just *assume* this; rather, one applies the multiplication rule (1.16) and checks it. A further property of the quaternions, again derivable from the rules given above, is that the conjugation of the product AB gives

$$\overline{AB} = \bar B \bar A\,. \tag{1.21}$$

Notice that all the features of the quaternions, including the non-commutativity of multiplication, and the conjugation property (1.21), are familiar from the way in which matrices work (with conjugation now understood to be Hermitean conjugation). In fact, as we shall discuss in section 1.1.1 below, one can represent quaternions by 2×2 matrices.

If we now go to octonions, then it is straightforward to check from the universal rules given above in (1.16) and (1.17), together with (1.1) and (1.2), that not only is their multiplication non-commutative, but it is also non-associative; i.e. in general

$$A(BC) \ne (AB)C \tag{1.22}$$

This non-associativity is a direct consequence of the non-commutativity of the multiplication rule for the quaternions that now appear on the right-hand side of (1.16) and (1.17). Notice that the *order* in which the symbols appear in the expressions on the right-hand side of (1.16) is absolutely crucial now. Of course it did not matter when we were defining quaternions, because the multiplication rule for the complex numbers is commutative. But it does matter now when we multiply octonions, because the constituent quaternions do not multiply commutatively.

Notice that from the universal rules given above we can also derive that the conjugation of a product of two octonions satisfies the same property (1.21) that we saw above for the quaternions. However, it should be emphasised that unlike the quaternions, which can be represented by 2×2 matrices, the octonions can *not* be represented by matrices. To see this, we need look no further than the multiplicative non-associativity displayed in equation (1.22): Matrix multiplication is associative, and therefore octonions cannot be represented by matrices.

The crucial feature that all four of these number systems have is that they are *Division Algebras*. First, notice that in all cases we have

$$|A|^2 \equiv A\bar{A} = \bar{A}A = (a\bar{a} + b\bar{b}, 0)\,, \tag{1.23}$$

Furthermore, these quantity is real, and so we can just write it as

$$A\bar{A} = \bar{A}A = a\bar{a} + b\bar{b}\,. \tag{1.24}$$

We have the property that $|A|^2 \ge 0$, with equality if and only if $A = 0$. At the risk of sounding like the proverbial broken gramaphone record (for those old enough to remember what they are), we can emphasise again here that *all* the statements being made here can be verified directly using just (1.16) and (1.17), together with (1.1) and (1.2).

For the quaternions, it is then obvious that they form a division algebra; we just multiply $AB = C$ on the left by $\bar{A}$, and use

$$\bar{A}(AB) = (\bar{A}A)B = |A|^2\, B = \bar{A}C\,, \tag{1.25}$$

and hence, dividing out by the non-zero real number $|A|^2$, we get

$$B = \frac{\bar{A}C}{|A|^2}\,. \tag{1.26}$$

For the octonions, we can actually perform the identical calculation, but in order to do this we must first check a slightly non-obvious fact. We noted that in general, octonionic multiplication is non-associative. Looking at (1.25) we see that we assumed that $\bar{A}(AB) = (\bar{A}A)B$. This is no problem for quaternions, since their multiplication rule *is* associative, but it looks dangerous for octonions. However, one can check (from, as always, (1.16) and

(1.17)), that in the special case of multiplying the three octonions $\bar{A}AB$ (in other words, the special case when two of the adjacent octonions are conjugates of each other), the multiplication *is* associative, and so the steps in (1.25) are still valid. This then proves that the octonions do indeed form a division algebra.

It is worth noting, however, that we only just "got away with it" for the octonions. If we try going to the next natural generalisation of the octonions, which we might call the "hexadecions," defined as ordered pairs of octonions obeying the rules (1.16) and (1.17), then we find that our luck has run out; we do not get a division algebra.

1.1.1 Further details on quaternions and octonions

As we have already emphasised, the rules described above specify the four division algebras (i.e. reals, complex, quaternions and octonions) completely, and one does not need any other knowledge or definitions in order to manipulate them. However, it is sometimes useful to have other ways to think about them, analogous to the way in which we introduce the symbol i, as the square root of minus 1, for the complex numbers.

We give some more details about these other ways of thinking of quaternions and octonions in this section. Some of the discussion is a little repetitive, since it is drawn from an earlier version of these lecture notes. This entire section can be ignored, if desired.

The "next" extension beyond the complex numbers is, as has been said above, to the *quaternions*. Another way of thinking about them is to say that one now has three independent imaginary units, usually denoted by i, j and k, subject to the rules

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1\,, \qquad \mathrm{i}\,\mathrm{j} = -\mathrm{j}\,\mathrm{i} = \mathrm{k}\,, \qquad \mathrm{j}\,\mathrm{k} = -\mathrm{k}\,\mathrm{j} = \mathrm{i}\,, \qquad \mathrm{k}\,\mathrm{i} = -\mathrm{i}\,\mathrm{k} = \mathrm{j}\,. \tag{1.27}$$

A quaternion q is then a quantity of the form

$$q = q_0 + q_1\,\mathrm{i} + q_2\,\mathrm{j} + q_3\,\mathrm{k}\,, \tag{1.28}$$

where q_0, q_1, q_2 and q_3 are all real numbers. There is again an operation of complex conjugation, $\bar{q}$, in which the signs of all three of i, j and k are reversed

$$\bar{q} = q_0 - q_1\,\mathrm{i} - q_2\,\mathrm{j} - q_3\,\mathrm{k}\,, \tag{1.29}$$

The modulus $|q|$ of a quaternion q is a real number, defined to be the positive square root of

$$|q|^2 \equiv \bar{q}\,q = q\,\bar{q} = q_0^2 + q_1^2 + q_2^2 + q_3^2\,. \tag{1.30}$$

Clearly $|q| \ge 0$, with equality if and only if $q = 0$.

Which of the previously-stated properties of complex numbers still hold for the quaternions? It is not so obvious, until one goes through and checks. It is perfectly easy to do this, of course; the point is, though, that it does now need a bit of careful checking, and the value of setting up a formalised structure that defines the rules becomes apparent. The answer is that for the quaternions, one has now lost multiplicative commutativity, so $q\,q' \neq q'\,q$ in general. A consequence of this is that there is no longer a unique definition of the quotient of quaternions. However, a very important point is that we *do* keep the following property. If q and q' are two quaternions, then we have

$$|q\,q'| = |q|\,|q'|\,, \tag{1.31}$$

as one can easily verify from the previous definitions.

Let us note that for the quaternions, if we introduce the notation γ_a for $a = 1, 2, 3$ by

$$\gamma_1 = \mathrm{i}\,, \qquad \gamma_2 = \mathrm{j}\,, \qquad \gamma_3 = \mathrm{k}\,, \tag{1.32}$$

then the algebra of the quaternions, given in (1.27), can be written as

$$\gamma_a\,\gamma_b = -\delta_{ab} + \epsilon_{abc}\,\gamma_c\,, \tag{1.33}$$

where ϵ_{abc} is the totally antisymmetric tensor with

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1\,, \quad \epsilon_{132} = \epsilon_{321} = \epsilon 213 = -1\,. \tag{1.34}$$

Note that the Einstein summation convention for the repeated c index is understood, so (1.33) really means

$$\gamma_a\,\gamma_b = -\delta_{ab} + \sum_{c=1}^{3} \epsilon_{abc}\,\gamma_c\,. \tag{1.35}$$

In fact, one can recognise this as the multiplication algebra of $-\sqrt{-1}$ times the Pauli matrices σ_a of quantum mechanics, $\gamma_a = -\sqrt{-1}\,\sigma_a$, which can be represented as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.36}$$

(We use the rather clumsy notation $\sqrt{-1}$ here to distinguish this "ordinary" square root of minus one from the i quaternion.) In this representation, the quaternion defined in (1.28) is therefore written as

$$q = \begin{pmatrix} q_0 - \sqrt{-1}\,q_3 & -\sqrt{-1}\,q_1 - q_2 \\ -\sqrt{-1}\,q_1 + q_2 & q_0 + \sqrt{-1}\,q_3 \end{pmatrix}. \tag{1.37}$$

Since the quaternions are now represented by matrices, it is immediately clear that we shall have associativity, $A(BC) = (AB)C$, but not commutativity, under multiplication.

As a final remark about the quaternions, note that we can equally well view them as we did previously, as an ordered pair of complex numbers. Thus we may define

$$q = (a, b) = a + b\,\mathrm{j} = a_0 + a_1\,\mathrm{i} + b_0\,\mathrm{j} + b_1\,\mathrm{k}\,, \tag{1.38}$$

where $a = a_0 + a_1\,\mathrm{i}$, $b = b_0 + b_1\,\mathrm{i}$. Here, we assign to i and j the multiplication rules given in (1.27), and k is, by definition, nothing but $\mathrm{i}\,\mathrm{j}$. Quaternionic conjugation is given by $\bar{q} = (\bar{a}, -b)$. The multiplication rule for the quaternions $q = (a, b)$ and $q' = (c, d)$ can then easily be seen to be

$$q\,q' = (a\,c - b\,\bar{d}, a\,d + b\,\bar{c})\,. \tag{1.39}$$

To see this, we just expand out $(a + b\,\mathrm{j})(c + d\,\mathrm{j})$:

$$\begin{aligned}(a + b\,\mathrm{j})(c + d\,\mathrm{j}) &= a\,c + b\,\mathrm{j}\,d\,\mathrm{j} + a\,d\,\mathrm{j} + b\,\mathrm{j}\,c \\ &= a\,c + b\,\bar{d}\,\mathrm{j}^2 + a\,d\,\mathrm{j} + b\,\bar{c}\,\mathrm{j} \\ &= (a\,c - b\,\bar{d}) + (a\,d + b\,\bar{c})\,\mathrm{j} \\ &= (a\,c - b\,\bar{d}, a\,d + b\,\bar{c})\,. \end{aligned} \tag{1.40}$$

Note that the complex conjugations in this expression arise from taking the quaternion j through the quaternion i, which generates a minus sign, thus

$$\begin{aligned}\mathrm{j}\,c &= \mathrm{j}\,(c_0 + c_1\,\mathrm{i}) = c_0\,\mathrm{j} + c_1\,\mathrm{j}\,\mathrm{i} \\ &= c_0\,\mathrm{j} - c_1\,\mathrm{i}\,\mathrm{j} = (c_0 - c_1\,\mathrm{i})\,\mathrm{j} = \bar{c}\,\mathrm{j}\,. \end{aligned} \tag{1.41}$$

Notice that the way quaternions are defined here as ordered pairs of complex numbers is closely analogous to the definition of the complex numbers themselves as ordered pairs of real numbers. The multiplication rule (1.39) is also very like the multiplication rule in the last line in (1.2) for the complex numbers. Indeed, the only real difference is that for the quaternions, the notion of complex conjugation of the constituent complex numbers arises. It is because of this that commutativity of the quaternions is lost.

The next stage after the quaternions is the *octonions*, where one has 7 independent imaginary units. The rules for how these combine is quite intricate, leading to the rather splendidly-named *Zorn Product* between two octonions. It turns out that for the octonions, not only does one not have multiplicative commutativity, but multiplicative associativity is also lost, meaning that $A\,(B\,C) \ne (A\,B)\,C$ in general.

For the octonions, let us denote the 7 imaginary units by γ_a, where now $1 \le a \le 7$. Their multiplication rule is reminiscent of (1.33), but instead is

$$\gamma_a \, \gamma_b = -\delta_{ab} + c_{abc} \, \gamma_c \,, \tag{1.42}$$

where c_{abc} are a set of *totally-antisymmetric* constant coefficients, and the Einstein summation convention is in operation, meaning that the index c in the last term is understood to be summed over the range 1 to 7. Note that the total antisymmetry of c_{abc} means that the interchange of *any* pair of indices causes a sign change; for example, $c_{abc} = -c_{bac}$. A convenient choice for the c_{abc}, which are known as the *structure constants* of the octonion algebra, is

$$c_{147} = c_{257} = c_{367} = c_{156} = c_{264} = c_{345} = -1 \,, \qquad c_{123} = +1 \,. \tag{1.43}$$

Here, it is to be understood that all components related to these by the antisymmetry of c_{abc} will take the values implied by the antisymmetry, while all other components not yet specified are zero. For example, we have $c_{174} = +1$, $c_{321} = -1$, $c_{137} = 0$.

We may think of an octonion w as an object built from 8 real numbers w_0 and w_a, with

$$w = w_0 + w_a \, \gamma_a \,. \tag{1.44}$$

There is a notion of an *octonionic conjugate*, where the signs of the 7 imaginary units are reversed:

$$\bar w = w_0 - w_a \, \gamma_a \,, \tag{1.45}$$

and there is a modulus $|w|$, which is a real number defined by

$$|w|^2 \equiv \bar w \, w = w_0^2 + \sum_{a=1}^{7} w_a^2 \,. \tag{1.46}$$

Note that $|w| \ge 0$, and $|w|$ vanishes if and only if $w = 0$.

One can verify from (1.43) that

$$c_{abc} \, c_{ade} = \delta_{bd} \, \delta_{ce} - \delta_{be} \, \delta_{cd} - c_{bcde} \,, \tag{1.47}$$

where an absolutely crucial point is that c_{bcde} is also *totally antisymmetric*. In fact,

$$c_{bcde} = \tfrac{1}{6} \epsilon_{bcdefgh} \, c_{fgh} \,, \tag{1.48}$$

where $\epsilon_{bcdefgh}$ is the totally-antisymmetric tensor of 7 dimensions, with $\epsilon_{1234567} = +1$.

It is straightforward to see that the octonions are non-associative. For example, from the rules given above we can see that

$$\gamma_3 \, (\gamma_1 \, \gamma_7) = \gamma_3 \, c_{174} \, \gamma_4 = \gamma_3 \, \gamma_4 = c_{345} \, \gamma_5 = -\gamma_5 \,, \tag{1.49}$$

whilst

$$(\gamma_3\,\gamma_1)\,\gamma_7 = c_{312}\,\gamma_2\,\gamma_7 = \gamma_2\,\gamma_7 = c_{275}\,\gamma_5 = +\gamma_5\,. \tag{1.50}$$

So what *does* survive? An important thing that is still true for the octonions is that any two of them, say w and w', will satisfy

$$|w\,w'| = |w|\,|w'|\,. \tag{1.51}$$

Most importantly, all of the real, complex, quaternionic and octonionic algebras are *division algebras*. This means that the concept of division makes sense, which is perhaps quite surprising in the case of the octonions. Suppose that A, B and C are any three numbers in any one of these four number systems. First note that we have

$$\bar A\,(A\,B) = (\bar A\,A)\,B\,. \tag{1.52}$$

This is obvious from the associativity for the real, complex or quaternionic algebras. It is not obvious for the octonions, since they are not associative (i.e. $A\,(B\,C) \ne (A\,B)\,C$), but a straightforward calculation using the previously-given properties shows that it is true for the special case $\bar A\,(A\,B) = (\bar A\,A)\,B$. Thus we can consider the following manipulation. If $A\,B = C$, then we will have

$$\bar A\,(A\,B) = |A|^2\,B = \bar A\,C\,. \tag{1.53}$$

Hence we have

$$B = \frac{\bar A\,C}{|A|^2}\,, \tag{1.54}$$

where we are allowed to divide by the real number $|A|^2$, provided that it doesn't vanish. Thus as long as $A\ne 0$, we can give meaning to the division of C by A. This shows that all four of the number systems are division algebras.

Finally, note that again we can define the octonions as an ordered pair of the previous objects, i.e. quaternions, in this chain of real, complex, quaternionic and octonionic division algebras. Thus we define the octonion $w = (a,b) = a + b\,\gamma_7$, where $a = a_0 + a_1\,\mathrm{i} + a_2\,\mathrm{j} + a_3\,\mathrm{k}$ and $b = b_0 + b_1\,\mathrm{i} + b_2\,\mathrm{j} + b_3\,\mathrm{k}$ are quaternions, and $\mathrm{i} = \gamma_1$, $\mathrm{j} = \gamma_2$ and $\mathrm{k} = \gamma_3$. The conjugate of w is given by $\bar w = (\bar a, -b)$. It is straightforward to show, from the previously-given multiplication rules for the imaginary octonions, that the rule for multiplying octonions $w = (a,b)$ and $w' = (c,d)$ is

$$w\,w' = (a\,c - \bar d\,b, d\,a + b\,\bar c)\,. \tag{1.55}$$

This is very analogous to the previous multiplication rule (1.39) that we found for the quaternions. Note, however, that the issue of ordering of the constituent quaternions in these octonionic products is now important, and indeed a careful calculation from the multiplication rules shows that the ordering must be as in (1.55). In fact (1.55) is rather general, and encompasses all three of the multiplication rules that we have met. As a rule for the quaternions, the ordering of the complex-number constituents in (1.55) would be unimportant, and as a rule for the complex numbers, not only the ordering but also the complex conjugation of the real-number constituents would be unimportant.

1.1.2 The Complex Plane

After discussing the generalities of division algebras, let us return now to the complex numbers, which is the subject we wish to develop further here. Since a complex number z is an ordered pair of real numbers, $z = (x, y)$, it is natural to represent it as a point in the two-dimensional plane, whose Cartesian axes are simply x and y. This is known as the *Complex Plane*, or sometimes the *Argand Diagram.* Of course it is also often convenient to employ polar coordinates r and θ in the plane, related to the Cartesian coordinates by

$$x = r\,\cos\theta\,, \qquad y = r\,\sin\theta\,. \tag{1.56}$$

Since we can also write $z = x + \mathrm{i}\,y$, we therefore have

$$z = r\,(\cos\theta + \mathrm{i}\,\sin\theta)\,. \tag{1.57}$$

Note that $|z|^2 = r^2\,(\cos^2\theta + \sin^2\theta) = r^2$.

Recalling that the power-series expansions of the exponential function, the cosine and the sine functions are given by

$$e^x = \sum_{n\ge 0} \frac{x^n}{n!}\,, \qquad \cos x = \sum_{p\ge 0} \frac{(-1)^p\,x^{2p}}{(2p)!}\,, \qquad \sin x = \sum_{p\ge 0} \frac{(-1)^p\,x^{2p+1}}{(2p+1)!}\,, \tag{1.58}$$

we can see that in particular, in the power series expansion of $e^{\mathrm{i}\,\theta}$ the real terms (even powers of θ assemble into the power series for $\cos\theta$, whilst the imaginary terms (odd powers of θ) assemble into the series for $\sin\theta$. In other words

$$e^{\mathrm{i}\,\theta} = \cos\theta + \mathrm{i}\,\sin\theta\,. \tag{1.59}$$

Turning this around, which can be achieved by adding or subtracting the comlex conjugate, we find

$$\cos\theta = \tfrac{1}{2}(e^{\mathrm{i}\,\theta} + e^{-\mathrm{i}\,\theta})\,, \qquad \sin\theta = \tfrac{1}{2\mathrm{i}}(e^{\mathrm{i}\,\theta} - e^{-\mathrm{i}\,\theta})\,. \tag{1.60}$$

Combining (1.57) and (1.59), we therefore have

$$z = r\,e^{\mathrm{i}\,\theta}\,. \tag{1.61}$$

Note that we can also write this as $z = |z|\, e^{\mathrm{i}\,\theta}$. The angle θ is known as the *phase*, or the *argument*, of the complex number z. When complex numbers are multiplied together, the phases are additive, and so if $z_1 = |z_1|\, e^{\mathrm{i}\,\theta_1}$ and $z_2 = |z_2|\, e^{\mathrm{i}\,\theta_2}$, then

$$z_1\, z_2 = |z_1|\, |z_2|\, e^{\mathrm{i}\,(\theta_1+\theta_2)}\,. \tag{1.62}$$

We may note that the following inequality holds:

$$|z_1 + z_2| \le |z_1| + |z_2|\,. \tag{1.63}$$

This can be seen by calculating the square:

$$\begin{aligned} |z_1 + z_2|^2 &= (\bar z_1 + \bar z_2)(z_1 + z_2) = |z_1|^2 + |z_2|^2 + \bar z_1\, z_2 + \bar z_2\, z_1\,, \\ &= |z_1|^2 + |z_2|^2 + 2|z_1\, |z_2|\, \cos(\theta_1 - \theta_2)\,, \\ &\le |z_1|^2 + |z_2|^2 + 2|z_1\, |z_2| = (|z_1| + |z_2|)^2\,, \end{aligned} \tag{1.64}$$

where we write $z_1 = |z_1|\, e^{\mathrm{i}\theta_1}$ and $z_2 = |z_2|\, e^{\mathrm{i}\theta_2}$. (The inequality follows from the fact that $\cos\theta \le 1$.) By induction, the inequality (1.63) extends to any finite number of terms:

$$|z_1 + z_2 + \cdots + z_n| \le |z_1| + |z_2| + \cdots + |z_n|\,. \tag{1.65}$$

1.2 Analytic or Holomorphic Functions

Having introduced the notion of complex numbers, we can now consider situations where we have a complex function depending on a complex argument. The most general kind of possibility would be to consider a complex function $f = u + \mathrm{i}\, v$, where u and v are themselves real functions of the complex variable $z = x + \mathrm{i}\, y$;

$$f(z) = u(x, y) + \mathrm{i}\, v(x, y)\,. \tag{1.66}$$

As it stands, this notion of a function of a complex variable is too broad, and consequently of limited value. If functions are to be at all interesting, we must be able to differentiate them. Suppose the function $f(z)$ is defined in some region, or domain, D in the complex plane (the two-dimensional plane with Cartesian axes x and y). We would naturally define the derivative of f at a point z_0 in D as the limit of

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{\delta f}{\delta z} \tag{1.67}$$

as z approaches z_0. The key point here, though, is that in order to be able to say "the limit," we must insist that the answer is independent of how we let z approach the point

z_0. The complex plane, being 2-dimensional, allows z to approach z_0 on any of an infinity of different trajectories. We would like the answer to be unique.

A classic example of a function of z whose derivative is not unique is $f(z) = |z|^2 = \bar{z}\, z$. Thus from (1.67) we would attempt to calculate the limit

$$\frac{|z|^2 - |z_0|^2}{z - z_0} = \frac{z\,\bar{z} - z_0\,\bar{z}_0}{z - z_0} = \bar{z} + z_0 \frac{\bar{z} - \bar{z}_0}{z - z_0} \tag{1.68}$$

as z approaches z_0. Now, if we write $z - z_0 = |z - z_0|\, e^{\mathrm{i}\,\theta}$, we see that this becomes

$$\bar{z} + z_0\, e^{-2\mathrm{i}\,\theta} = \bar{z} + z_0\,(\cos 2\theta - \mathrm{i}\, \sin 2\theta)\,, \tag{1.69}$$

which shows that, except at $z_0 = 0$, the answer depends on the angle θ at which z approaches z_0 in the complex plane. One say that the function $|z|^2$ is not *differentiable* except at $z = 0$.

The interesting functions $f(z)$ to consider are those which *are* differentiable in some domain D in the complex plane. Placing the additional requirement that $f(z)$ be *single valued* in the domain, we have the definition of an *analytic* function, sometimes known as a *holomorphic* function. Thus:

A function $f(z)$ is analytic or holomorphic in a domain D in the complex plane if it is single-valued and differentiable everywhere in D.

Let us look at the conditions under which a function is analytic in D. It is easy to derive *necessary* conditions. Suppose first we take the limit in (1.67) in which $z + \delta z$ approaches z along the direction of the real axis (the x axis), so that $\delta z = \delta x$;

$$\frac{\delta f}{\delta z} = \frac{\delta u + \mathrm{i}\,\delta v}{\delta x + \mathrm{i}\,\delta y} = \frac{u_x\,\delta x + \mathrm{i}\, v_x\,\delta x}{\delta x} = u_x + \mathrm{i}\, v_x\,. \tag{1.70}$$

(Clearly for this to be well-defined the partial derivatives $u_x \equiv \partial u/\partial x$ and $v_x \equiv \partial v/\partial x$ must exist.)

Now suppose instead we approach along the imaginary axis, $\delta z = \mathrm{i}\,\delta y$ so that now

$$\frac{\delta f}{\delta z} = \frac{\delta u + \mathrm{i}\,\delta v}{\delta x + \mathrm{i}\,\delta y} = \frac{u_y\,\delta y + \mathrm{i}\, v_y\,\delta y}{\mathrm{i}\,\delta y} = -\mathrm{i}\, u_y + v_y\,. \tag{1.71}$$

(This time, we require that the partial derivatives u_y and v_y exist.) If this is to agree with the previous result from approaching along x, we must have $u_x + \mathrm{i}\, v_x = v_y - \mathrm{i}\, u_y$, which, equating real and imaginary parts, gives

$$u_x = v_y\,, \qquad u_y = -v_x\,. \tag{1.72}$$

These conditions are known as the *Cauchy-Riemann equations*. It is easy to show that we would derive the same conditions if we allowed δz to lie along any ray that approaches z at any angle.

The Cauchy-Riemann equations by themselves are *necessary* but not *sufficient* for the analyticity of the function f. The full statement is the following:

A continuous single-valued function $f(z)$ is analytic or holomorphic in a domain D if the four derivatives u_x, u_y, v_x and v_y exist, are continuous and satisfy the Cauchy-Riemann equations.[1]

There is a nice alternative way to view the Cauchy-Riemann equations. Since $z = x + \mathrm{i}\, y$, and hence $\bar{z} = x - \mathrm{i}\, y$, we may solve to express x and y in terms of z and $\bar{z}$:

$$x = \tfrac{1}{2}(z + \bar{z})\,, \qquad y = -\tfrac{\mathrm{i}}{2}\,(z - \bar{z})\,. \tag{1.73}$$

Formally, we can think of z and $\bar{z}$ as being independent variables. Then, using the chain rule, we shall have

$$\begin{aligned}
\partial_z &\equiv \frac{\partial}{\partial z} = \frac{\partial x}{\partial z}\,\frac{\partial}{\partial x} + \frac{\partial y}{\partial z}\,\frac{\partial}{\partial y} = \tfrac{1}{2}\partial_x - \tfrac{\mathrm{i}}{2}\,\partial_y\,,\\
\partial_{\bar{z}} &\equiv \frac{\partial}{\partial \bar{z}} = \frac{\partial x}{\partial \bar{z}}\,\frac{\partial}{\partial x} + \frac{\partial y}{\partial \bar{z}}\,\frac{\partial}{\partial y} = \tfrac{1}{2}\partial_x + \tfrac{\mathrm{i}}{2}\,\partial_y\,,
\end{aligned} \tag{1.74}$$

where $\partial_x \equiv \partial/\partial x$ and $\partial_y \equiv \partial/\partial y$. (Note that ∂_z means a partial derivative holding $\bar{z}$ fixed, *etc.*) So if we have a complex function $f = u + \mathrm{i}\, v$, then $\partial_{\bar{z}} f$ is given by

$$\partial_{\bar{z}} f = \tfrac{1}{2} u_x + \tfrac{\mathrm{i}}{2}\, u_y + \tfrac{\mathrm{i}}{2} v_x - \tfrac{1}{2}\, v_y\,, \tag{1.75}$$

which vanishes by the Cauchy-Riemann equations (1.72).[2] So the Cauchy-Riemann equations are equivalent to the statement that the function $f(z)$ depends on z but not on $\bar{z}$. We now see instantly why the function $f = |z|^2 = \bar{z}\, z$ was not in general analytic, although it was at the origin, $z = 0$.

We have seen that the real and imaginary parts u and v of an analytic function satisfy the Cauchy-Riemann equations (1.72). From these, it follows that $u_{xx} = v_{yx} = v_{xy} = -u_{yy}$, and similarly for v. In other words, u and v each satisfy Laplace's equation in two dimensions:

$$\nabla^2 u = 0\,, \qquad \nabla^2 v = 0\,, \qquad \text{where } \nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\,. \tag{1.76}$$

[1] A function $f(z)$ is *continuous* at z_0 if, for any given $\epsilon > 0$ (however small), we can find a number δ such that $|f(z) - f(z_0)| < \epsilon$ for all points z in D satisfying $|z - z_0| < \delta$.

[2] One might feel uneasy with treating z and $\bar{z}$ as independent variables, since one is actually the complex conjugate of the other. The proper way to show that it is a valid procedure is temporarily to introduce a genuinely independent complex variable $\tilde{z}$, and to write functions as depending on z and $\tilde{z}$, rather than z and $\bar{z}$. After performing the differentiations in this enlarged complex 2-plane, one then sets $\tilde{z} = \bar{z}$, which amounts to taking the standard "section" that defines the complex plane. It then becomes apparent that one can equally well just treat z and $\bar{z}$ as independent, and cut out the intermediate step of enlarging the dimension of the complex space.

This is a very useful property, since it provides us with ways of solving Laplace's equation in two dimensions. It has applications in 2-dimensional electrostatics and gravity, and in hydrodynamics.

Another very important consequence is that we can use the properties (1.76) in reverse, in the following sense. We have shown that if u is the real part of an analytic function, then it satisfies $\nabla^2 u = 0$. In fact the implication goes in the other direction too; if $u(x, y)$ satisfies the Laplace equation $u_{xx} + u_{yy} = 0$ then it follows that it can be taken to be the real part of some analytic function. We can say that $u_{xx} + u_{yy} = 0$ is the *integrability condition* for the pair of equations $u_x = v_y$, $u_y = -v_x$ to admit a solution for $v(x, y)$.

To solve for $v(x, y)$, one differentates $u(x, y)$ with respect to x or y, and integrates with respect to y or x respectively, to construct the function $v(x, y)$ using (1.72):

$$\begin{aligned} v(x,y) &= \int_{y_0}^{y} \frac{\partial u(x,y')}{\partial x}\, dy' + \alpha(x)\,, \\ v(x,y) &= -\int_{x_0}^{x} \frac{\partial u(x',y)}{\partial y}\, dx' + \beta(y)\,. \end{aligned} \tag{1.77}$$

The first integral, which comes from integrating $u_x = v_y$, leaves an arbitrary function of x unresolved, while the second, coming from integrating $u_y = -v_x$, leaves an arbitrary function of y unresolved. Consistency between the two resolves everything, up to an additive constant in $v(x, y)$. This constant never can be determined purely from the given data, since clearly if $f(z)$ is analytic then so is $f(z) + \mathrm{i}\,\gamma$, where γ is a real constant. But the real parts of $f(z)$ and $f(z) + \mathrm{i}\,\gamma$ are identical, and so clearly we cannot deduce the value of γ, merely from the given $u(x, y)$. Note that we do need both equations in (1.77), in order to determine $v(x, y)$ up to the additive constant γ. Of course the freedom to pick different constant lower limits of integration y_0 and x_0 in (1.77) just amounts to changing the arbitrary functions $\alpha(x)$ and $\beta(y)$, so we can choose y_0 and x_0 in any way we wish.

Let us check this with an example. Suppose we are given $u(x, y) = e^x \cos y$, and asked to find $v(x, y)$. A quick check shows that $u_{xx} + u_{yy} = 0$, so we will not be wasting our time by searching for $v(x, y)$. We have

$$u_x = v_y = e^x \cos y\,, \qquad u_y = -v_x = -e^x \sin y\,, \tag{1.78}$$

and so integrating as in (1.77) we get

$$v(x,y) = e^x \sin y + \alpha(x)\,, \qquad v(x,y) = e^x \sin y + \beta(y)\,. \tag{1.79}$$

Sure enough, the two expressions are compatible, and we see that $\alpha(x) = \beta(y)$. By the standard argument that is the same as one uses in the separation of variables, it must be that $\alpha(x) = \beta(y) = \gamma$, where γ is a (real) constant. Thus we have found that $v(x, y) = e^x \sin y + \gamma$, and so

$$
\begin{aligned}
f(z) &= u + \mathrm{i}\, v = e^x (\cos x + \mathrm{i} \sin y) + \mathrm{i}\,\gamma = e^x\, e^{\mathrm{i}\, y} + \mathrm{i}\,\gamma = e^{x+\mathrm{i}\, y} + \mathrm{i}\,\gamma \\
&= e^z + \mathrm{i}\,\gamma\,. \qquad (1.80)
\end{aligned}
$$

Note that another consequence of the Cauchy-Riemann equations (1.72) is that

$$u_x\, v_x + u_y\, v_y = 0\,, \qquad (1.81)$$

or, in other words,

$$\vec{\nabla} u \cdot \vec{\nabla} v = 0\,, \qquad (1.82)$$

where

$$\vec{\nabla} \equiv (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}) \qquad (1.83)$$

is the 2-dimensional gradient operator. Equation (1.82) says that families of curves in the (x, y) plane corresponding to $u =$ constant and $v =$ constant intersect at right-angles at all points of intersection. This is because $\vec{\nabla} u$ is perpendicular to the lines of constant u, while $\vec{\nabla} v$ is perpendicular to the lines of constant v. Since we are in two dimensions here, it follows that if the perpendicular to line $u =$ constant is perpendicular to the perpendicular to the line $v =$ constant, then the lines $u =$ constant must be perpendicular to the lines $v =$ constant wherever they intersect.

1.2.1 Power Series

A very important concept in complex variable theory is the idea of a power series, and its radius of convergence. We could consider the infinite series $\sum_{n=0}^\infty a_n\, (z - z_0)^n$, but since a simple shift of the origin in the complex plane allows us to take $z_0 = 0$, we may as well make life a little bit simpler by assuming this has been done. Thus, let us consider

$$f(z) = \sum_{n=0}^\infty a_n\, z^n\,, \qquad (1.84)$$

where the a_n are constant coefficients, which may in general be complex.

A useful criterion for convergence of a series is the Cauchy test. This states that if the terms b_n in an infinite sum $\sum_n b_n$ are all non-negative, then $\sum_n b_n$ converges or diverges according to whether the limit of

$$(b_n)^{\frac{1}{n}} \qquad (1.85)$$

is less than or greater than 1, as n tends to infinity.

We can apply this to determine the conditions under which the series (1.84) is *absolutely convergent*. Taking the modulus of (1.84), and using the inequality (1.65), we shall have

$$|f(z)| = \Big|\sum_{n=0}^\infty a_n\, z^n\Big| \le \sum_{n=0}^\infty |a_n\, z^n|\,. \tag{1.86}$$

Thus we consider the series

$$\sum_{n=0}^\infty |a_n|\, |z|^n\,, \tag{1.87}$$

which is clearly a sum of non-negative terms. If this converges, then $|f(z)|$ is finite, and so the series (1.84) is clearly convergent. If

$$|a_n|^{\frac{1}{n}} \longrightarrow 1/R \tag{1.88}$$

as $n \longrightarrow \infty$, then it is evident that the power series (1.84) is absolutely convergent if $|z| < R$, and divergent if $|z| > R$. (As always, the borderline case $|z| = R$ is trickier, and depends on finer details of the coefficients a_n.) The quantity R is called the *radius of convergence* of the series. The circle of radius R (centred on the expansion point $z = 0$ in our case) is called the *circle of convergence*. The series (1.84) is absolutely convergent for any z that lies within the circle of convergence.

We can now establish the following theorem, which is of great importance.

If $f(z)$ is defined by the power series (1.84), then $f(z)$ is an analytic function at every point within the circle of convergence.

This is all about establishing that the power series defining $f(z)$ is differentiable within the circle of convergence. Thus we define

$$\phi(z) = \sum_{n=1}^\infty n\, a_n\, z^{n-1}\,, \tag{1.89}$$

without yet prejudging that $\phi(z)$ is the derivative of $f(z)$. Suppose the series (1.84) has radius of convergence R. It follows that for any ρ such that $0 < \rho < R$, $|a_n\, \rho^n|$ must be bounded, since we know that even the entire infinite sum is bounded. We may say, then, that $|a_n\, \rho^n| < K$ for any n, where K is some positive number. Then, defining $r = |z|$, and $\eta = |h|$, it follows that if $r < \rho$ and $r + \eta < \rho$, we have

$$\frac{f(z+h) - f(z)}{h} - \phi(z) = \sum_{n=0}^\infty a_n \Big(\frac{(z+h)^n - z^n}{h} - n\, z^{n-1}\Big)\,. \tag{1.90}$$

Now expand $(z+h)^n$ using the binomial theorem, i.e.

$$(z+h)^n = z^n + n\, z^{n-1}\, h + \frac{1}{2!}\, n(n-1)\, z^{n-2}\, h^2 + \cdots + n\, z\, h^{n-1} + h^n\,. \tag{1.91}$$

Using the inequality (1.65), we have

$$\begin{aligned}
\Big|\frac{(z+h)^n - z^n}{h} - n\, z^{n-1}\Big| &= \Big|\tfrac{1}{2!}\, n(n-1)\, z^{n-2}\, h + \tfrac{1}{3!}\, n(n-1)(n-2)\, z^{n-3}\, h^2 + \cdots + h^{n-1}\Big|\,,\\
&\le \tfrac{1}{2!}\, n(n-1)\, r^{n-2}\, \eta + \tfrac{1}{3!}\, n(n-1)(n-2)\, r^{n-3}\, \eta^2 + \cdots + \eta^{n-1}\,,\\
&= \frac{(r+\eta)^n - r^n}{\eta} - n\, r^{n-1}\,. \qquad (1.92)
\end{aligned}$$

Hence

$$\begin{aligned}
\sum_{n=0}^\infty |a_n|\, \Big|\frac{(z+h)^n - z^n}{h} - n\, z^{n-1}\Big| &\le K \sum_{n=0}^\infty \frac{1}{\rho^n}\Big[\frac{(r+\eta)^n - r^n}{\eta} - n\, r^{n-1}\Big]\,,\\
&= K\Big[\frac{1}{\eta}\Big(\frac{\rho}{\rho - r - \eta} - \frac{\rho}{\rho - r}\Big) - \frac{\rho}{(\rho-r)^2}\Big]\,,\\
&= \frac{K\, \rho\, \eta}{(\rho - r - \eta)(\rho - r)^2}\,. \qquad (1.93)
\end{aligned}$$

(The summations involved in getting to the second line are simply geometric series, and can be seen from

$$\frac{1}{1-x} = \sum_{n\ge 0} x^n\,, \qquad (1.94)$$

and the series obtained by differentiating this with respect to x.) Clearly the last line in (1.93) tends to zero as η goes to zero. This proves that $\phi(z)$ given in (1.89) is indeed the derivative of $f(z)$. Thus $f(z)$, defined by the power series (1.84), is differentiable within its circle of convergence. Since it is also manifestly single-valued, this means that it is analytic with the circle of convergence.

It is also clear that the derivative $f'(z)$, given, as we now know, by (1.89), is has the same radius of convergence as the original series for $f(z)$. This is because the limit of $|n\, a_n|^{1/n}$ as n tends to infinity is clearly the same as the limit of $|a_n|^{1/n}$. The process of differentiation can therefore be continued to higher and higher derivatives. In other words, a power series can be differentiated term by term as many times as we wish, at any point within its circle of convergence.

1.3 Contour Integration

1.3.1 Cauchy's Theorem

A very important result in the theory of complex functions is *Cauchy's Theorem*, which states:

- *If a function $f(z)$ is analytic, and it is continuous within and on a smooth closed contour C, then*

$$\oint_C f(z)\, dz = 0\,. \qquad (1.95)$$

The symbol $\oint$ denotes that the integration is taken around a closed contour; sometimes, when there is no ambiguity, we shall omit the subscript C that labels this contour.

To see what (1.95) means, consider first the following. Since $f(z) = u(x,y) + \mathrm{i}\, v(x,y)$, and $z = x + \mathrm{i}\, y$, we may write (1.95) as

$$\oint_C f(z)\, dz = \oint_C (u\, dx - v\, dy) + \mathrm{i} \oint_C (v\, dx + u\, dy)\,, \tag{1.96}$$

where we have separated the original integral into its real and imaginary parts. Written in this way, each of the contour integrals can be seen to be nothing but a closed line integral of the kind familiar, for example, in electromagnetism. The only difference here is that we are in two dimensions rather than three. However, we still have the concept Stokes' Theorem, known as *Green's Theorem* in two dimensions, which asserts that

$$\oint_C \vec{E} \cdot d\vec{\ell} = \int_S \vec{\nabla} \times \vec{E} \cdot d\vec{S}\,, \tag{1.97}$$

where C is a closed curve bounding a domain S, and $\vec{E}$ is any vector field defined in S and on C, with well-defined derivatives in S. In two dimensions, the curl operator $\vec{\nabla}\times$ just means

$$\vec{\nabla} \times \vec{E} = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\,. \tag{1.98}$$

(It is effectively just the z component of the three-dimensional curl.) $\vec{E} \cdot d\vec{\ell}$ means $E_x\, dx + E_y\, dy$, and the area element $d\vec{S}$ will just be $dx\, dy$.

Applying Green's theorem to the integrals in (1.96), we therefore obtain

$$\oint_C f(z)\, dz = -\int_S \Big(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\Big)\, dx\, dy + \mathrm{i} \int_S \Big(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\Big)\, dx\, dy\,. \tag{1.99}$$

But the integrands here are precisely the quantities that vanish by virtue of the Cauchy-Riemann equations (1.72), and thus we see that $\oint f(z)\, dz = 0$, verifying Cauchy's theorem.

An alternative proof of Cauchy's theorem can be given as follows. Define first the slightly more general integral

$$F(\lambda) \equiv \lambda \oint f(\lambda z)\, dz\,; \qquad 0 \le \lambda \le 1\,, \tag{1.100}$$

where λ is a constant parameter that can be freely chosen in the interval $0 \le \lambda \le 1$. Cauchy's theorem is therefore the statement that $F(1) = 0$. To show this, first differentiate $F(\lambda)$ with respect to λ:

$$F'(\lambda) = \oint f(\lambda z)\, dz + \lambda \oint z\, f'(\lambda z)\, dz\,. \tag{1.101}$$

(The prime symbol $'$ always means the derivative of a function with respect to its argument.) Now integrate the second term by parts, giving

$$\begin{aligned} F'(\lambda) &= \oint f(\lambda z)\, dz + \lambda\Big([\lambda^{-1}\, z\, f(\lambda z)] - \lambda^{-1} \oint f(\lambda z)\, dz\Big) \\ &= [z\, f(\lambda z)]\,, \end{aligned} \tag{1.102}$$

where the square brackets indicate that we take the difference between the values of the enclosed quantity at the beginning and end of the integration range. But since we are integrating around a closed curve, and since $z\, f(\lambda z)$ is a single-valued function, this must vanish. Thus we have established that $F'(\lambda) = 0$, implying that $F(\lambda) =$ constant. We can determine this constant by considering any value of λ we wish. Taking $\lambda = 0$, it is clear from (1.100) that $F(0) = 0$, whence $F(1) = 0$, proving Cauchy's theorem.

Why did we appear not to need the Cauchy-Riemann equations (1.72) in this proof? The answer, of course, is that effectively we did, since we assumed that we could sensibly talk about the derivative of f, called f'. As we saw when we discussed the Cauchy-Riemann equations, they are the consequence of requiring that $f'(z)$ have a well-defined meaning.

Cauchy's theorem has very important implications in the theory of integration of complex functions. One of these is that if $f(z)$ is an analytic function in some domain D, then if we integrate $f(z)$ from points z_1 to z_2 within D the answer

$$\int_{z_1}^{z_2} f(z)\, dz \tag{1.103}$$

is independent of the path of integration within D. This follows immediately by noting that if we consider two integration paths P_1 and P_2 then the total path consisting of integration from z_1 to z_2 along P_1, and then back to z_1 in the negative direction along P_2 constitutes a closed curve $C = P_1 - P_2$ within D. Thus Cauchy's theorem tells us that

$$0 = \oint_C f(z)\, dz = \int_{P_1} f(z)\, dz - \int_{P_2} f(z)\, dz\,. \tag{1.104}$$

Another related implication from Cauchy's theorem is that it is possible to define an *indefinite integral* of $f(z)$, by

$$g(z) = \int_{z_0}^{z} f(z')\, dz'\,, \tag{1.105}$$

where the contour of integration can be taken to be any path within the domain of analyticity. Notice that the integrated function, $g(z)$, has the same domain of analyticity as the integrand $f(z)$. To show this, we just have to show that the derivative of $g(z)$ is unique. This (almost self-evident) property can be made evident by considering

$$\frac{g(z) - g(\zeta)}{z - \zeta} - f(\zeta)\,, \tag{1.106}$$

and noting that

$$
\begin{aligned}
g(z)-g(\zeta) &= \int_{z_0}^{z} f(z')dz' - \int_{z_0}^{\zeta} f(z')dz' = \int_{\zeta}^{z} f(z')dz'\,, \\
f(\zeta) &= \frac{f(\zeta)}{z-\zeta}\,(z-\zeta) = \frac{f(\zeta)}{z-\zeta}\int_{\zeta}^{z} dz' = \frac{1}{z-\zeta}\int_{\zeta}^{z} f(\zeta)dz'\,, \qquad (1.107)
\end{aligned}
$$

which means that

$$
\frac{g(z)-g(\zeta)}{z-\zeta} - f(\zeta) = \frac{\int_{\zeta}^{z}[f(z')-f(\zeta)]\,dz'}{z-\zeta}\,. \qquad (1.108)
$$

Since $f(z)$ is continuous and single-valued, it follows that $|f(z')-f(\zeta)|$ will tend to zero at least as fast as $|z-\zeta|$ for any point z' on a direct path joining ζ to z, as z approaches ζ. Together with the fact that the integration range itself is tending to zero in this limit, proportional to $|z-\zeta|$, it is evident that the right-hand side in (1.108) will tend to zero as ζ approaches ζ (since the single power of $(z-\zeta)$ in the denominator is outweighed by at least two powers in the numerator), implying therefore that $g'(z)$ exists and is equal to $f(z)$.

A third very important implication from Cauchy's theorem is that *if a function $f(z)$ that* does *contain some sort of singularities within a closed curve C is integrated around C, then the result will be unchanged if the contour is deformed in any way, provided that it does not cross any singularity of $f(z)$ during the deformation.* This property will prove to be invaluable later, when we want to perform explicit evaluations of contour integrals. We prove it by deforming the closed contour C into the new closed contour $\widetilde{C}$, and then joining the two by a narrow "causeway" of infinitesimally-separated parallel paths. This creates a total contour which, by construction, contains no singularities at all. The integrals in and out along the causeway cancel each other in the limit when the separation of the two paths becomes zero, and hence, taking into account the orientations of the two contributions C and $\widetilde{C}$ in the total closed path, we find from Cauchy's theorem that

$$
\oint_C f(z)dz = \oint_{\widetilde{C}} f(z)dz\,. \qquad (1.109)
$$

Finally, on the subject of Cauchy's theorem, let us note that we can turn it around, and effectively use it as a *definition* of an analytic function. This is the content of *Morera's Theorem*, which states:

- *If $f(z)$ is continuous and single-valued within a closed contour C, and if $\oint f(z)\,dz = 0$ for any closed contour within C, then $f(z)$ is analytic within C.*

This can provide a useful way of testing whether a function is analytic in some domain.

1.3.2 Cauchy's Integral Formula

Suppose that the function $f(z)$ is analytic in a domain D. Consider the integral

$$G(a) = \oint_C \frac{f(z)}{z-a}\, dz\,, \tag{1.110}$$

where the contour C is any closed curve within D. There are three cases to consider, depending on whether the point a lies inside, on, or outside the contour of integration C.

Consider first the case when a lies within C. By an observation in the previous section, we know that the value of the integral (1.110) will not alter if we deform the contour in any way provided that the deformation does not cross over the point $z = a$. We can exploit this in order to make life simple, by deforming the contour into a small circle C', of radius ϵ, centred on the point a. Thus we may write

$$z - a = \epsilon\, e^{\mathrm{i}\theta}\,, \tag{1.111}$$

with the deformed contour C' being parameterised by taking θ from 0 to 2π.[3]

Thus we have $dz = \mathrm{i}\,\epsilon\, e^{\mathrm{i}\theta}\, d\theta$, and so

$$G(a) = \mathrm{i} \int_0^{2\pi} f(a + \epsilon\, e^{\mathrm{i}\theta})\, d\theta = \mathrm{i}\, f(a) \int_0^{2\pi} d\theta + \mathrm{i} \int_0^{2\pi} [f(a + \epsilon\, e^{\mathrm{i}\theta}) - f(a)]\, d\theta\,. \tag{1.112}$$

In the limit as ϵ tends to zero, the continuity of the function $f(z)$ implies that the last integral will vanish, since $f(a + \epsilon\, e^{\mathrm{i}\,\theta}) = f(a) + f'(a)\, \epsilon\, e^{\mathrm{i}\,\theta} + \cdots$, and so we have that if $f(z)$ is analytic within and on any closed contour C then

$$\oint_C \frac{f(z)}{z-a}\, dz = 2\pi\, \mathrm{i}\, f(a)\,, \tag{1.113}$$

provided that C contains the point $z = a$. This is *Cauchy's integral formula.*

Obviously if the point $z = a$ were to lie outside the contour C, then we would, by Cauchy's theorem, have

$$\oint_C \frac{f(z)}{z-a}\, dz = 0\,, \tag{1.114}$$

since then the integrand would be a function that was analytic within C.

The third case to consider is when the point a lies exactly on the path of the contour C. It is somewhat a matter of definition, as to how we should handle this case. The most reasonable thing is to decide, like in the Judgement of Solomon, that the point is to be

[3]Note that this means that we define a *positively-oriented* contour to be one whose path runs *anti-clockwise*, in the direction of *increasing* θ. Expressed in a coordinate-invariant way, a positively-oriented closed contour is one for which the interior lies to the *left* as you walk along the path.

viewed as being split into two, with half of it lying inside the contour, and half outside. Thus if a lies on C we shall have

$$\oint_C \frac{f(z)}{z-a}\, dz = \pi\, \mathrm{i}\, f(a)\,. \tag{1.115}$$

We can view the Cauchy integral formula as a way of evaluating an analytic function at a point z in terms of a contour integral around any closed curve C that contains z:

$$f(z) = \frac{1}{2\pi\, \mathrm{i}} \oint_C \frac{f(\zeta)\, d\zeta}{\zeta - z}\,. \tag{1.116}$$

A very useful consequence from this is that we can use it also to express the derivatives of $f(z)$ in terms of contour integrals. Essentially, one just differentiates (1.116) with respect to z, meaning that on the right-hand side it is only the function $(\zeta - z)^{-1}$ that is differentiated. We ought to be a little careful just once to verify that this "differentiation under the integral" is justified, so that having established the validity, we can be cavalier about it in the future. The demonstration is in any case pretty simple. We have

$$\begin{aligned} \frac{f(z+h) - f(z)}{h} &= \frac{1}{2\pi\, \mathrm{i}} \oint \frac{f(\zeta)}{h} \Big(\frac{1}{\zeta - z - h} - \frac{1}{\zeta - z}\Big)\, d\zeta\,, \\ &= \frac{1}{2\pi\, \mathrm{i}} \oint \frac{f(\zeta)\, d\zeta}{(\zeta - z)(\zeta - z - h)}\,. \end{aligned} \tag{1.117}$$

Now in the limit when $h \longrightarrow 0$ the left-hand side becomes $f'(z)$, and thus we get

$$f'(z) = \frac{1}{2\pi\, \mathrm{i}} \oint \frac{f(\zeta)\, d\zeta}{(\zeta - z)^2}\,. \tag{1.118}$$

The question of the validity of this process, in which we have taken the limit $h \longrightarrow 0$ under the integration, comes down to whether it was valid to assume that

$$\begin{aligned} T &\equiv -\oint f(\zeta) \Big(\frac{1}{(\zeta - z)^2} - \frac{1}{(\zeta - z - h)(\zeta - z)}\Big)\, d\zeta \\ &= h \oint \frac{f(\zeta)\, d\zeta}{(\zeta - z)^2\, (\zeta - z - h)} \end{aligned} \tag{1.119}$$

vanishes as h tends to zero.

To show this, let us consider the more general question of obtaining an upper bound on the modulus of a contour integral. Consider the integral

$$G \equiv \oint_C g(z)\, dz\,, \tag{1.120}$$

for some function $g(z)$ integrated around the closed curve C. Using the inequality (1.65), generalised to a sum over an infinity of infinitesimal complex numbers, we have

$$|G| = \Big| \oint_C g(z)\, dz \Big| \le \oint_C |g(z)|\, |dz|\,. \tag{1.121}$$

Now, clearly this last integral, which sums up quantities that are all positive or zero, must itself be less than or equal to the integral

$$\oint_C |g|_{\rm max}\, |dz|\,, \tag{1.122}$$

where we define $|g|_{\rm max}$ to be the largest value that $|g(z)|$ attains anywhere on the contour C. Since this maximum value is of course just a constant, we have that

$$|G| \le |g|_{\rm max} \oint_C |dz|\,. \tag{1.123}$$

Finally, we note that

$$\oint_C |dz| = \oint_C \sqrt{dx^2+dy^2} = L\,, \tag{1.124}$$

where L is the total length of the contour C. Thus we conclude that

$$\Big|\oint_C g(z)\, dz\Big| \le |g|_{\rm max}\, L\,. \tag{1.125}$$

One further useful inequality may be written down in the case that $g(z)$ is itself the ratio of complex functions: $g(z) = A(z)/B(z)$. Obviously, we can now say that the maximum value of $|g(z)|$ anywhere on the contour is bounded by

$$|g|_{\rm max} \le \frac{|A|_{\rm max}}{|B|_{\rm min}}\,, \tag{1.126}$$

where $|A|_{\rm max}$ is the maximum value of $|A(z)|$ anywhere on the contour C, and $|B|_{\rm min}$ is the minimum value of $|B((z)|$ anywhere on the contour C. (In general, these two extrema will occur at different points on the contour.) Thus we may say that

$$\Big|\oint \frac{A(z)}{B(z)}\, dz\Big| \le \frac{|A|_{\rm max}}{|B|_{\rm min}}\, L\,, \tag{1.127}$$

where again L is the total length of the contour.[4]

Returning now to our discussion above, it is evident using (1.127) that

$$|T| \le \frac{|h|\, M\, L}{b^2\, (b-|h|)}\,, \tag{1.128}$$

where M is the maximum value of $|f(\zeta)|$ on the contour, L is the length of the contour, and b is the minimum value of of $|\zeta - z|$ on the contour. These are all fixed numbers, independent of h, and so we see that indeed T must vanish as h is taken to zero.

[4]Note, by the way, that although we presented the argument above for the case of a closed contour, the bounds (1.125) and (1.127) apply equally well to the case of an open contour that does not close on itself. Of course, they will only be useful bounds if the length L is finite.

More generally, by continuing the above procedure, we can show that the n'th derivative of $f(z)$ is given by

$$f^{(n)}(z) = \frac{1}{2\pi\,\mathrm{i}} \oint f(\zeta)\, \frac{d^n}{dz^n}\Big(\frac{1}{\zeta - z}\Big)\, d\zeta\,, \tag{1.129}$$

or, in other words,

$$f^{(n)}(z) = \frac{n!}{2\pi\,\mathrm{i}} \oint_C \frac{f(\zeta)\, d\zeta}{(\zeta - z)^{n+1}}\,. \tag{1.130}$$

Note that since all the derivatives of $f(z)$ exist, for all point C within the contour C, it follows that $f^{(n)}(z)$ is analytic within C for any n.

1.3.3 The Taylor Series

We can use Cauchy's integral formula to derive Taylor's theorem for the expansion of a function $f(z)$ around a point $z = a$ at which $f(z)$ is analytic. An important outcome from this will be that we shall see that the radius of convergence of the Taylor series extends up to the singularity of $f(z)$ that is nearest to $z = a$.

From Cauchy's integral formula we have that if $f(z)$ is analytic inside and on a contour C, and if $z = a + h$ lies inside C, then

$$f(a+h) = \frac{1}{2\pi\,\mathrm{i}} \oint \frac{f(\zeta)\, d\zeta}{\zeta - a - h}\,. \tag{1.131}$$

Now, bearing in mind that the geometric series $\sum_{n=0}^N x^n$ sums to give $(1 - x^{N+1})\,(1-x)^{-1}$, we have that

$$\sum_{n=0}^N \frac{h^n}{(\zeta - a)^{n+1}} = \frac{1}{\zeta - a - h} - \frac{h^{N+1}}{(\zeta - a - h)\,(\zeta - a)^{N+1}}\,. \tag{1.132}$$

We can use this identity as an expression for $(\zeta - a - h)^{-1}$ in (1.131), implying that

$$f(a+h) = \sum_{n=0}^N \frac{h^n}{2\pi\,\mathrm{i}} \oint \frac{f(\zeta)\, d\zeta}{(\zeta - a)^{n+1}} + \frac{h^{N+1}}{2\pi\,\mathrm{i}} \oint \frac{f(\zeta)\, d\zeta}{(\zeta - a - h)\,(\zeta - a)^{N+1}}\,. \tag{1.133}$$

In other words, in view of our previous result (1.130), we have

$$f(a+h) = \sum_{n=0}^N \frac{h^n}{n!}\, f^{(n)}(a) + R_N\,, \tag{1.134}$$

where the "remainder" term R_N is given by

$$R_N = \frac{h^{N+1}}{2\pi\,\mathrm{i}} \oint_C \frac{f(\zeta)\, d\zeta}{(\zeta - a - h)\,(\zeta - a)^{N+1}}\,. \tag{1.135}$$

Now, if M denotes the maximum value of $|f(\zeta)|$ on the contour C, then by taking C to be a circle of radius r centred on $\zeta = a$, we shall have

$$|R_N| \le \frac{|h|^{N+1}\, M\, r}{(r - |h|)\, r^{N+1}} = \frac{M\, r}{r - |h|}\, \Big(\frac{|h|}{r}\Big)^{N+1}\,. \tag{1.136}$$

Thus if we choose h such that $|h| < r$, it follows that as N is sent to infinity, R_N will go to zero. This means that the Taylor series

$$f(a+h) = \sum_{n=0}^{\infty} \frac{h^n}{n!}\, f^{(n)}(a)\,, \tag{1.137}$$

or in other words,

$$f(z) = \sum_{n=0}^{\infty} \frac{(z-a)^n}{n!}\, f^{(n)}(a)\,, \tag{1.138}$$

will be convergent for any z lying within the circle of radius r centred on $z = a$. But we can choose this circle to be as large as we like, provided that it does not enclose any singularity of $f(z)$. Therefore, it follows that the radius of convergence of the Taylor series (1.137) is precisely equal to the distance between $z = a$ and the nearest singularity of $f(z)$.

1.3.4 The Laurent Series

Suppose now that we want to expand $f(z)$ around a point $z = a$ where $f(z)$ has a singularity. Clearly the previous Taylor expansion will no longer work. However, depending upon the nature of the singularity at $z = a$, we may be able to construct a more general kind of series expansion, known as a Laurent series. To do this, consider two concentric circles C_1 and C_2, centred on the point $z = a$, where C_1 has a larger radius that can be taken out as far as possible before hitting the next singularity of $f(z)$, while C_2 is an arbitrarily small circle enclosing a. Take the path C_1 to be anticlockwise, while the path C_2 is clockwise. We can make C_1 and C_2 into a single closed contour C, by joining them along a narrow "causeway," as shown in Figure 1.

The idea is that we will take a limit where the width of the "causeway" joining the inner and outer circles shrinks to zero. In the region of the complex plane under discussion, the function $f(z)$ has, by assumption, only an isolated singularity at $z = a$.

Now consider Cauchy's integral formula for this contour, which will give

$$f(a+h) = \frac{1}{2\pi\,\mathrm{i}} \oint_C \frac{f(\zeta)\, d\zeta}{\zeta - a - h}\,. \tag{1.139}$$

The reason for this is that the closed contour C encloses no singularities except for the pole at $\zeta = a + h$. In particular, it avoids the singularity of $f(z)$ at $z = a$. Since the integrand is non-singular in the neighbourhood of the "causeway," we see that when the width of the causeway is taken to zero, we shall find that the integration along the lower "road" heading in from C_1 to C_2 will be exactly cancelled by the integration in the opposite direction along the upper "road" heading out from C_2 to C_1. Furthermore, in the limit when the width

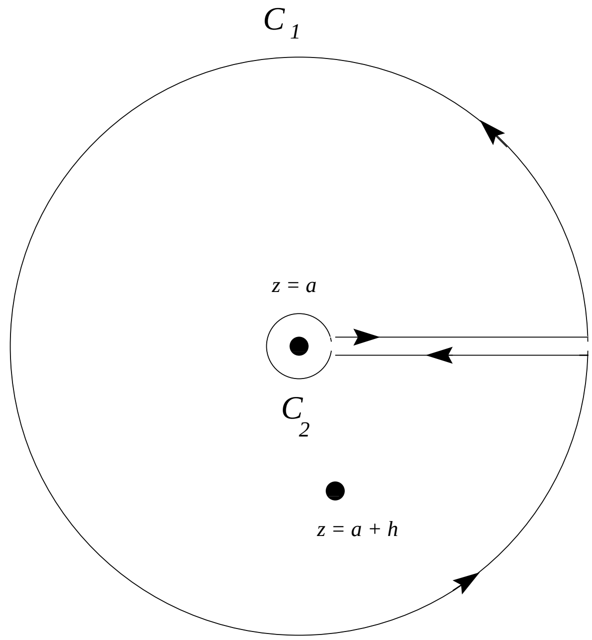

Figure 1: The contour $C = C_1 + C_2$ for Cauchy's integral

of the causeway goes to zero, the gaps in the contours C_1 and C_2 shrink to zero, and they can be replaced by closed circular contours. In this sense, therefore, we can disregard the contribution of the causeway, and just make the replacement that

$$\oint_C \longrightarrow \oint_{C_1} + \oint_{C_2} \,. \tag{1.140}$$

We can therefore write (1.139) as

$$f(a+h) = \frac{1}{2\pi\,\mathrm{i}} \oint_{C_1} \frac{f(\zeta)\, d\zeta}{\zeta - a - h} + \frac{1}{2\pi\,\mathrm{i}} \oint_{C_2} \frac{f(\zeta)\, d\zeta}{\zeta - a - h} \,. \tag{1.141}$$

For the first integral, around the large circle C_1, we can use the same expansion for $(\zeta - a - h)^{-1}$ as we used in the Taylor series previously, obtained by setting $N = \infty$ in (1.132), and using the fact that the second term on the right-hand side then vanishes, since $h^{N+1}/|\zeta - a|^{N+1}$ goes to zero on C_1 when N goes to infinity, as a result of the radius of C_1 being larger than $|h|$. In other words, we expand $(\zeta - a - h)^{-1}$ as

$$\begin{aligned} \frac{1}{\zeta - a - h} &= \frac{1}{(\zeta - a)(1 - h\,(\zeta - a)^{-1})}\,, \\ &= \frac{1}{\zeta - a}\Big(1 + \frac{h}{\zeta - a} + \frac{h^2}{(\zeta - a)^2} + \cdots\Big)\,, \\ &= \sum_{n=0}^{\infty} \frac{h^n}{(\zeta - a)^{n+1}}\,. \end{aligned} \tag{1.142}$$

On the other hand, in the second integral in (1.141) we can expand $(\zeta - a - h)^{-1}$ in a series valid for $|\zeta - a| << |h|$, namely

$$\begin{aligned}
\frac{1}{\zeta - a - h} &= -\frac{1}{h(1 - (\zeta - a)\, h^{-1})}\,, \\
&= -\frac{1}{h}\Big(1 + \frac{\zeta - a}{h} + \frac{(\zeta - a)^2}{h^2} + \cdots\Big)\,, \\
&= -\sum_{m=1}^{\infty} \frac{(\zeta - a)^{m-1}}{h^m}\,.
\end{aligned} \tag{1.143}$$

Thus we find

$$f(a+h) = \frac{1}{2\pi\,\mathrm{i}} \sum_{n=0}^{\infty} h^n \oint_{C_1} \frac{f(\zeta)\, d\zeta}{(\zeta - a)^{n+1}} + \frac{1}{2\pi\,\mathrm{i}} \sum_{m=1}^{\infty} \frac{1}{h^m} \oint_{C_2^+} f(\zeta)\, (\zeta - a)^{m-1}\, d\zeta\,, \tag{1.144}$$

where we define C_2^+ to mean the contour C_2 but with the direction of the integration path reversed, i.e. C_2^+ runs *anti-clockwise* around the point $\zeta = a$, which means it is now the standard *positive* direction for a contour. Thus we have

$$f(a+h) = \sum_{n=-\infty}^{\infty} a_n\, h^n\,, \tag{1.145}$$

where the coefficients a_n are given by

$$a_n = \frac{1}{2\pi\,\mathrm{i}} \oint \frac{f(\zeta)\, d\zeta}{(\zeta - a)^{n+1}}\,. \tag{1.146}$$

Here, the integration contour is C_1 when evaluating a_n for $n \ge 0$, and C_2^+ when evaluating a_n for $n < 0$. Notice that we can in fact just as well choose to use the contour C_1 for the $n < 0$ integrals too, since the deformation of the contour C_2^+ into C_1 does not cross any singularities of the integrand when $n < 0$.

Note that using the original variable $z = a + h$, (1.145) is written as

$$f(z) = \sum_{n=-\infty}^{\infty} a_n\, (z - a)^n\,. \tag{1.147}$$

The expansion in (1.147) is known as the *Laurent Series*. By similar arguments to those we used for the Taylor series, one can see that the series converges in an annulus whose larger radius is defined by the contour C_1. This contour could be chosen to be the largest possible circle centred on the singularity of $f(z)$ at $z = a$ that does not enclose any other singularity of $f(z)$.

In the Laurent series, the function $f(z)$ has been split as the sum of two parts:

$$f(z) = f_+(z) + f_-(z)\,, \tag{1.148}$$

$$f_+(z) \equiv \sum_{n \ge 0} a_n\, (z - a)^n\,, \qquad f_-(z) \equiv \sum_{m \ge 1} \frac{a_{-m}}{(z - a)^m}\,. \tag{1.149}$$

The part $f_+(z)$ (the terms with $n \ge 0$ in (1.147)) is analytic everywhere inside the larger circle C_1. The part $f_-(z)$ (the terms with $n \le -1$ in (1.147)) is analytic everywhere outside the small circle C_2 enclosing the singularity as $z = a$.

In practice, one commonly wants to work out just the first few terms in the Laurent expansion of a function around a singular point. For example, it is often of interest to know the singular terms, corresponding to the inverse powers of $(z-a)$ in (1.147). If the function in question has a pole of degree N at the expansion point, then there will just be N singular terms, corresponding to the powers $(z-a)^{-N}$ down to $(z-a)^{-1}$. For reasons that we shall see later, the coefficient of $(z-a)^{-1}$ is often of particular interest.

Determining the first few terms in the expansion of a function with a pole at $z = a$ is usually pretty simple, and can just be done by elementary methods. Suppose, for example, we have the function

$$f(z) = \frac{g(z)}{z^N}\,, \tag{1.150}$$

where $g(z)$ is analytic, and that we want to find the Laurent expansion around the point $z = 0$. Since $g(z)$ is analytic, it has a Taylor expansion, which we can write as

$$g(z) = \sum_{m\ge 0} b_m\, z^m\,. \tag{1.151}$$

The Laurent expansion for $f(z)$ is therefore

$$\begin{aligned} f(z) &= \frac{1}{z^N} \sum_{m\ge 0} b_m\, z^m \\ &= \sum_{m\ge 0} b_m\, z^{m-N} \\ &= \sum_{n\ge -N} b_{n+N}\, z^n\,. \end{aligned} \tag{1.152}$$

For example, the Laurent expansion of $f(z) = z^{-2}\, e^z$ is given by

$$\begin{aligned} f(z) &= \frac{1}{z^2}\, \Big(1 + z + \tfrac12 z^2 + \tfrac16 z^3 + \cdots\Big) \\ &= \frac{1}{z^2} + \frac{1}{z} + \tfrac12 + \tfrac16 z + \cdots\,. \end{aligned} \tag{1.153}$$

In more complicated examples, there might be an analytic function that goes to zero in the denominator of the function $f(z)$. We can still work out the first few terms in the Laurent expansion by elementary methods, by writing out the Taylor expansion of the function in the denominator. Consider, for example, the function $f(z) = 1/\sin z$, to be expanding in a Laurent series around $z = 0$. We just write out the first few terms in the

Taylor series for $\sin z$,

$$\sin z = z - \tfrac{1}{6} z^3 + \tfrac{1}{120} z^5 + \cdots$$
$$= z \left(1 - \tfrac{1}{6} z^2 + \tfrac{1}{120} z^4 + \cdots\right) . \qquad (1.154)$$

Notice that on the second line, we have pulled out the overall factor of z, so that what remains inside the parentheses is an analytic function that does not go to zero at $z = 0$. Now, we write

$$f(z) = \frac{1}{\sin z} = \frac{1}{z} \left(1 - \tfrac{1}{6} z^2 + \tfrac{1}{120} z^4 + \cdots\right)^{-1} , \qquad (1.155)$$

and the problem has reduced to the kind we discussed previously. Making the expansion of the term in parentheses using $(1 + x)^{-1} = 1 - x + x^2 - x^3 + \cdots$, we get

$$f(z) = \frac{1}{z} \left(1 + \tfrac{1}{6} z^2 + \tfrac{7}{360} z^4 + \cdots\right) , \qquad (1.156)$$

and hence the Laurent expansion is

$$\frac{1}{\sin z} = \frac{1}{z} + \tfrac{1}{6} z + \tfrac{7}{360} z^3 + \cdots . \qquad (1.157)$$

Note that if we had only wanted to know the pole term, we would not have needed to push the series expansion as far as we just did. So as a practical tip, time can be saved by working just to the order needed, and no more, when performing the expansion. (One must take care, though, to be sure to take the expansion far enough.)

1.4 Classification of Singularities

We are now in a position to classify the types of singularity that a function of a complex variable may possess.

Suppose that $f(z)$ has a singularity at $z = a$, and that its Laurent expansion for $f(a+h)$, given in general in (1.145), actually terminates at some specific negative value of n, say $n = -N$. Thus we have

$$f(a + h) = \sum_{n=-N}^{\infty} a_n \, h^n . \qquad (1.158)$$

We then say that $f(z)$ has a *pole* of order N at $z = a$. In other words, as z approaches a the function $f(z)$ has the behaviour

$$f(z) = \frac{a_{-N}}{(z - a)^N} + \text{ less singular terms} . \qquad (1.159)$$

If, on the other hand, the sum over negative values of n in (1.145) does not terminate, but goes on to $n = -\infty$, then the function $f(z)$ has an *essential singularity* at $z = a$. A classic example is the function

$$f(z) = e^{\frac{1}{z}}\,. \tag{1.160}$$

This has the Laurent expansion

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{n!\, z^n} \tag{1.161}$$

around $z = 0$, which is obtained simply by taking the usual Taylor expansion of

$$e^w = \sum_{n\ge 0} \frac{w^n}{n!} \tag{1.162}$$

and setting $w = 1/z$. The Laurent series (1.161) has terms in arbitrarily negative powers of z, and so $z = 0$ is an essential singularity.

Functions have quite a complicated behaviour near an essential singularity. For example, if z approaches zero along the positive real axis, $e^{1/z}$ tends to infinity. On the other hand, if the approach to zero is along the negative real axis, $e^{1/z}$ instead tends to zero. An approach to $z = 0$ along the imaginary axis causes $e^{1/z}$ to have unit modulus, but with an ever-increasing phase rotation. In fact a function $f(z)$ with an essential singularity can take on any value, for z near to the singular point.

Note that the Laurent expansion (1.145) that we have been discussing here is applicable only if the singularity of $f(z)$ is an *isolated* one.[5] There can also exist singularities of a different kind, which are neither poles nor essential singularities. Consider, for example, the functions $f(z) = \sqrt{z}$, or $f(z) = \log z$. Neither of these can be expanded in a Laurent series around $z = 0$. They are both discontinuous along an entire semi-infinite line starting from the point $z = 0$. Thus the singularity at $z = 0$ is not an isolated one; it is called a *branch point*. We shall discuss these in more detail later.

For now, just take note of the fact that a singularity in a function does not necessarily mean that the function is infinite there. *By definition, a function $f(z)$ is singular at $z = a$ if it is not analytic at $z = a$.* Thus, for example, $f(z) = z^{1/2}$ is singular at $z = 0$, even though $f(0) = 0$. This can be seen from the fact that we cannot expand $z^{1/2}$ as a power series around $z = 0$, and therefore $z^{1/2}$ cannot be analytic there. It is also the case that although $f(0)$ is finite, the derivatives of $f(z)$ are infinite at $z = 0$.

For now, let us look in a bit more detail at functions with isolated singularities.

[5]By definition, if a function $f(z)$ has a singularity at $z = a$, then it is an isolated singularity if $f(z)$ can be expanded in a Laurent series around $z = a$.

1.4.1 Entire Functions

A very important, and initially perhaps rather surprising, result is the following, known as *Liouville's Theorem*:

> *A function $f(z)$ that is analytic for all finite values of z and is bounded everywhere is a constant.*

Note that when we say $f(z)$ is bounded everywhere (at finite z), we mean that there exists some positive number S, which is independent of z, such that

$$|f(z)| \le S \tag{1.163}$$

for all finite z.

We can prove Liouville's theorem using the result obtained earlier from Cauchy's integral formula, for $f'(a)$:

$$f'(a) = \frac{1}{2\pi\,\mathrm{i}} \oint \frac{f(z)\,dz}{(z-a)^2}\,. \tag{1.164}$$

Take the contour of integration to be a circle of radius R centred on $z = a$, which means that the points z on the contour are defined by $|z - a| = R$. Since we are assuming that $f(z)$ is bounded, we may take $|f(z)| \le M$ for all points z on the contour, where M is some finite positive number. Then, using (1.164), we must have

$$|f'(a)| \le \Big(\frac{M}{2\pi\,R^2}\Big)\,(2\pi\,R) = \frac{M}{R}\,. \tag{1.165}$$

Thus by taking R to infinity, and recalling our assumption that $f(z)$ remains bounded for all finite z (meaning that M is finite, in fact $M \le S$, no matter how large R is), we see that $f'(a)$ must be zero. Thus $f(a)$ is a constant, independent of a. Thus Liouville's theorem is established.

An illustration of Liouville's theorem can be given with the following example. Suppose we try to construct an analytic function that is well-behaved, and bounded, everywhere. If we were considering real functions as opposed to complex analytic functions, we might consider a function such as

$$f(x) = \frac{1}{1+x^2}\,, \tag{1.166}$$

which rather boringly falls off to zero as x tends to $+\infty$ or $-\infty$, having attained the exciting peak of $f = 1$ at the origin. Thus as a real function of x, we have $|f(x)| \le 1$ everywhere. However, viewed as a function of the variable z in the complex plane, it is unbounded:

$$f(z) = \frac{1}{1+z^2} = \frac{1}{(z-\mathrm{i})(z+\mathrm{i})} = \frac{\mathrm{i}}{2(z+\mathrm{i})} - \frac{\mathrm{i}}{2(z-\mathrm{i})},\,. \tag{1.167}$$

Thus the function $f(z)$ actually has poles at $z = \pm \mathrm{i}$, away from the z axis. Of course we could consider instead the function

$$g(z) = \frac{1}{1 + |z|^2}\,, \tag{1.168}$$

which certainly satisfies $|g(z)| \le 1$ everywhere. But $g(z)$ is not analytic (since it depends on $\bar{z}$ as well as z.)

Liouville's theorem tells us that any bounded analytic function we try to construct is inevitably going to have singularities somewhere, unless we are content with the humble constant function.

A similar argument to the above allows us to extend Liouville's theorem to the following:

> *If $f(z)$ is analytic for all finite z, and if $|f(z)|$ is bounded by $S\,|z|^k$ for some integer k and some constant S*, i.e. $|f(z)/z^k| \le S$ for all finite z, then $f(z)$ is a polynomial of degree $\le k$.

To show this, we follow the same strategy as before, by using the higher-derivative consequences of Cauchy's integral:

$$f^{(n)}(a) = \frac{n!}{2\pi\,\mathrm{i}} \oint \frac{f(z)\,dz}{(z-a)^{n+1}}\,. \tag{1.169}$$

Assume that $|f(z)| \le M\,|z|^k$ on the contour at radius R centred on $z = a$. Then we have

$$|f^{(n)}(a)| \le \Big(\frac{n!\,M\,R^k}{2\pi\,R^{n+1}}\Big)\,(2\pi\,R) = n!\,M\,R^{k-n}\,. \tag{1.170}$$

Thus we see that as R tends to infinity, all the terms with $k < n$ will vanish (since we shall always have $M \le S$, where S is some fixed number), and so

$$f^{(n)}(a) = 0\,, \qquad \text{for } \; n > k\,. \tag{1.171}$$

But this is just telling us that $f(z)$ is a polynomial in z with z^k as its highest power, which proves the theorem. Liouville's theorem itself is just the special case $k = 0$.

A function $f(z)$ that is a polynomial in z of degree k,

$$f(z) = \sum_{n=0}^{k} a_n\, z^n\,, \tag{1.172}$$

is clearly analytic for all finite values of z. However, if $k > 0$ it will inevitably have a pole at infinity. To see this, we use the usual trick of making the coordinate transformation

$$\zeta = \frac{1}{z}\,, \tag{1.173}$$

and then looking at the behaviour of the function $f(1/\zeta)$ at $\zeta = 0$. Clearly, for a polynomial of degree k of the form (1.172), we shall get

$$f(1/\zeta) = \sum_{n=0}^{k} a_n \, \zeta^{-n} \,, \tag{1.174}$$

implying that there are poles of orders up to and including k at $z = \infty$.

Complex functions that are analytic in every finite region in the complex plane are called *entire functions*. All polynomials, as we have seen, are therefore entire functions. Another example is the exponential function e^z, defined by the power-series expansion

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \,. \tag{1.175}$$

By the Cauchy test for the convergence of series, we see that $(|z|^n/n!)^{1/n} = |z| \, (n!)^{-1/n}$ tends to zero as n tends to infinity,[6] for any finite $|z|$, and so the exponential is analytic for all finite z. Of course the situation at infinity is another story; here, one has to look at $e^{1/\zeta}$ as ζ tends to zero, and as we saw previously this has an essential singularity, which is more divergent than any finite-order pole. Other examples of entire functions are $\cos z$, and the Bessel function of integer order, $J_n(z)$. The Bessel function has the power-series expansion

$$J_n(z) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell! \, (n+\ell)!} \Big(\frac{z}{2}\Big)^{n+2\ell} \,. \tag{1.176}$$

Of course we know from Liouville's theorem that any interesting entire function (i.e. anything except the purely constant function) must have some sort of singularity at infinity.

1.4.2 Meromorphic Functions

Entire functions are analytic everywhere except at infinity. Next on the list are *meromorphic functions*:

> *A Meromorphic Function $f(z)$ is analytic everywhere in the complex plane (including infinity), except for isolated poles.*

[6]To see this, we may, for convenience, take n to be even, in which case we may write

$$n! = [1 \cdot n][2 \cdot (n-1)][3 \cdot (n-2)] \cdots [\tfrac{1}{2}n \cdot (\tfrac{1}{2}n + 1)] \,.$$

Each of the $\frac{1}{2}n$ square brackets is $\ge n$, and so we have $n! \ge n^{n/2}$. It follows that $(n!)^{1/n} \ge n^{1/2}$, and hence $(n!)^{-1/n} \le n^{-1/2}$, which tends to zero as n tends to infinity.

We insist, in the definition of a meromorphic function, that the only singularities that are allowed are *poles*, and not, for example, *essential singularities.* Note that we also insist, in this definition of a strictly meromorphic function, that it either be analytic also at infinity, or at worst, have a pole at infinity.

The number of poles in a meromorphic function must be finite. This follows from the fact that if there were an infinite number then there would exist some singular point, either at finite z or at $z = \infty$, which would not be isolated, thus contradicting the definition of an everywhere-meromorphic function. For example; suppose we had a function with poles at all the integers along the real axis. These would appear to be isolated, since each one is unit distance from the next. However, these poles actually have an accumulation point at infinity, as can be seen by writing $z = 1/\zeta$ and looking near $\zeta = 0$. Thus a function of this type will actually have a bad singularity at infinity, We shall in fact be studying such an example later.

Any meromorphic function $f(z)$ can be written as a ratio of two polynomials. Such a ratio is known as a *rational function.* To see why we can always write $f(z)$ in this way, we have only to make use of the observation above that the number of poles must be finite. Let the number of poles at finite z be N. Thus at a set of N points z_n in the complex plane, the function $f(z)$ has poles of orders d_n. It follows that the function

$$g(z) \equiv f(z) \prod_{n=1}^{N} (z - z_n)^{d_n} \tag{1.177}$$

must be analytic everywhere (except possibly at infinity), since we have cleverly arranged to cancel out every pole at finite z. Even if $f(z)$ does have a pole at infinity, it follows from (1.177) that $g(z)$ will diverge no faster than $|z|^k$ for some finite integer k. But we saw earlier, in the generalisation of Liouville's theorem, that any such function must be a polynomial of degree $\le k$. Thus we conclude that $f(z)$ is a ratio of polynomials:

$$f(z) = \frac{g(z)}{\prod_{n=1}^{N} (z - z_n)^{d_n}} \,. \tag{1.178}$$

The fact that a meromorphic function can be expressed as a ratio of polynomials can be extremely useful.

A ratio of two polynomials can be expanded out as a sum of partial fractions. For example

$$\frac{1+z^2}{1-z^2} = \frac{1}{z+1} - \frac{1}{z-1} - 1 \,. \tag{1.179}$$

Therefore it follows that a function $f(z)$ that is meromorphic can be expanded out as a sum of partial fractions in that region. For a strictly meromorphic function, this sum will be a finite one (since there are only finitely many poles, each of finite order).

Having introduced the notion of a strictly meromorphic function, it is also useful to introduce a slightly less strict notion of meromorphicity. Thus, we can define the notion of a function that is meromorphic *within a restricted region.* Thus a function is said to be meromorphic in a domain D in the complex plane if it is analytic except for pole singularities in the domain D. The previous definition of a meromorphic function thus corresponds to the case where D is the entire complex plane, *including infinity.* A very common situation for a more restricted meromorphic function is when we consider functions that are *meromorphic in the finite complex plane.* Such functions are analytic except for isolated pole singularities everywhere in the finite complex plane, but they are allowed to have "worse" singularities (such as essential singularities) at infinity. Notice in particular that such a function *is* now allowed to have an infinite number of isolated poles in the finite complex plane (since we are now allowing there to be an accumulation point at infinity).

Let us consider an example of a function $f(z)$ that is meromorphic in some region, and furthermore where every pole is of order 1. This is in fact a very common circumstance. As a piece of terminology, a pole of order 1 is also known as a *simple pole.* Let us assume that the poles are located at the points z_n, numbered in increasing order of distance from the origin. Thus near $z = z_n$, we shall have

$$f(z) \sim \frac{b_n}{z - z_n}\,, \tag{1.180}$$

where the constant b_n characterises the "strength" of the pole. In fact b_n is known as the *residue* at the pole $z = z_n$.

Consider a circle C_p centred on $z = 0$ and with radius R_p chosen so that it encloses p of the poles. To avoid problems, we choose R_p so that it does not pass through any pole. Then the function

$$G_p(z) \equiv f(z) - \sum_{n=1}^{p} \frac{b_n}{z - z_n} \tag{1.181}$$

will be analytic within the circle, since we have explicitly arranged to subtract out all the poles (which we are assuming all to be of order 1). Using Cauchy's integral, we shall therefore have

$$G_p(z) = \frac{1}{2\pi\,\mathrm{i}} \oint_{C_p} \frac{G_p(\zeta)\, d\zeta}{\zeta - z} = \frac{1}{2\pi\,\mathrm{i}} \oint_{C_p} \frac{f(\zeta)\, d\zeta}{\zeta - z} - \frac{1}{2\pi\,\mathrm{i}} \sum_{n=1}^{p} b_n \oint_{C_p} \frac{d\zeta}{(\zeta - z)(\zeta - z_n)}\,. \tag{1.182}$$

Now, each term in the sum here integrates to zero. This is because the integrand is

$$\frac{1}{(\zeta - z)(\zeta - z_n)} = \frac{1}{z - z_n} \Big[\frac{1}{\zeta - z} - \frac{1}{\zeta - z_n}\Big] \tag{1.183}$$

The integral (over ζ) is taken around a contour that encloses both the simple pole at $\zeta = z$ and the simple pole at $\zeta = z_n$. We saw earlier, in the proof of Cauchy's integral formula, that a contour integral running anti-clockwise around a simple pole $c/(\zeta - \zeta_0)$ gives the answer $2\pi\, c\,\mathrm{i}$, and so the result of integrating (1.183) around our contour is $(2\pi\,\mathrm{i} - 2\pi\,\mathrm{i})/(z - z_n) = 0$. Thus we conclude that

$$G_p(z) = \frac{1}{2\pi\,\mathrm{i}} \oint_{C_p} \frac{f(\zeta)\, d\zeta}{\zeta - z}\,. \tag{1.184}$$

Now, consider a sequence of ever-larger circles C_p, enclosing larger and larger numbers of poles. This defines a sequence of functions $G_p(z)$ for increasing p, each of which is analytic within R_p. We want to show that $G_p(z)$ is bounded as p tends to infinity, which will allow us to invoke Liouville's theorem and deduce that $G_\infty(z) =$ constant. By a now-familiar method of argument, we suppose that M_p is the maximum value that $|f(\zeta)|$ attains anywhere on the circular contour of radius R_p. Then from (1.184), and using (1.125), we shall have

$$|G_p(z)| \le \frac{M_p\, R_p}{R_p - |z|}\,. \tag{1.185}$$

Consider first the case of a function f for which M_p is bounded in value as R_p goes to infinity. Then, we see from (1.185) that $|G_p(z)|$ is bounded as p goes to infinity. By Liouville's theorem, it follows that $G_\infty(z)$ must just be a constant, c. Thus in this case we have

$$f(z) = c + \sum_{n=1}^{\infty} \frac{b_n}{z - z_n}\,. \tag{1.186}$$

We are left with one undetermined constant, c. This can be fixed by looking at one special value of z, and then equating the two sides in (1.186). Suppose, for example, that $f(z)$ is analytic at $z = 0$. We can then determine c by setting $z = 0$:

$$f(0) = c - \sum_{n=1}^{\infty} \frac{b_n}{z_n}\,, \tag{1.187}$$

and then plugging the solution for c back into (1.186), giving

$$f(z) = f(0) + \sum_{n=1}^{\infty} \Big[\frac{b_n}{z - z_n} + \frac{b_n}{z_n}\Big]\,. \tag{1.188}$$

(If $f(z)$ happens to have a pole at $z = 0$, then we just choose some other special value of z instead, when solving for c.)

We obtained this result by assuming that $f(z)$ was bounded on the circle of radius R_p, as R_p was sent to infinity. Even if this is not the case, one can often construct a related function, for example $f(z)/z^k$ for some suitable integer k, which *is* bounded on the circle. With appropriate minor modifications, a formula like (1.188) can then be obtained.

An example is long overdue. Consider the function $f(z) = \tan z$. which is, of course $(\sin z)/\cos z$. Now we have

$$\begin{aligned}\sin z &= \sin(x + \mathrm{i}\, y) = \sin x \, \cosh y + \mathrm{i}\, \cos x \, \sinh y\,,\\ \cos z &= \cos(x + \mathrm{i}\, y) = \cos x \, \cosh y - \mathrm{i}\, \sin x \, \sinh y\,,\end{aligned} \tag{1.189}$$

where we have used the standard results that $\cos(\mathrm{i}\, y) = \cosh y$ and $\sin(\mathrm{i}\, y) = \mathrm{i}\, \sinh y$. Thus we have

$$\begin{aligned}|\sin z|^2 &= \sin^2 x \, \cosh^2 y + \cos^2 x \, \sinh^2 y = \sin^2 x + \sinh^2 y\,,\\ |\cos z|^2 &= \cos^2 x \, \cosh^2 y + \sin^2 x \, \sinh^2 y = \cos^2 x + \sinh^2 y\,.\end{aligned} \tag{1.190}$$

It is evident that $|\sin z|$ is finite for all finite z, and that therefore $\tan z$ can have poles only when $\cos z$ vanishes. From the second expression for $|\cos z|^2$ in (1.190), we see that this can happen only if $y = 0$ and $\cos x = 0$, i.e. at

$$z = (n + \tfrac{1}{2})\pi\,, \tag{1.191}$$

where n is an integer.

Near $z = (n + \frac{1}{2})\pi$, say $z = \zeta + (n + \frac{1}{2})\pi$, where $|\zeta|$ is small, we shall have

$$\begin{aligned}\sin z &\longrightarrow \sin(n + \tfrac{1}{2})\pi = (-1)^n\,,\\ \cos z &\longrightarrow -\sin(n + \tfrac{1}{2})\pi \, \sin\zeta \longrightarrow -(-1)^n\, \zeta\,,\end{aligned} \tag{1.192}$$

and so the pole at $z = z_n = (n + \frac{1}{2})\pi$ has residue $b_n = -1$.

We also need to examine the boundedness of $f(z) = \tan z$ on the circles R_p. These circles are most conveniently taken to go precisely half way between the poles, so we should take $R_p = p\,\pi$. Now from (1.190) we have

$$|\tan z|^2 = \frac{\sin^2 x \, \cosh^2 y + \cos^2 x \, \sinh^2 y}{\cos^2 x \, \cosh^2 y + \sin^2 x \, \sinh^2 y} = \frac{\sin^2 x + \sinh^2 y}{\cos^2 x + \sinh^2 y}\,. \tag{1.193}$$

Bearing in mind that $\sin x$ and $\cos x$ are bounded by 1, that $\cos p\pi = (-1)^p \neq 0$, and that $\sinh^2 y$ and $\cosh^2 y$ both diverge like $\frac{1}{4}e^{2|y|}$ as $|y|$ tends to infinity, we see that $|\tan z|$ is indeed bounded on the circles R_p of radius $p\,\pi$, as p tends to infinity. Thus we can now invoke our result (1.188), to deduce that

$$\tan z = -\sum_{n=-\infty}^{\infty} \Big[\frac{1}{z - (n + \frac{1}{2})\,\pi} + \frac{1}{(n + \frac{1}{2})\pi)}\Big]\,. \tag{1.194}$$

We can split the summation range into the poles at positive and at negative values of x, by using

$$\sum_{n=-\infty}^{\infty} u_n = \sum_{n=0}^{\infty} u_n + \sum_{n=0}^{\infty} u_{-n-1}\,. , \tag{1.195}$$

Thus (1.194) gives

$$\tan z = -\sum_{n=0}^{\infty}\Big[\frac{1}{z-(n+\frac{1}{2})\,\pi} + \frac{1}{(n+\frac{1}{2})\pi)}\Big] - \sum_{n=0}^{\infty}\Big[\frac{1}{z+(n+\frac{1}{2})\,\pi} - \frac{1}{(n+\frac{1}{2})\pi)}\Big] \tag{1.196}$$

which, grouping the summands together, becomes

$$\tan z = \sum_{n=0}^{\infty} \frac{2z}{(n+\frac{1}{2})^2\,\pi^2 - z^2}\,. \tag{1.197}$$

This gives our series for the function $f(z) = \tan z$. Note that it displays the expected poles at all the places where the $\cos z$ denominator vanishes, namely at $z = (m+\frac{1}{2})\,\pi$, where m is any integer.

Another application of the result (1.188) is to obtain an expansion of an *entire function* as an infinite product. Suppose $f(z)$ is entire, meaning that it is analytic everywhere except at infinity. It follows that $f'(z)$ is an analytic function too, and so the function

$$g(z) \equiv \frac{f'(z)}{f(z)} = \frac{d}{dz}\,\log f(z) \tag{1.198}$$

is meromorphic for all finite z. (Its only singularities are poles at the places where $f(z)$ vanishes, i.e. at the *zeros* of $f(z)$.)

Let us suppose that $f(z)$ has only *simple* zeros, i.e. it vanishes like $c_n\,(z-z_n)$ near the zero at $z = z_n$, and furthermore, suppose that $f(0) \ne 0$. Thus we can apply the formula (1.188) to $g(z)$, implying that

$$\frac{d}{dz}\,\log f(z) = \frac{f'(0)}{f(0)} + \sum_{n=1}^{\infty}\Big[\frac{1}{z-z_n} + \frac{1}{z_n}\Big]\,. \tag{1.199}$$

This can be integrated to give

$$\log f(z) = \log f(0) + \frac{f'(0)}{f(0}\,z + \sum_{n=1}^{\infty}\Big[\log\Big(1-\frac{z}{z_n}\Big) + \frac{z}{z_n}\Big]\,. \tag{1.200}$$

Finally, exponentiating this, we get

$$f(z) = f(0)\,e^{[f'(0)/f(0)]\,z}\,\prod_{n=1}^{\infty}\Big(1-\frac{z}{z_n}\Big)\,e^{z/z_n}\,. \tag{1.201}$$

This infinite-product expansion is valid for any entire function $f(z)$ with simple zeros at $z = z_n$, none of which is located at $z = 0$, whose logarithmic derivative f'/f is bounded on a set of circles R_p. Obviously, without too much trouble, generalisations can be obtained where some of these restrictions are removed.

Let us apply this result in an example. Consider the function $\sin z$. From (1.190) we see that it has zeros only at $y = 0$, $x = n\,\pi$. The zero at $z = 0$ is unfortunate, since in the derivation of (1.201) we required our entire function $f(z)$ to be non-zero at $z = 0$. But this is easily handled, by taking our entire function to be $f(z) = (\sin z)/z$, which tends to 1 at $z = 0$. We now have a function that satisfies all the requirements, and so from (1.201) we shall have

$$\frac{\sin z}{z} = \prod_{n=-\infty}^{\infty} \Big(1 - \frac{z}{n\,\pi}\Big)\, e^{z/(n\,\pi)}\,, \tag{1.202}$$

where the term $n = 0$ in the product is to be omitted. Combining the positive-n and negative-n terms pairwise, we therefore find that

$$\sin z = z \prod_{n=1}^{\infty} \Big[1 - \Big(\frac{z}{n\,\pi}\Big)^2\Big]\,. \tag{1.203}$$

It is manifest that this has zeros in all the right places.

1.4.3 Branch Points, and Many-valued Functions

All the functions we have considered so far have been single-valued ones; given a point z, the function $f(z)$ has a unique value. Many functions do not enjoy this property. A classic example is the function $f(z) = z^{1/2}$. This can take two possible values for each non-zero point z, for the usual reason that there is an ambiguity of sign in taking the square root. This can be made more precise here, by considering the representation of the point z as $z = r\, e^{\mathrm{i}\theta}$. Thus we shall have

$$f(z) = (r\, e^{\mathrm{i}\theta})^{\frac12} = r^{\frac12}\, e^{\frac{\mathrm{i}}{2}\,\theta}\,. \tag{1.204}$$

But we can also write $z = r\, e^{\mathrm{i}(\theta+2\pi)}$, since θ is periodic, with period 2π, on the complex plane. Now we obtain

$$f(z) = (r\, e^{\mathrm{i}\,(\theta+2\pi)})^{\frac12} = r^{\frac12}\, e^{\frac{\mathrm{i}}{2}\,\theta + \mathrm{i}\,\pi} = -r^{\frac12}\, e^{\frac{\mathrm{i}}{2}\,\theta}\,. \tag{1.205}$$

In fact, if we look at the value of $f(z) = z^{1/2}$ on the circle $z = r\, e^{\mathrm{i}\,\theta}$, taking θ from $\theta = 0$ to $\theta_0 = 2\pi - \epsilon$, where ϵ is a small positive constant, we see that

$$f(r\, e^{\mathrm{i}\,\theta}) \longrightarrow -f(r)\,, \tag{1.206}$$

as θ approaches θ_0. But since we are back essentially to where we started in the complex plane, it follows that $f(z)$ must be *discontinuous*; it undergoes a jump in its value, on completing a circuit around the origin.

Of course although in this description we seemed to attach a particular significance to the positive real axis there is not really anything especially distinguished about this line. We could just as well have re-oriented our discussion, and concluded that the jump in the value of $f(z) = z^{1/2}$ occurred on some other axis emanating from the origin. The important invariant statement is that if you trace around any closed path that encircles the origin, the value of $z^{1/2}$ will have changed, by an overall factor of (-1), on returning to the starting point. The function $f(z) = z^{1/2}$ is double-valued on the complex plane.

If we continue on and take a second trip around the closed path, we will return again with a factor of (-1) relative to the previous visitation of the starting point. So after two rotations, we are back where we started and the function $f(z) = z^{1/2}$ is back to its original value too. This is expressed mathematically by the fact that

$$f(r\,e^{\mathrm{i}\,(\theta+4\pi)}) = r^{\frac{1}{2}}\,e^{\frac{\mathrm{i}}{2}\,\theta}\,e^{2\pi\,\mathrm{i}} = r^{\frac{1}{2}}\,e^{\frac{\mathrm{i}}{2}\,\theta} = f(r\,e^{\mathrm{i}\,\theta})\,. \tag{1.207}$$

An elegant way to deal with a multi-valued function such as $f(z) = z^{1/2}$ is to consider an enlarged two-dimensional surface on which the function is defined. In the case of the double-valued function $f(z) = z^{1/2}$, we can do it as follows. Imagine taking the complex plane, and making a semi-infinite cut along the real axis, from $x = 0$ to $x = +\infty$. Now, stack a second copy of the complex plane above this one, again with a cut from $x = 0$ to $x = +\infty$. Now, identify (i.e. glue) the lower edge of the cut of the underneath complex plane with the upper edge of the cut of the complex plane that sits on top. Finally (a little trickier to imagine!), identify the lower cut edge of the complex plane on top with the upper cut edge of the complex plane that sits underneath. We have created something a bit like a multi-story car-park (with two levels, in this case). As you drive anti-clockwise around the origin, starting on the lower floor, you find, after one circuit, that you have driven up onto the upper floor. Carrying on for one more circuit, you are back on the lower floor again.[7] What has been achieved is the creation of a two-sheeted surface, called a *Riemann Surface*, on which one has to take z around the origin through a total phase of 4π before before it returns to its starting point. The function $f(z) = z^{1/2}$ is therefore single-valued on this two-sheeted surface. "Ordinary" functions, i.e. ones that were single-valued on the original complex plane, simply have the property of taking the same value on each of the two sheets, at $z = r\,e^{\mathrm{i}\,\theta}$ and $z = r\,e^{\mathrm{i}\,(\theta+2\pi)}$.

We already noted that the choice of where to run the cut was arbitrary. The important thing is that for the function $f(z) = z^{1/2}$, it must run from $z = 0$ out to $z = \infty$, along any

[7] Of course multi-story car-parks don't work quite like that in real life, owing to the need to be able to embed them in three dimensions!

arbitrarily specifiable path. It is often convenient to take this to be the cut along the real positive axis, but any other choice will do.

The reason why the origin is so important here is that it is at $z = 0$ that the actual branch point of the function $f(z) = z^{1/2}$ lies. It is easy to see this, by following the value of $f(z) = z^{1/2}$ as z is taken around various closed paths (it is simplest to choose circles) in the complex plane. One easily sees that the $f(z) \longrightarrow -f(z)$ discontinuity is encountered for any path that encloses the origin, but no discontinuity arises for any closed path that does not enclose the origin.

If one encircles the origin, one also encircles the point at infinity, so $f(z) = z^{1/2}$ also has a branch point at infinity. (Clearly $f(1/\zeta) = \zeta^{-1/2}$ is also double valued on going around $\zeta = 0$.) So in fact, the *branch cut* that we must introduce is running from one branch point to the other. This is a general feature of multi-valued functions. In more complicated cases, this can mean that there are various possible choices for how to select the branch cuts. In the present case, choosing the branch cut along any arbitrary path from $z = 0$ to $z = \infty$ will do. Then, as one follows around a closed path, there is a discontinuity in $f(z)$ each time the branch cut is crossed. If a closed path crosses it twice (in opposite directions), then the two cancel out, and the function returns to its original value without any discontinuity.[8]

Consider another example, namely the function

$$f(z) = (z^2 - 1)^{\frac{1}{2}} = (z - 1)^{\frac{1}{2}}\,(z + 1)^{\frac{1}{2}}\,. \tag{1.208}$$

It is easy to see that since $z^{1/2}$ has a branch point at $z = 0$, here we shall have branch points at $z = 1$ and $z = -1$. Any closed path encircling either $z = -1$ or $z = +1$ (but not both) will reveal a discontinuity associated with the two-valuedness of $(z + 1)^{\frac{1}{2}}$ or $(z - 1)^{\frac{1}{2}}$ respectively. On the other hand, a circuit that encloses *both* of the points $z = 1$ and $z = -1$ will not encounter any discontinuity. The minus sign coming from encircling one branch point is cancelled by that coming from encircling the other. The upshot is that we can choose our branch cuts in either of two superficially-different ways. One of the choices is to run the branch cut from $z = -1$ to $z = +1$. Another quite different-looking choice is to

[8]In the special case of $z^{1/2}$, for which the function is exactly two-valued, then crossing over the cut twice even both in the *same* direction will cause a cancellation of the discontinuity. But more generally, a double crossing of the branch will cause the discontinuities to cancel only if the crossings are in *opposite* directions. Of course *multiple* crossings of the cut in the same direction might lead to a cancellation, if the function is only finitely-many valued. For example, $f(z) = z^{1/n}$ is n-valued, so winding n times around in the same direction gets back to the original value, if n is an integer. On the other hand $f(z) = z^{1/\pi}$ will never return to its original value, no matter how many complete circuits of the origin are made.

run a branch cut from $z = 1$ to $z = +\infty$ along the real positive axis, and another cut from $z = -1$ to $z = -\infty$ along the real negative axis.

For either of these choices, one gets the right conclusion. Namely, as one follows along any path, there is a discontinuity whenever a branch cut is crossed. Crossing twice in a given path will cause the two discontinuities to cancel out. so even if we consider the second choice of branch cuts, with two cuts running out to infinity from the points $z = -1$ and $z = +1$, we get the correct conclusion that a closed path that encircles both $z = -1$ and $z = +1$ will reveal no discontinuity after returning to its starting point.

Actually the two apparently-different choices for the branch cuts are not so very different, topologically-speaking. Really, $z = \infty$ is like a single point, and one effectively should view the complex plane as the surface of a sphere, with everywhere out at infinity corresponding to the same point on the sphere. Think of making a stereographic projection from the north pole of the sphere onto the infinite plane tangent to the south pole. We think of this plane as the complex plane. A straight line joining the north pole to a given point in the complex plane therefore passes through a single point on the sphere. This gives a mapping from each point in the complex plane into a point on the sphere. Clearly, things get a bit degenerate as we go further and further out in the complex plane. Eventually, we find that all points at $|z| = \infty$, regardless of their direction out from the origin, map onto a single point on the sphere, namely the north pole. This sphere, known as the *Riemann Sphere*, is really what the complex plane is like. Of course as we have seen, a lot of otherwise well-behaved functions tend to have more severe singularities at $z = \infty$, but that doesn't detract from the usefulness of the picture. Figure 2 below show the mapping of a point Q in the complex plane into a corresponding point P on the Riemann sphere.

As it happens, our function $f(z) = (z^2 - 1)^{1/2}$ is rather moderately behaved at $z = \infty$; it has a Laurent expansion with just a simple pole:

$$\begin{aligned} f(1/\zeta) &= (\zeta^{-2} - 1)^{\frac{1}{2}} = \zeta^{-1}\,(1 - \zeta^2)^{\frac{1}{2}}\,, \\ &= \frac{1}{\zeta} - \tfrac{1}{2}\zeta - \tfrac{1}{8}\zeta^3 - \tfrac{1}{16}\,\zeta^5 + \cdots\,. \end{aligned} \tag{1.209}$$

Since it has no branch point there, we can actually take the second choice of branch cuts, where the two cuts ran from $z = -1$ and $z = +1$ to infinity (in other words a single line from $z = -1$ to the north pole and back to $z = +1$), and deform it continuously into the first choice, where the branch cut simply runs from $z = -1$ to $z = +1$. If you think of the branch cut as an elastic band joining $z = -1$ to $z = +1$ via the north pole, it only takes

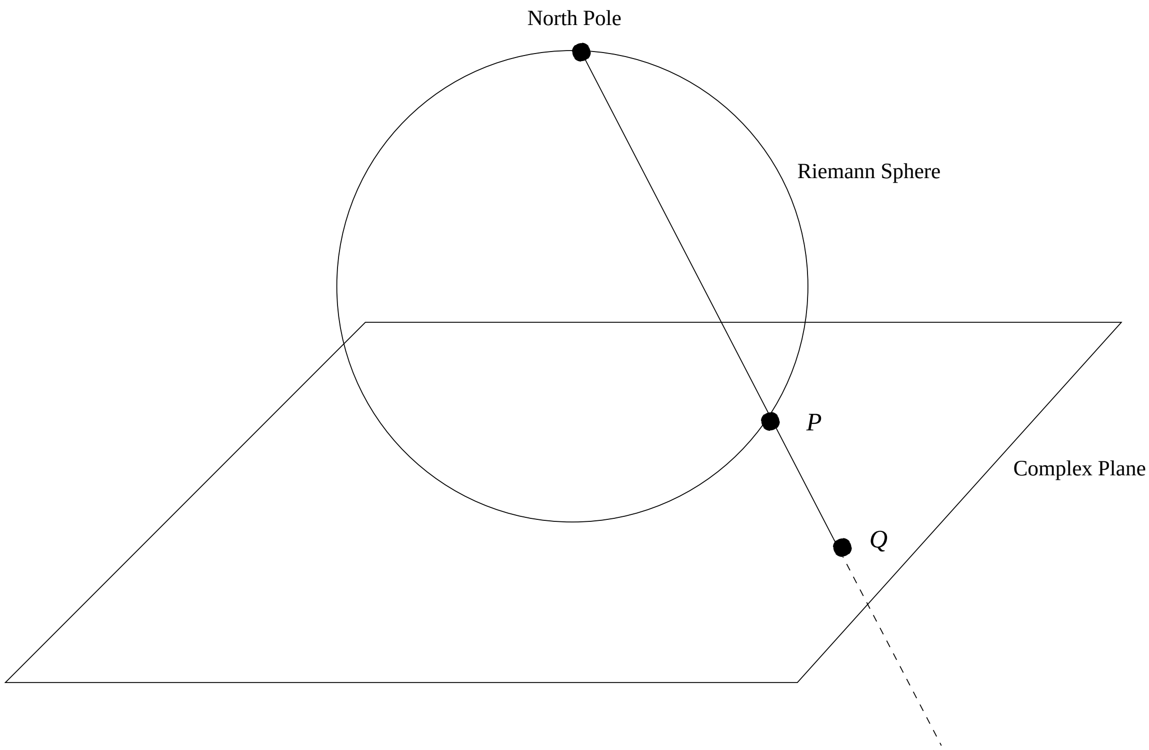

Figure 2: The point Q in the complex plane projects onto P on the Riemann sphere.

someone like Superman wandering around at the north pole to give it a little tweak, and it can contract smoothly and continuously from the second choice of branch cut to the first.

1.5 The Oppenheim Formula

Before proceeding with the mainstream of the development, let us pause for an interlude on a rather elegant and curious topic. It is a rather little-known method for solving the following problem. Suppose we are given the real part $u(x, y)$ of an analytic function $f(z) = u(x, y) + \mathrm{i}\, v(x, y)$. It is a classic exercise, to work out the imaginary part $v(x, y)$, and hence to learn what the full analytic function $f(z)$ is, by making use of the Cauchy-Riemann equations. We discussed this problem a while back, showing how one solves for v by integrating the Cauchy-Riemann equations.

What is not so well known is that one can do the job of finding $v(x, y)$ from $u(x, y)$ without ever needing to differentiate or integrate at all. This makes a nice party trick, if you go to the right (or maybe wrong!) sort of parties. The way it works is absurdly simple, and so, in the best traditions of a conjuring trick, here first is the "show." Unlike the conjuror's trick, however, we shall see afterwards how the rabbit was slipped into the hat. I have not been able to find very full references to it; the earliest I came across is to a certain Prof. A. Oppenheim, so I shall refer to it as the "Oppenheim Method."

The way to derive the analytic function $f(z)$, given its real part $u(x,y)$, is the following:

$$f(z) = 2u(\frac{z}{2}, \frac{z}{2\,\mathrm{i}}) + c\,, \tag{1.210}$$

where c is a constant. The real part of c can be fixed by using the known given expression for the real part of $f(z)$. The imaginary part of c is not determinable. Of course this is always the case; $f(z)$ and $f(z) + \mathrm{i}\,\gamma$, where γ is a real constant, have the same real parts and the same analyticity properties, so *no* method could tell us what γ is, in the absence of further specification. (In the usual Cauchy-Riemann derivation of $v(x,y)$, this arbitrariness arose as a constant of integration.)

Just to show that it really does work, consider the same example that we treated above using the traditional method. Suppose we are given that $u(x,y) = e^x\,\cos y$ is the real part of an analytic function $f(z)$. What is $f(z)$? According to (1.210), the answer is

$$\begin{aligned} f(z) &= 2e^{\frac{1}{2}z}\,\cos(-\frac{\mathrm{i}}{2}\,z) + c = 2e^{\frac{1}{2}z}\,\cosh(\frac{1}{2}z) + c\,, \\ &= e^z + 1 + c\,. \end{aligned} \tag{1.211}$$

Now, we fix c by noting, for example, that from the original $u(x,y)$ we have $u(0,0) = 1$, and so we should choose c so that $f(z)$ has real part 1 at $z = 0$. Thus we have $c = -1$, and hence $f(z) = e^z$. (There is no need to be tedious about always adding $\mathrm{i}\,\gamma$, since this trivial point about the arbitrariness over the imaginary constant is now well understood.) Finally, we can easily verify that indeed $f(z) = e^z$ is the answer we were looking for, since

$$e^z = e^{x+\mathrm{i}\,y} = e^x\,(\cos y + \mathrm{i}\,\sin y)\,, \tag{1.212}$$

and so sure enough, this analytic function has real part $e^x\,\cos y$.

How did it work? Like all the best conjuring tricks, the explanation is ludicrously simple. Since $f(z)$ is analytic, we can expand it as a power series, $f(z) = \sum_{n\ge 0} a_n\, z^n$. Note that we are assuming here that it is in particular analytic at $z = 0$; we shall show later how to remove this assumption. If we write the expansion coefficients a_n as $a_n = \alpha_n + \mathrm{i}\,\beta_n$, where α_n and β_n are real, then from the series expansion we shall have

$$2u(x,y) = f(z) + \overline{f(z)} = \sum_{n\ge 0}\Big[(\alpha_n + \mathrm{i}\,\beta_n)\,(x + \mathrm{i}\,y)^n + (\alpha_n - \mathrm{i}\,\beta_n)\,(x - \mathrm{i}\,y)^n\Big]\,. \tag{1.213}$$

Now plug in the values $x = z/2$, $y = z/(2\mathrm{i})$, as required in the Oppenheim formula:

$$2u(\frac{z}{2}, \frac{z}{2\mathrm{i}}) \;=\; \sum_{n\ge 0}\Big[(\alpha_n + \mathrm{i}\,\beta_n)\,\Big(\frac{z}{2} + \frac{z}{2}\Big)^n + (\alpha_n - \mathrm{i}\,\beta_n)\,\Big(\frac{z}{2} - \frac{z}{2}\Big)^n\Big]\,,$$

$$
\begin{aligned}
&= \sum_{n\ge 0} (\alpha_n + \mathrm{i}\,\beta_n)\, z^n + \alpha_0 - \mathrm{i}\,\beta_0\,, && (1.214)\\
&= f(z) + \alpha_0 - \mathrm{i}\,\beta_0\,.
\end{aligned}
$$

That's all there is to it! The result is proven. *Omne ignotum pro magnifico.*

We assumed in the proof that $f(z)$ was analytic at $z = 0$. If it's not, then in its present form the procedure can sometimes break down. For example, suppose we consider the function $u(x, y) = \frac{1}{2}\log(x^2 + y^2)$. (Secretly, we know that this is the real part of the function $f(z) = \log z$, which of course is analytic for all finite z except for the branch point at $z = 0$.) Trying the Oppenheim formula (1.210), we get

$$
f(z) = \log(\frac{1}{4}z^2 - \frac{1}{4}z^2) + c = \log 0 + c\,. \tag{1.215}
$$

Oooppps!! Not to worry, we know why it has failed. We need to find a generalisation of the Oppenheim formula, to allow for such cases where the function we are looking for happens to be non-analytic at $z = 0$. The answer is the following:

$$
f(z) = 2u(\frac{z+a}{2}, \frac{z-a}{2\mathrm{i}}) + c\,, \tag{1.216}
$$

where a is an arbitrary constant, to be chosen to avoid any unpleasantness. Let's try this in our function $u(x, y) = \frac{1}{2}\log(x^2 + y^2)$:

$$
\begin{aligned}
f(z) &= \log\Big[\Big(\frac{z+a}{2}\Big)^2 - \Big(\frac{z-a}{2}\Big)^2\Big] + c\,,\\
&= \log(a\, z) + c = \log z + \log a + c\,. && (1.217)
\end{aligned}
$$

So for any value of a other than $a = 0$, everything is fine. As usual, an elementary examination of a special case fixes the real part of the constant c, to give $c = -\log a$.

It is easy to see why the generalisation (1.216) works. We just repeat the derivation in (1.214), but now consider an expansion of the function $f(z)$ around $z = a$ rather than $z = 0$; $f(z) = \sum_{n\ge 0} a_n\, (z-a)^n$. Provided we don't choose a so that we are trying to expand around a singular point of $f(z)$, all must then be well:

$$
\begin{aligned}
2u(\frac{z+a}{2}, \frac{z-a}{2\mathrm{i}}) &= \sum_{n\ge 0}\Big[(\alpha_n + \mathrm{i}\,\beta_n)\Big(\frac{z+a}{2} + \frac{z-a}{2} - a\Big)^n + (\alpha_n - \mathrm{i}\,\beta_n)\Big(\frac{z+a}{2} - \frac{z-a}{2} - a\Big)^n\Big]\,,\\
&= \sum_{n\ge 0} (\alpha_n + \mathrm{i}\,\beta_n)\,(z-a)^n + \alpha_0 - \mathrm{i}\,\beta_0\,, && (1.218)\\
&= f(z) + \alpha_0 - \mathrm{i}\,\beta_0\,.
\end{aligned}
$$

Just to show off the method in one further example, suppose we are given

$$
u(x, y) = e^{\frac{x}{x^2+y^2}}\, \cos\frac{y}{x^2+y^2}\,. \tag{1.219}
$$

Obviously we shall have to use (1.216) with $a \neq 0$ here. Thus we get

$$
\begin{aligned}
f(z) &= 2e^{\frac{z+a}{2a\,z}}\,\cos\frac{z-a}{2\mathrm{i}\,a\,z} + c = 2e^{\frac{z+a}{2a\,z}}\,\cosh\frac{z-a}{2a\,z}\,, \\
&= e^{\frac{z+a}{2a\,z}}\left(e^{\frac{z-a}{2a\,z}} + e^{\frac{a-z}{2a\,z}}\right) + c\,, \\
&= e^{\frac{1}{a}} + e^{\frac{1}{z}} + c\,.
\end{aligned} \tag{1.220}
$$

Fixing the constant c from a special case, we get

$$
f(z) = e^{\frac{1}{z}}\,. \tag{1.221}
$$

The method has even worked for a function with an essential singularity, provided that we take care not to try using $a = 0$. (Try doing the calculation by the traditional procedure using (1.77) to see how much simpler it is to use the generalised Oppenheim formula.)

Having shown how effective the Oppenheim method is, it is perhaps now time to admit to why in some sense a little bit of a cheat is being played here. This is not to say that anything was incorrect; all the formulae derived are perfectly valid. It is a slightly unusual kind of trick that has been played, in fact.

Normally, when a conjuror performs a trick, it is he who "slips the rabbit into the hat," and then pulls it out at the appropriate moment to astound his audience. Ironically enough, in the case of the Oppenheim formula it is the audience itself that unwittingly slips the rabbit into the hat, and yet nevertheless it is duly amazed when the rabbit reappears.

The key point is that if one were actually working with a realistic problem, in which only the real part of an analytic function were known, one would typically be restricted to knowing it only as an "experimental result" from a set of observations. Indeed, in a common circumstance such information about the real part of an analytic function might arise precisely from an experimental observation of, for example, the refractive index of a medium as a function of frequency. The *imaginary part*, on the other hand, is related to the decay of the wave as it moves through the medium. There are quite profound *Dispersion Relations* that can be derived that relate the imaginary part to the real part. They are derived precisely by making use of the Cauchy-Riemann relations, to derive $v(x,y)$ from $u(x,y)$ by taking the appropriate derivatives and integrals of $u(x,y)$, as in (1.77).

So why was the Oppenheim formula a cheat? The answer is that it assumes that one knows what happens if one inserts complex values like $x = (z+a)/2$ and $y = (z-a)/(2\mathrm{i})$ into the "slots" of $u(x,y)$ that are designed to take the real numbers x and y. In a real-life experiment one cannot do this; one cannot set the frequency of a laser to a complex value! So the knowledge about the function $u(x,y)$ that the Oppenheim formula requires one to have is knowledge that is not available in practical situations. In those real-life cases, one would instead have to use (1.77) to calculate $v(x,y)$. And the process of integration is "non-local," in the sense that the value for the integral depends upon the values that the integrand takes in an entire region in the (x,y) plane. This is why dispersion relations actually contain quite subtle information.

The ironic thing is that although the Opennheim formula is therefore in some sense a "cheat," it nevertheless works, and works correctly, in any example that one is likely to check it with. The point is that when we want to test a formula like that, we tend not to go out and start measuring refractive indices; rather, we reach into our memories and drag out some familiar function whose properties have already been established. So it is a formula that is "almost never" usable, and yet it works "almost always" when it is tested with toy examples. It is a bit like asking someone to pick a random number. Amongst the set of all numbers, the chance that an arbitrarily chosen number will be rational is zero, and yet the chance that the person's chosen number will be rational is pretty close to unity.

1.6 Calculus of Residues

After some rather lengthy preliminaries, we have now established the groundwork for the further development of the subject of complex integration. First, we shall derive a general result about the integration of functions with poles.

If $f(z)$ has an isolated pole of order n at $z=a$, then by definition, it can be expressed as

$$f(z) = \frac{a_{-n}}{(z-a)^n} + \frac{a_{-n+1}}{(z-a)^{n-1}} + \cdots + \frac{a_{-1}}{z-a} + \phi(z)\,, \tag{1.222}$$

where $\phi(z)$ is analytic at and near $z=a$. The coefficient of a_{-1} in this expansion is called the *residue* of $f(z)$ at the pole at $z=a$.

Let us consider the integral of $f(z)$ around a closed contour C which encloses the pole at $z=a$, but within which $\phi(z)$ is analytic. (So C encloses no other singularities of $f(z)$ except the pole at $z=a$.) We have

$$\oint_C f(z)\,dz = \sum_{k=1}^n a_{-k} \oint_C \frac{dz}{(z-a)^k} + \oint \phi(z)\,dz\,. \tag{1.223}$$

By Cauchy's theorem we know that the last integral vanishes, since $\phi(z)$ is analytic within C. To evaluate the integrals under the summation, we may deform the contour C to a circle of radius ρ centred on $z=a$, respecting the previous condition that no other singularities of $f(z)$ shall be encompassed.

Letting $z - a = \rho\, e^{\mathrm{i}\theta}$, the deformed contour C is then parameterised by allowing θ to range from 0 to 2π, while holding ρ fixed. Thus we shall have $dz = \mathrm{i}\,\rho\, e^{\mathrm{i}\,\theta}\, d\theta$ on the contour, and so

$$\oint_C \frac{dz}{(z-a)^k} = \int_0^{2\pi} \frac{\mathrm{i}\,\rho\, e^{\mathrm{i}\theta}\, d\theta}{\rho^k e^{\mathrm{i}\,k\,\theta}} = \mathrm{i}\,\rho^{1-k} \int_0^{2\pi} e^{(1-k)\,\mathrm{i}\,\theta}\, d\theta = \rho^{1-k} \Big[\frac{e^{(1-k)\,\mathrm{i}\,\theta}}{1-k}\Big]_0^{2\pi}\,. \tag{1.224}$$

When the integer k takes any value other than $k = 1$, this clearly gives zero. On the other hand, when $k = 1$ we have

$$\oint_C \frac{dz}{z-a} = \mathrm{i} \int_0^{2\pi} d\theta = 2\pi\,\mathrm{i} \tag{1.225}$$

as we saw when deriving Cauchy's integral formula. Thus we arrive at the conclusion that

$$\oint_C f(z)\, dz = 2\pi\,\mathrm{i}\, a_{-1}\,. \tag{1.226}$$

The result (1.226) gives the value of the integral when the contour C encloses only the pole in $f(z)$ located at $z = a$. Clearly, if the contour were to enclose several poles, at locations $z = a$, $z = b$, $z = c$, *etc.*, we could smoothly deform C so that it described circles around each of the poles, joined by narrow "causeways" of the kind that we encountered previously, which would contribute nothing to the total integral.

Thus we arrive at the *Theorem of Residues*, which asserts that if $f(z)$ be analytic everywhere within a contour C, except at a number of isolated poles inside the contour, then

$$\oint_C F(z)\, dz = 2\pi\,\mathrm{i} \sum_s \mathcal{R}_s\,, \tag{1.227}$$

where $\mathcal{R}_s$ denotes the residue at pole number s.

It is useful to note that if $f(z)$ has a simple pole at $z = a$, then the residue at $z = a$ is given by taking the limit of $(z - a)\, f(z)$ as z tends to a.

1.7 Evaluation of real integrals

The theorem of residues can be used in order to evaluate many kinds of integrals. Since this is an important application, we shall look at a number of examples. First, a list of three main types of real integral that we shall be able to evaluate:

(1) Integrals of the form

$$\int_0^{2\pi} R(\cos\theta, \sin\theta)\, d\theta\,, \tag{1.228}$$

where R is a rational function of $\cos\theta$ and $\sin\theta$. (Recall that if $f(z)$ is a *rational* function, it means that it is the ratio of two polynomials.)

(2) Integrals of the form

$$\int_{-\infty}^{\infty} f(x)\, dx\,, \tag{1.229}$$

where $f(z)$ is analytic in the upper half plane ($y > 0$) except for poles that do not lie on the real axis. The function $f(z)$ is also required to have the property that $z\, f(z)$ should tend to zero as $|z|$ tends to infinity whenever $0 \le \arg(z) \le \pi$. ($\arg(z)$ is the *phase* of z. This condition means that $z\, f(z)$ must go to zero for all points z that go to infinity in the upper half plane.)

(3) Integrals of the form

$$\int_0^{\infty} x^{\alpha-1}\, f(x)\, dx\,, \tag{1.230}$$

where $f(z)$ is a rational function, analytic at $z = 0$, with no poles on the positive real axis. Furthermore, $z^\alpha\, f(z)$ should tend to zero as z approaches 0 or infinity.

First, consider the type (1) integrals. We introduce z as the complex variable $z = e^{\mathrm{i}\,\theta}$. Thus we have

$$\cos\theta = \tfrac{1}{2}(z + z^{-1})\,, \qquad \sin\theta = \tfrac{1}{2\mathrm{i}}(z - z^{-1})\,. \tag{1.231}$$

Recalling that R is a rational function of $\cos\theta$ and $\sin\theta$, it follows that the integral (1.228) will become a contour integral of some rational function of z, integrated around a unit circle centred on the origin. It is a straightforward procedure to evaluate the residues of the poles in the rational function, and so, by using the theorem of residues, the result follows.

Let us consider an example. Suppose we wish to evaluate

$$I(p) \equiv \int_0^{2\pi} \frac{d\theta}{1 - 2p\, \cos\theta + p^2}\,, \tag{1.232}$$

where $0 < p < 1$. Writing $z = e^{\mathrm{i}\,\theta}$, we shall have $d\theta = -\mathrm{i}\, z^{-1}\, dz$, and hence

$$I(p) = \oint_C \frac{dz}{\mathrm{i}\,(1 - p\, z)(z - p)}\,. \tag{1.233}$$

This has simple poles, at $z = p$ and $z = 1/p$. Since we are assuming that $0 < p < 1$, it follows from the fact that C is the unit circle that the pole at $z = 1/p$ lies outside C, and so the only pole enclosed is the simple pole at $z = p$. Thus the residue of the integrand at $z = p$ is given by taking the limit of

$$(z - p)\Big[\frac{1}{\mathrm{i}\,(1 - p\, z)(z - p)}\Big] \tag{1.234}$$

as z tends to p, i.e. $-\mathrm{i}/(1-p^2)$. Thus from the theorem of residues (1.227), we get

$$\int_0^{2\pi} \frac{d\theta}{1-2p\,\cos\theta+p^2} = \frac{2\pi}{1-p^2}\,, \qquad 0<p<1\,. \tag{1.235}$$

Note that if we consider the same integral (1.232), but now take the constant p to be greater than 1, the contour C (the unit circle) now encloses only the simple pole at $z=1/p$. Multiplying the integrand by $(z-1/p)$, and taking the limit where z tends to $1/p$, we now find that the residue is $+\mathrm{i}/(1-p^2)$, whence

$$\int_0^{2\pi} \frac{d\theta}{1-2p\,\cos\theta+p^2} = \frac{2\pi}{p^2-1}\,, \qquad p>1\,. \tag{1.236}$$

In fact the results for all real p can be combined into the single formula

$$\int_0^{2\pi} \frac{d\theta}{1-2p\,\cos\theta+p^2} = \frac{2\pi}{|p^2-1|}\,. \tag{1.237}$$

For a more complicated example, consider

$$I(p) \equiv \int_0^{2\pi} \frac{\cos^2 3\theta\, d\theta}{1-2p\,\cos 2\theta+p^2}\,, \tag{1.238}$$

with $0<p<1$. Now, we have

$$I(p) = \oint_C \frac{\left(\frac{1}{2}z^3+\frac{1}{2}z^{-3}\right)^2 dz}{\mathrm{i}\,z\,(1-p\,z^2)(1-p\,z^{-2})} = \oint_C \frac{(z^6+1)^2\,dz}{4\mathrm{i}\,z^5\,(1-p\,z^2)(z^2-p)}\,. \tag{1.239}$$

The integrand has poles at $z=0$, $z=\pm p^{\frac{1}{2}}$ and $z=\pm p^{-\frac{1}{2}}$. Since we are assuming $0<p<1$, it follows that only the poles at $z=0$ and $z=\pm p^{\frac{1}{2}}$ lie within the unit circle corresponding to the contour C. The poles at $z=\pm p^{\frac{1}{2}}$ are simple poles, and the only slight complication in this example is that the pole at $z=0$ is of order 5, so we have to work a little harder to extract the residue there. Shortly, we shall derive a general formula that can sometimes be useful in such circumstances. An alternative approach, often in practise preferrable when one is working out the algebra by hand (as opposed to using an algebraic computing program), is simply to factor out the singular behaviour and then make a Taylor expansion of the remaining regular function, as we described earlier. In this case, it is quite simple. We have

$$\begin{aligned}
\frac{(z^6+1)^2}{z^5\,(1-p\,z^2)(z^2-p)} &= -\frac{1}{p\,z^5}\,(1+z^6)^2\,(1-p\,z^2)^{-1}\,(1-z^2\,p^{-1})^{-1} \\
&= -\frac{1}{p\,z^5}\,(1+p\,z^2+p^2\,z^4+\cdots)(1+z^2\,p^{-1}+z^4\,p^{-2}+\cdots) \\
&= -\frac{1}{p\,z^5}\,(1+p\,z^2+p^2\,z^4+z^2\,p^{-1}+z^4+z^4\,p^{-2}+\cdots) \\
&= -\frac{1}{p\,z^5} - \frac{(p^2+1)}{p^2\,z^3} - \frac{(p^4+p^2+1)}{p^3\,z} + \cdots\,,
\end{aligned} \tag{1.240}$$

from which we read off the residue of this function at $z = 0$ as the coefficient of $1/z$. Notice that to make these expansions we just used $(1-x)^{-1} = 1 + x + x^2 + \cdots$, and that we only needed to push these expansions far enough to collect the terms at order z^4 that are then multiplied by $1/z^5$.

After a little further algebra for the two simple poles, we find that the residues of the integrand in (1.239) at $z = 0$, $z = p^{\frac{1}{2}}$ and $z = -p^{\frac{1}{2}}$ are given by

$$\frac{\mathrm{i}\,(1+p^2+p^4)}{4p^3}\,, \qquad -\frac{\mathrm{i}\,(1+p^3)^2}{8p^3\,(1-p^2)}\,, \qquad -\frac{\mathrm{i}\,(1+p^3)^2}{8p^3\,(1-p^2)}\,, \tag{1.241}$$

respectively. Plugging into the theorem of residues (1.227), we therefore obtain the result

$$\int_0^{2\pi} \frac{\cos^2 3\theta\, d\theta}{1 - 2p\,\cos 2\theta + p^2} = \frac{\pi\,(1-p+p^2)}{(1-p)}\,, \tag{1.242}$$

when $0 < p < 1$.

It is sometimes useful to have a general result for the evaluation of the residue at an n'th-order pole. It really just amounts to formalising the procedure we used above, of extracting the singular behaviour and then Taylor expanding the remaining analytic factor:

If $f(z)$ has a pole of order n at $z = a$, it follows that it will have the form

$$f(z) = \frac{g(z)}{(z-a)^n}\,, \tag{1.243}$$

where $g(z)$ is analytic in the neighbourhood of $z = a$. Thus we may expand $g(z)$ in a Taylor series around $z = a$, giving

$$\begin{aligned} f(z) &= \frac{1}{(z-a)^n}\Big(g(a) + (z-a)\,g'(a) + \cdots + \frac{1}{(n-1)!}\,(z-a)^{n-1}\,g^{(n-1)}(a) + \cdots\Big)\,,\\ &= \frac{g(a)}{(z-a)^n} + \frac{g'(a)}{(z-a)^{n-1}} + \cdots + \frac{g^{(n-1)}(a)}{(n-1)!\,(z-a)} + \cdots\,. \end{aligned} \tag{1.244}$$

We then read off the residue, namely the coefficient of the first-order pole term $1/(z-a)$, finding $g^{(n-1)}(a)/(n-1)!$. Re-expressing this in terms of the original function $f(z)$, using (1.243), we arrive at the general result that

If $f(z)$ has a pole of order n at $z = a$, then the residue $\mathcal{R}$ is given by

$$\mathcal{R} = \frac{1}{(n-1)!}\Big[\frac{d^{n-1}}{dz^{n-1}}\,\big((z-a)^n\,f(z)\big)\Big]_{z=a}\,. \tag{1.245}$$

It is straightforward to check that when applied to our example in (1.239), the formula (1.245) reproduces our previous result for the residue at the 5'th-order pole at $z = 0$. In practise, though, it is usually more convenient in a hand calculation to use the method described previously, rather than slogging out the derivatives needed for (1.245).

As a final example of the type (1) class of integrals, consider

$$I(a,b) \equiv \int_0^{2\pi} \frac{d\theta}{(a+b\,\cos\theta)^2} = \oint_C \frac{4z\,dz}{\mathrm{i}\,(b+2a\,z+b\,z^2)^2}\,, \tag{1.246}$$

where $a > b > 0$. The integrand has (double) poles at

$$z = \frac{-a \pm \sqrt{a^2-b^2}}{b}\,, \tag{1.247}$$

and so just the pole at $z = (-a+\sqrt{a^2-b^2})/b$ lies inside the unit circle. After a little calculation, one finds the residue there, and hence, from (1.227), we get

$$\int_0^{2\pi} \frac{d\theta}{(a+b\,\cos\theta)^2} = \frac{2\pi\,a}{(a^2-b^2)^{\frac{3}{2}}}\,. \tag{1.248}$$

Turning now to integrals of type 2 (1.229), the approach here is to consider a contour integral of the form

$$I \equiv \oint_C f(z)\,dz\,, \tag{1.249}$$

where the contour C is taken to consist of the line from $x=-R$ to $x=+R$ along the x axis, and then a semicircle of radius R in the upper half plane, thus altogether forming a closed path.

The condition that $z\,f(z)$ should go to zero as $|z|$ goes to infinity with $0 \le \arg(z) \le \pi$ ensures that the contribution from integrating along the semicircular arc will vanish when we send R to infinity. (On the arc we have $dz = \mathrm{i}\,R\,e^{\mathrm{i}\,\theta}\,d\theta$, and so we would like $R\,f(R\,e^{\mathrm{i}\,\theta})$ to tend to zero as R tends to infinity, for all θ in the range $0 \le \theta \le \pi$, whence the condition that we placed on $f(z)$.) Thus we shall have that

$$\int_{-\infty}^{\infty} f(x)\,dx = 2\pi\,\mathrm{i}\sum_s \mathcal{R}_s\,, \tag{1.250}$$

where the sum is taken over the residues $\mathcal{R}_s$ at all the poles of $f(z)$ in the upper half plane. The contour is depicted in Figure 3 below.

Consider, as a simple example,

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2}\,. \tag{1.251}$$

Clearly, the function $f(z) = (1+z^2)^{-1}$ fulfils all the requirements for this type of integral. Since $f(z) = (z+\mathrm{i})^{-1}\,(z-\mathrm{i})^{-1}$, we see that there is just a single pole in the upper half

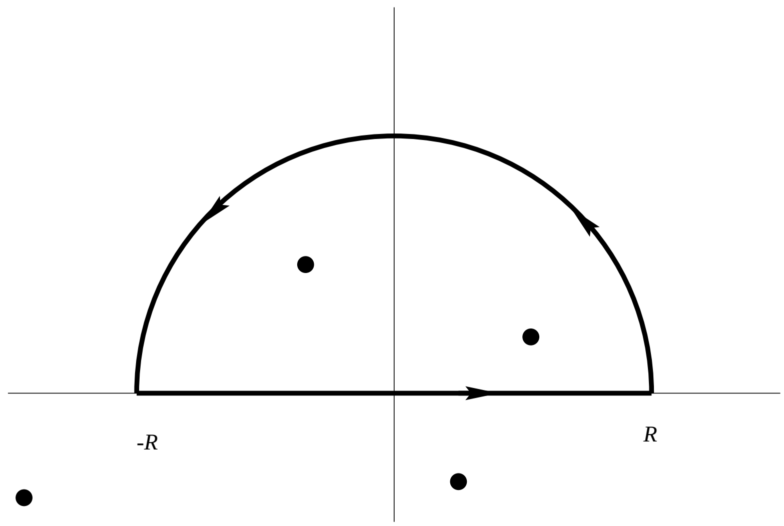

Figure 3: The contour encloses poles of $f(z)$ in the upper half plane

plane, at $z = \mathrm{i}$. It is a simple pole, and so the residue of $f(z)$ there is $1/(2\mathrm{i})$. Consequently, from (1.250) we derive

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi\,. \tag{1.252}$$

Of course in this simple example we could perfectly well have evaluated the integral instead by more "elementary" means. A substitution $x = \tan\theta$ would convert (1.251) into

$$\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} d\theta = \pi\,. \tag{1.253}$$

However, in more complicated examples the contour integral approach is often much easier to use. Consider, for instance, the integral

$$\int_{-\infty}^{\infty} \frac{x^4\,dx}{(a+b\,x^2)^4}\,, \tag{1.254}$$

where $a > 0$ and $b > 0$. The function $f(z) = z^4\,(a+b\,z^2)^{-4}$ has poles of order 4 at $z = \pm\mathrm{i}(a/b)^{\frac{1}{2}}$, and so there is just one pole in the upper half plane. Using either the formula (1.245), or the direct approach of extracting the singular factor and Taylor-expanding "by hand" to calculate the residue, and multiplying by $2\pi\,\mathrm{i}$, we get

$$\int_{-\infty}^{\infty} \frac{x^4\,dx}{(a+b\,x^2)^4} = \tfrac{1}{16}\,\pi\,a^{-\frac{3}{2}}\,b^{-\frac{5}{2}}\,. \tag{1.255}$$

Just to illustrate the point, we may note that we could in principle have worked out (1.254) by "elementary means," but the procedure would be quite unpleasant to implement. By means of an appropriate trigonometric substitution, one eventually concludes that the integral (1.254) gives

$$\int_{-\infty}^{\infty} \frac{x^4\, dx}{(a+b\,x^2)^4} = \Big[\frac{3b^2\,x^5 - 8a\,b\,x^3 - 3a^2\,x}{48a\,b^2\,(a+b\,x^2)^3} + \frac{1}{16a^{3/2}\,b^{5/2}}\ \arctan\Big(\frac{\sqrt{b}\,x}{\sqrt{a}}\Big)\Big]_{-\infty}^{\infty}\,, \qquad (1.256)$$

from which the result (1.256) follows. If you try it, you will find that the labour involved is *much* more than in the contour integral method.

One reason for the great saving of labour when using the contour integral method is that when using the old-fashioned approach of first evaluating the indefinite integral, and then substituting the limits of integration, one is actually working out much more than is ever needed. It is intrinsically a more complicated exercise to find the function whose derivative is $f(x)$ than it is to find the result of integrating $f(x)$ over a fixed interval such as $-\infty$ to ∞. If we look at the first term in (1.255) (the rational function of x), we see that they disappear altogether when one sets x to its limiting values of $-\infty$ and $+\infty$. And yet by the old-fashioned method it is necessary first to thrash out the entire integral, including these terms, since we don't know in advance how to recognise that some of the terms in the final result will end up getting thrown away when the limits are substituted. In our example above, the indefinite integral *is* still doable, albeit with a struggle. In more complicated examples there may be no closed-form expression for the indefinite integral, and yet the definite integral may have a simple form, easily found by contour-integral methods.

Finally, consider integrals of type 3 (1.230). In general, α is assumed to be a real number, but not an integer. We consider the function $(-z)^{\alpha-1}\,f(z)$, which therefore has a branch-point singularity at $z=0$. We shall choose to run the branch cut along the positive real z-axis, and consider a contour C of exactly the form given in Figure 1, with $a=0$. Note that the branch cut therefore does not lie within the total closed contour. Eventually, we allow the radius of the larger circle C_1 to become infinite, while the radius of the smaller circle C_2 will go to zero. In view of the assumption that $z^\alpha\,f(z)$ goes to zero as z goes to 0 or infinity, it follows that the contributions from integrating around these two circles will each give zero.

Unlike the situation when we used the contour of Figure 1 for deriving the Laurent series, we are now faced with a function $(-z)^{\alpha-1}\,f(z)$ with a branch point at $z=0$, and we have chosen to run the branch cut between the two paths forming the causeway. Consequently, there is a discontinuity as one traces the value of $(-z)^{\alpha-1}\,f(z)$ around a closed path that encircles the origin. This means that the results of integrating along the two sides of the "causeway" connecting the circles C_1 and C_2 will not cancel.

We shall define $(-z)^{\alpha-1}$ to be *real and positive* when z lies on the negative real axis, such as at the point where the small circle C_2 intersects the negative real axis.[9] Consequently, on the *lower* part of the causeway (below the positive real axis), the phase will be $e^{\mathrm{i}\pi\,(\alpha-1)}$. On the other hand, on the *upper* part of the causeway (above the positive real axis), the phase will be $e^{-\mathrm{i}\,\pi\,(\alpha-1)}$. Thus we find that

$$\begin{aligned}\oint_C (-z)^{\alpha-1}\,f(z)\,dz &= -e^{\mathrm{i}\pi\,(\alpha-1)}\int_0^\infty x^{\alpha-1}\,f(x)\,dx + e^{-\mathrm{i}\pi\,(\alpha-1)}\int_0^\infty x^{\alpha-1}\,f(x)\,dx\,,\\ &= 2\mathrm{i}\,\sin(\pi\,\alpha)\int_0^\infty x^{\alpha-1}\,f(x)\,dx\,,\end{aligned} \tag{1.257}$$

where the minus sign on the first term on the right in the top line comes from the fact that the integral from $x=0$ to $x=\infty$ is running in the direction opposite to the indicated direction of the contour in Figure 1. The contour integral on the left-hand side picks up all the contributions from the poles of $f(z)$. Thus we have the result that

$$\int_0^\infty x^{\alpha-1}\,f(x)\,dx = \frac{\pi}{\sin\pi\alpha}\sum_s \mathcal{R}_s\,, \tag{1.258}$$

where $\mathcal{R}_s$ is the residue of $(-z)^{\alpha-1}\,f(z)$ at pole number s of the function $f(z)$.

As an example, consider the integral

$$\int_0^\infty \frac{x^{\alpha-1}\,dx}{1+x}\,. \tag{1.259}$$

Here, we therefore have $f(z)=1/(z+1)$, which just has a simple pole, at $z=-1$. The residue of $(-z)^{\alpha-1}\,f(z)$ is therefore just 1, and so from (1.258) we obtain that when $0<\alpha<1$,

$$\int_0^\infty \frac{x^{\alpha-1}\,dx}{1+x} = \frac{\pi}{\sin\pi\alpha}\,. \tag{1.260}$$

(The restriction $0<\alpha<1$ is to ensure that the fall-off conditions for type 3 integrands at $z=0$ and $z=\infty$ are satisfied.)

A common circumstance is when there is in fact a pole in the integrand that lies exactly on the path where we wish to run the contour. An example would be an integral of the type

[9] Note that this is a *definition* – it is not automatically true. To see this, observe that we can write $1=e^{2\pi\,\mathrm{i}n}$, where n is any integer. If we were dealing with a single-value function $g(z)$, then $g(1)=g(e^{2\pi\,\mathrm{i}\,n})$ for any n and there is no ambiguity. But if we raise 1 to the fractional power b, then we would have all the possible choices $e^{2\pi\,\mathrm{i}\,n\,b}$ for what we mean by 1^b. If b is rational, i.e. $b=p/q$ for integers p and q, then there will be a finite number of different "b'th roots of 1," while if b is irrational, such as $b=\sqrt{2}$ or $b=\pi$, then there will be infinitely many different b'th roots of 1. They are in general complex numbers of unit magnitude. Our specific choice we are making here, which must be stated, and not merely left to chance, is to say that we shall define $(-z)^b$ to be real and positive, when z lies on the negative real axis.

(2) discussed above, but where the integrand now has poles on the real axis. If these are *simple* poles, then the following method can be used. Consider a situation where we wish to evaluate $\int_{-\infty}^{\infty} f(x)\,dx$, and $f(z)$ has a single simple pole on the real axis, at $z=a$. What we do is to make a little detour in the contour, to skirt around the pole, so the contour C in Figure 3 now aquires a little semicircular "bypass" γ, of radius ρ, taking it into the upper half plane around the point $z=a$. This is shown in Figure 4 below. Thus before we take the limit where $R\longrightarrow\infty$, we shall have

$$\int_{-R}^{a-\rho} f(x)\,dx + \int_\gamma f(z)\,dz + \int_{a+\rho}^{R} f(x)\,dx = 2\pi\,\mathrm{i}\sum_j \mathcal{R}_j\,, \tag{1.261}$$

where as usual $\mathcal{R}_j$ is the residue of $f(z)$ at its j'th pole in the upper half plane.

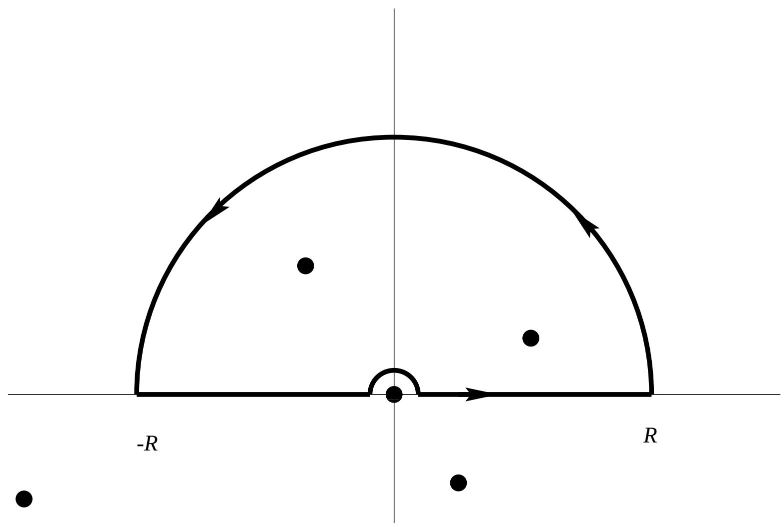

Figure 4: The contour bypasses a pole at the origin

To evaluate the contribution on the semicircular contour γ, we let $z-a=\rho\,e^{\mathrm{i}\,\theta}$, implying that the contour is parameterised (in the direction of the arrow) by taking θ to run from π to 0. Thus near $z=a$ we shall have $f(z)\sim\widetilde{\mathcal{R}}/(z-a)$, where $\widetilde{\mathcal{R}}$ is the residue of the simple pole at $z=a$, and $dz=\mathrm{i}\,\rho\,e^{\mathrm{i}\,\theta}\,d\theta$, whence

$$\int_\gamma f(z)\,dz = \mathrm{i}\,\widetilde{\mathcal{R}}\int_\pi^0 d\theta = -\mathrm{i}\,\pi\,\mathcal{R}\,. \tag{1.262}$$

Sending R to infinity, and ρ to zero, the remaining two terms on the left-hand side of (1.261) define what is called the *Cauchy Principal Value Integral*, denoted by $P\int$,

$$P\int_{-\infty}^{\infty} f(x)\,dx \equiv \int_{-\infty}^{a-\epsilon} f(x)\,dx + \int_{a+\epsilon}^{\infty} f(x)\,dx\,, \tag{1.263}$$

where one takes the limit where the small positive quantity ϵ goes to zero. Such a definition is necessary in order to give meaning to what would otherwise be an ill-defined integral.

In general, we therefore arrive at the result that if $f(z)$ has several simple poles on the real axis, with residues $\widetilde{\mathcal{R}}_k$, as well as poles in the upper half plane with residues $\mathcal{R}_j$, then

$$P\int_{-\infty}^{\infty} f(x)\,dx = 2\pi\,\mathrm{i}\sum_j \mathcal{R}_j + \mathrm{i}\,\pi\sum_k \widetilde{\mathcal{R}}_k\,. \tag{1.264}$$

Here, the principal-value prescription is used to give meaning to the integral, analogously to (1.263), at each of the simple poles on the real axis.

Consider, as an example, $\int_{-\infty}^{\infty}(\sin x)/x\,dx$. Actually, of course, this integrand has no pole on the real axis, since the pole in $1/x$ is cancelled by the zero of $\sin x$. But one way to do the calculation is to say that we shall calculate the imaginary part of

$$\int_{-\infty}^{\infty}\frac{e^{\mathrm{i}\,x}}{x}\,dx = \int_{-\infty}^{\infty}\frac{\cos x}{x}\,dx + \mathrm{i}\int_{-\infty}^{\infty}\frac{\sin x}{x}\,dx\,. \tag{1.265}$$

We must now use the principal-value prescription to give meaning to this integral, since the real part of the integrand in (1.265), namely $(\cos x)/x$, *does* have a pole at $x = 0$. But since we are after the imaginary part, the fact that we have "regulated" the real part of the integral will not upset what we want. Thus from (1.264) we find that

$$P\int_{-\infty}^{\infty}\frac{e^{\mathrm{i}\,x}}{x}\,dx = \mathrm{i}\,\pi\,, \tag{1.266}$$

and so from the imaginary part (which is all there is; the principal-value integral has regulated the ill-defined real part to be zero) we get

$$\int_{-\infty}^{\infty}\frac{\sin x}{x}\,dx = \pi\,. \tag{1.267}$$

Notice that there is another way that we could have handled a pole on the real axis. We could have bypassed aound it the other way, by taking a semicircular contour $\tilde\gamma$ that went into the lower half complex plane instead. Now, the integration (1.262) would be replaced by one where θ ran from $\theta = \pi$ to $\theta = 2\pi$ as one follows in the direction of the arrow, giving, eventually, a contribution $-\mathrm{i}\,\pi\,\widetilde{\mathcal{R}}$ rather than $+\mathrm{i}\,\pi\,\widetilde{\mathcal{R}}$ in (1.264). But all is actually well, because if we make a detour of this kind we should actually now also include the contribution of this pole as an honest pole enclosed by the full contour C, so it will also give a contribution $2\pi\,\mathrm{i}\,\widetilde{\mathcal{R}}$ in the first summation on the right-hand side of (1.264). So at the end of the day, we end up with the same conclusion no matter which way we detour around the pole.

Another common kind of real integral that can be evaluated using the calculus of residues involves the log function. Consider, for example, the following:

$$I \equiv \int_0^\infty \frac{\log x\, dx}{(1+x^2)^2}\,. \tag{1.268}$$

One way to evaluate this is by taking the usual large semicircular contour in the upper half plane, with a little semicircular detour γ (in the upper half plane) bypassing the branch point at $z=0$, as in Figure 4. We think of running the branch cut of $\log z$ from $z=0$ to $z=\infty$, just fractionally *below* the positive real axis. Thus for z on the positive real axis, we shall have simply $\log z = \log x$. If we look just *below* the branch cut, i.e. for $z = x - \mathrm{i}\,\epsilon$, where ϵ is a very small positive constant, we shall have $\log z = \log x + 2\mathrm{i}\,\pi$ in the limit when ϵ goes to zero, since we have effectively swung once around the origin, sending $z \longrightarrow z\, e^{2\mathrm{i}\,\pi}$, to get there.

Then we shall have

$$\int_{-\infty}^{-\rho} \frac{\log x\, dx}{(1+x^2)^2} + \int_\gamma \frac{\log z\, dz}{(1+z^2)^2} + \int_\rho^\infty \frac{\log x\, dx}{(1+x^2)^2} = 2\pi\,\mathrm{i}\,\mathcal{R}\,, \tag{1.269}$$

where $\mathcal{R}$ is the residue of $(\log z)/(1+z^2)^2$ at the double pole at $z=\mathrm{i}$ in the upper half plane. (As usual, we must check that the integrand indeed has the appropriate fall-off property so that the contribution from the large semicircular arc goes to zero; it does.) There are a couple of new features that this example illustrates.

First, consider the integral around the little semicircle γ. Letting $z = \rho\, e^{\mathrm{i}\,\theta}$ there we shall have

$$\int_\gamma \frac{\log z\, dz}{(1+z^2)^2} = -\mathrm{i}\,\rho \int_0^\pi \frac{\log(\rho\, e^{\mathrm{i}\,\theta})\, e^{\mathrm{i}\,\theta}\, d\theta}{(1+\rho^2\, e^{2\mathrm{i}\,\theta})^2}\,. \tag{1.270}$$

This looks alarming at first, but closer inspection reveals that it will give zero, once we take the limit $\rho \longrightarrow 0$. The point is that after writing $\log(\rho\, e^{\mathrm{i}\,\theta}) = \log\rho + \mathrm{i}\,\theta$, we see that the θ integrations will not introduce any divergences, and so the overall factors of ρ or $\rho\,\log\rho$ in the two parts of the answer will both nicely kill off the contributions, as $\rho \longrightarrow 0$.

Next, consider the first integral on the left-hand side of (1.269). For this, we can change variable from x, which takes negative values, to t, say, which is positive. But we need to take care, because of the multi-valuedness of the log function. So we should define

$$x = e^{\mathrm{i}\pi}\, t\,. \tag{1.271}$$

In all places except the log, we can simply interpret this as $x=-t$, but in the log we shall have $\log z = \log(e^{\mathrm{i}\pi}\, t) = \log t + \mathrm{i}\,\pi$. Thus the first integral in (1.269) gives

$$\int_{-\infty}^0 \frac{\log x\, dx}{(1+x^2)^2} = \int_0^\infty \frac{\log t\, dt}{(1+t^2)^2} + \mathrm{i}\,\pi \int_0^\infty \frac{dt}{(1+t^2)^2}\,. \tag{1.272}$$

(Now that we know that there is no contribution from the little semicircle γ, we can just take $\rho = 0$ and forget about it.) The first term on the right-hand side here is of exactly the same form as our original integral I defined in (1.268). The second term on the right is a simple integral. It itself can be done by contour integral methods, as we have seen. Since there is no new subtlety involved in evaluating it, let's just quote the answer, namely

$$\int_0^\infty \frac{dt}{(1+t^2)^2} = \tfrac{1}{4}\pi\,. \tag{1.273}$$

Taking stock, we have now arrived at the result that

$$2I + \tfrac{1}{4}\mathrm{i}\,\pi^2 = 2\pi\,\mathrm{i}\,\mathcal{R}\,. \tag{1.274}$$

It remains only to evaluate the residue of $(\log z)/(1+z^2)^2$ at the double pole at $z = \mathrm{i}$ in the upper half plane. We do this with the standard formula (1.245). Thus we have

$$\mathcal{R} = \frac{d}{dz}\Big[\frac{\log z}{(z+\mathrm{i})^2}\Big]\,, \tag{1.275}$$

to be evaluated at $z = \mathrm{i} = e^{\mathrm{i}\,\pi/2}$. (Note that we should write it explicitly as $e^{\mathrm{i}\,\pi/2}$ in order to know exactly what to do with the $\log z$ term.) Thus we get

$$\mathcal{R} = \tfrac{\mathrm{i}}{4} + \tfrac{1}{8}\pi\,. \tag{1.276}$$

Plugging into (1.274), we see that the imaginary term on the left-hand side is cancelled by the imaginary term in (1.276), leaving just $2I = -\pi/2$. Thus, eventually, we arrive at the result that

$$\int_0^\infty \frac{\log x\,dx}{(1+x^2)^2} = -\tfrac{1}{4}\pi\,. \tag{1.277}$$

Aside from the specifics of this example, there are two main general lessons to be learned from it. The first is that if an integrand has just a logarithmic divergence at some point $z = a$, then the contour integral around a little semicircle or circle centred on $z = a$ will give zero in the limit when its radius ρ goes to zero. This is because the logarithmic divergence of $\log\rho$ is outweighed by the linear factor of ρ coming from writing $dz = \mathrm{i}\,\rho\, e^{\mathrm{i}\,\theta}\,d\theta$.

The second general lesson from this example is that one should pay careful attention to how the a coordinate redefinition is performed, for example when re-expressing an integral along the negative real axis as an integral over a positive variable (like t in our example). In particular, one has to handle the redefinition with appropriate care in the multi-valued log function.

1.8 Summation of Series

Another application of the calculus of residues is for evaluating certain types of infinite series. The idea is the following. We have seen that the functions $\operatorname{cosec}\pi z$ and $\cot\pi z$ have the property of having simple poles at all the integers, whilst otherwise being analytic in the whole finite complex plane. In fact, they are bounded everywhere as one takes $|z|$ to infinity, except along the real axis where the poles lie. Using these functions, we can write down contour integrals that are related to infinite sums.

First, let us note that the residues of the two trigonometric functions are as follows:

- $\pi\,\cot\pi z$ has residue 1 at $z = n$
- $\pi\operatorname{cosec}\pi z$ has residue $(-1)^n$ at $z = n$

Consider the following integral:

$$I_p \equiv \oint_{C_p} f(z)\,\pi\,\cot\pi z\,, \tag{1.278}$$

where C_p is a closed contour that encloses the poles of $\cot\pi z$ at $z = 0, \pm 1, \pm 2, \ldots, \pm p$, but does not enclose any that lie at any larger value of $|z|$. A typical choice for the contour C_p is a square, centred on the origin, with side $2p+1$, or else a circle, again passing through the points $\pm(p+\frac{1}{2})$. (See Figure 5 below.) Then by the theorem of residues we shall have

$$I_p = 2\pi\,\mathrm{i}\sum_{n=-p}^{p} f(n) + 2\pi\,\mathrm{i}\sum_a \mathcal{R}_a\,, \tag{1.279}$$

where $\mathcal{R}_a$ denotes the residue of $f(z)\,\pi\,\cot\pi z$ at pole number a of the function $f(z)$, and the summation is over all such poles that lie within the contour C_p. In other words, we have simply split the total sum over residues into the first term, which sums over the residues at the known simple poles of $\cot\pi z$, and the second term, which sums over the poles associated with the function $f(z)$ itself. Of course, in the first summation, the residue of $f(z)\,\pi\,\cot\pi z$ at $z = n$ is simply $f(n)$, since the pole in $\pi\,\cot\pi z$ is simple, and itself has residue 1. (We are assuming here that $f(z)$ doesn't itself have poles at the integers.)

Now, it is clear that if we send p to infinity, so that the corresponding contour C_p grows to infinite size and encompasses the whole complex plane, we shall have

$$\oint_{C_\infty} f(z)\,\pi\,\cot\pi z = 2\pi\,\mathrm{i}\sum_{n=-\infty}^{\infty} f(n) + 2\pi\,\mathrm{i}\sum_a \mathcal{R}_a\,, \tag{1.280}$$

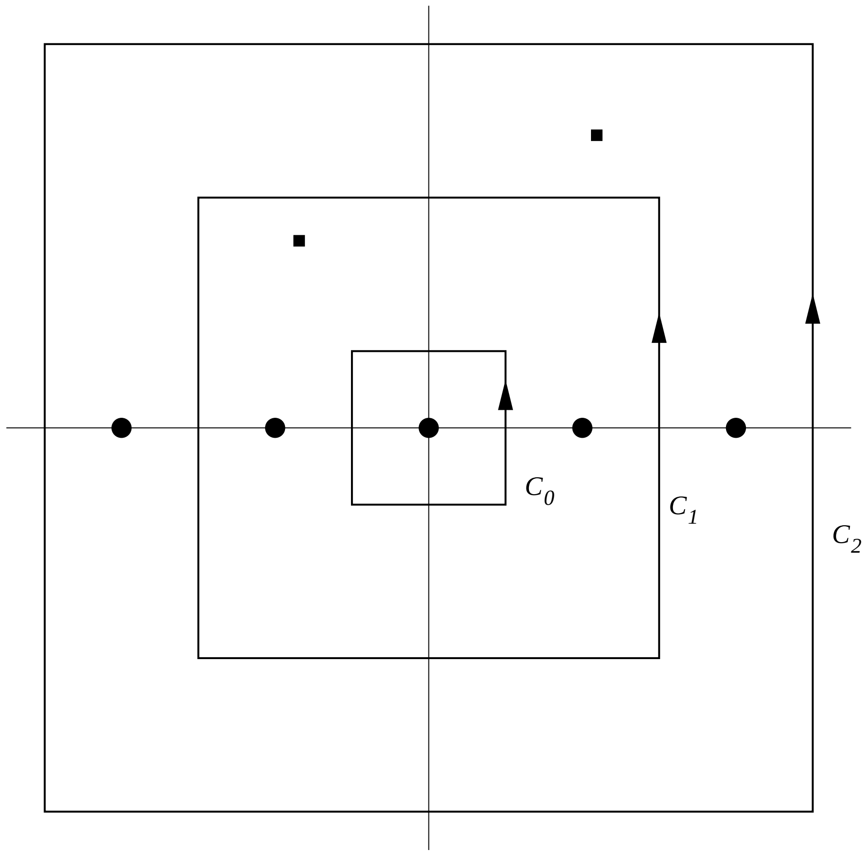

Figure 5: The square contours enclose the poles of $f(z)$ (square dots) and the poles of $\cot \pi z$ or $\operatorname{cosec} \pi z$ (round dots)

where the second sum now ranges over the residues $\mathcal{R}_a$ of $f(z)\, \pi\, \cot \pi z$ at all the poles of $f(z)$. Furthermore, let us suppose that the function $f(z)$ is such that

$$|z\, f(z)| \longrightarrow 0 \quad \text{as} \quad |z| \longrightarrow \infty\,. \tag{1.281}$$

It follows that the integral around the contour C_∞ out at infinity will be zero. Consequently, we obtain the result that

$$\sum_{n=-\infty}^{\infty} f(n) = -\sum_a \mathcal{R}_a\,, \tag{1.282}$$

where the right-hand sum is over the residues $\mathcal{R}_a$ of $f(z)\, \pi\, \cot \pi z$ at all the poles of $f(z)$.

In a similar fashion, using $\operatorname{cosec} \pi z$ in place of $\cot \pi z$, we have that

$$\sum_{n=-\infty}^{\infty} (-1)^n\, f(n) = -\sum_a \widetilde{\mathcal{R}}_a\,, \tag{1.283}$$

where the right-hand sum is over the residues of $f(z)\, \pi \operatorname{cosec} \pi z$ at all the poles of $f(z)$.

Consider an example. Suppose we take

$$f(z) = \frac{1}{(z+a)^2}\,. \tag{1.284}$$

This has a double pole at $z = -a$. Using (1.245), we therefore find that the residue of $f(z)\,\pi\,\cot\pi z$ at $z = -a$ is

$$\mathcal{R} = -\pi^2 \operatorname{cosec}^2(\pi a)\,, \tag{1.285}$$

and hence from (1.282) we conclude that

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+a)^2} = \frac{\pi^2}{\sin^2 \pi a}\,. \tag{1.286}$$

We can also evaluate the analgous sum with alternating signs, by using (1.283) instead. Now, we caluate the residue of $(z+a)^{-2}\,\pi \operatorname{cosec}\pi z$ at the double pole at $z = -a$, and conclude that

$$\sum_{n=-\infty}^{\infty} \frac{(-1)^n}{(n+a)^2} = \frac{\pi^2\,\cos\pi a}{\sin^2 \pi a}\,. \tag{1.287}$$

Clearly there are large classes of infinite series that can be summed using these techniques. We shall encounter another example later, in a discussion of the Riemann zeta function.

1.9 Analytic Continuation

Analyticity of a function of a complex variable is a very restrictive condition, and consequently it has many powerful implications. One of these is the concept of *analytic continuation.* Let us begin with an example.

Consider the function $g(z)$, which is *defined* by the power series

$$g(z) \equiv \sum_{n\ge 0} z^n\,. \tag{1.288}$$

It is easily seen, by applying the Cauchy test for convergence, that this series is absolutely convergent for $|z| < 1$. It follows, therefore, that the function $g(z)$ defined by (1.288) is analytic inside the unit circle $|z| < 1$. It is also true, of course, that $g(z)$ is singular outside the unit circle; the power series diverges.

Of couse (1.288) is a very simple geometric series, and we can see by inspection that it can be summed, when $|z| < 1$, to give

$$f(z) = \frac{1}{1-z}\,. \tag{1.289}$$

This is analytic everywhere except for a pole at $z = 1$. So we have two functions, $g(z)$ and $f(z)$, which are both analytic inside the unit circle, and indeed they are identical inside the unit circle. However, whereas the function $g(z)$ is singular outside the unit circle, the function $f(z)$ is well-defined and analytic in the entire complex plane, with the exception of the point $z = 1$ where it has a simple pole.

It is evident, therefore, that we can view $f(z) = 1/(1-z)$ as an extrapolation, or continuation, of the function $g(z) = 1 + z + z^2 + \cdots$ outside its circle of convergence. As we shall prove below, there is an enormously powerful statement that can be made; the function $1/(1-z)$ is the *unique* analyic continuation of the original function $g(z)$ defined in the unit circle by (1.288). This uniqueness is absolutely crucial, since it means that one can sensibly talk about *the* analytic continuation of a function that is initially defined in some restricted region of the complex plane. *A priori*, one might have imagined that there could be any number of ways of defining functions that coincided with $g(z)$ inside the unit circle, but that extrapolated in all sorts of different ways as one went outside the unit circle. And indeed, if we don't place the extra, and very powerful, restriction of *analyticity*, then that would be exactly the case. We could indeed dream up all sorts of non-analytic functions that agreed with $g(z)$ inside the unit circle, and that extrapolated in arbitrary ways outside the unit circle.[10] The amazing thing is that if we insist that the extrapolating function be analytic, then there is precisely one, and only one, analytic continuation.

In the present example, we have the luxury of knowing that the function $g(z)$, *defined* by the series expansion (1.288), actually sums to give $1/(1-z)$ for any z within the unit circle. This immediately allows us to deduce, in this example, that the analytic continuation of $g(z)$ is precisely given by

$$g(z) = \frac{1}{1-z}\,, \tag{1.290}$$

which is defined everywhere in the complex plane except at $z = 1$. So in this toy example, we know what the function "really is."

Suppose, for a moment, that we didn't know that the series (1.288) could be summed to give (1.290). We could, however, discover that $g(z)$ defined by (1.288) gave perfectly sensible results for any z within the unit circle. (For example, by applying the Cauchy test for absolute convergence of the series.) Suppose that we use these results to evaluate $f(z)$ in the neighbourhood of the point $z = -\frac{1}{2}$. This allows us, by using Taylor's theorem, to

[10] We could, for example, simply *define* a function $F(z)$ such that $F(z) \equiv g(z)$ for $|z| < 1$, and $F(z) \equiv h(z)$ for $|z| \ge 1$, where $h(z)$ is *any function we wish*. But the function will in general be horribly non-analytic on the unit circle $|z| = 1$ where the changeover occurs.

construct a series expansion for $g(z)$ around the point $z = -\frac{1}{2}$:

$$g(z) = \sum_{n\geq 0} \frac{g^{(n)}(-\frac{1}{2})}{n!}\, (z + \tfrac{1}{2})^n\,. \tag{1.291}$$

Where does this converge? We know from the earlier general discussion that it will converge within a circle of radius R centred on $z = -\frac{1}{2}$, where R is the distance from $z = -\frac{1}{2}$ to the nearest singularity. We know that actually, this singularity is at $z = 1$. Therefore our new Taylor expansion (1.291) is convergent in a circle of radius $\frac{3}{2}$, centered on $z = -\frac{1}{2}$. This circle of convergence, and the original one, are depicted in Figure 6 below. We see that this process has taken us outside the original unit circle; we are now able to evaluate "the function $g(z)$" in a region outside the unit circle, where its original power-series expansion (1.288) does not converge.[11]

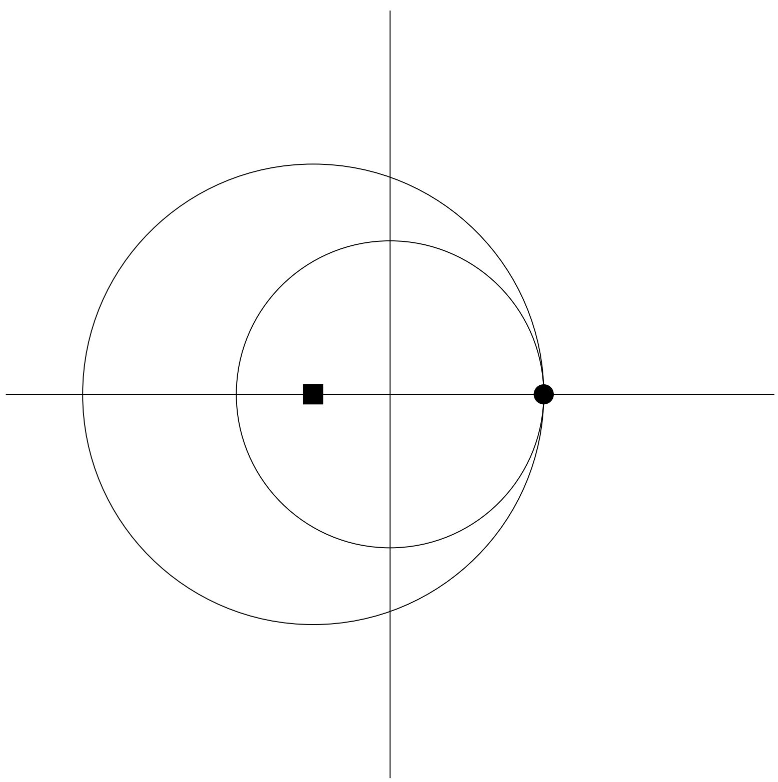

Figure 6: The circles of convergence for the two series

[11]Secretly, we know that the power series we will just have obtained is nothing but the standard Taylor expansion of $1/(1-z)$ around the point $z = -\frac{1}{2}$:

$$\frac{1}{1-z} = \tfrac{2}{3} + \tfrac{4}{9}(z + \tfrac{1}{2}) + \tfrac{8}{27}(z + \tfrac{1}{2})^2 + \tfrac{16}{81}(z + \tfrac{1}{2})^3 + \cdots\,, \tag{1.292}$$

which indeed converges in a circle of radius $\frac{3}{2}$.

It should be clear that be repeated use of this technique, we can eventually cover the entire complex plane, and hence construct the analytic continuation of $g(z)$ from its original definition (1.288) to a function defined everywhere except at $z = 1$.

The crucial point here is that the process of analytic continuation is a unique one. To show this, we can establish the following theorem:

> *Let $f(z)$ and $g(z)$ be two functions that are analytic in a region D, and suppose that they are equal on an infinite set of points having a limit point z_0 in D. Then $f(z) \equiv g(z)$ for all points z in D.*

In other words, if we know that the two analytic functions $f(z)$ and $g(z)$ agree on an arc of points ending at point[12] z_0 in D, then they must agree everywhere in D. (Note that we do not even need to know that they agree on a smooth arc; it is sufficient even to know that they agree on a discrete set of points that get denser and denser until the end of the arc at $z = z_0$ is reached.)

To prove this theorem, we first define $h(z) = f(z) - g(z)$. Thus we know that $h(z)$ is analytic in D, and it vanishes on an infinite set of points with limit point z_0. We are required to prove that $h(z)$ must be zero everywhere in D. We do this by expanding $h(z)$ in a Taylor series around $z = z_0$:

$$h(z) = \sum_{k=0}^{\infty} a_k\,(z - z_0)^k = a_0 + a_1\,(z - z_0) + \cdots\,, \tag{1.293}$$

which converges in some neighbourhood of z_0 since $h(z)$ is analytic in the neighbourhood of $z = z_0$. Since we want to prove that $h(z) = 0$, this means that we want to show that all the coefficients a_k are zero.

Of course since $h(z_0) = 0$ we know at least that $a_0 = 0$. We shall prove that all the a_k are zero by the time-honoured procedure of supposing that this is not true, and then arriving at a contradiction. Let us suppose that a_m, for some m, is the first non-zero a_k coefficient. This means that if we define

$$\begin{aligned} p(z) &\equiv (z - z_0)^{-m}\,h(z) = (z - z_0)^{-m} \sum_{k=m}^{\infty} a_k\,(z - z_0)^k\,, \\ &= a_m + a_{m+1}\,(z - z_0) + \cdots\,, \end{aligned} \tag{1.294}$$

then $p(z)$ is an analytic function, and its Taylor series is therefore also convergent, in the neighbourhood of $z = z_0$. Now comes the punch-line. We know that $h(z)$ is zero for all

[12] An example of such a set of points would be $z_n = z_0 + 1/n$, with $n = 1, 2, 3 \ldots$.

the points $z = z_n$ in that infinite set that has z_0 as a limit point. Thus in particular there are points z_n with n very large that are arbitrarily close to $z = z_0$, and at which $h(z)$ vanishes. It follows from its definition that $p(z)$ must also vanish at these points. But since the Taylor series for $p(z)$ is convergent for points z near to $z = z_0$, it follows that for $p(z_n)$ to vanish when n is very large we must have $a_m = 0$, since all the higher terms in the Taylor series would be negligible. But this contradicts our assumption that a_m was the first non-vanishing coefficient in (1.293). Thus the premise that there exists a first non-vanishing coefficient was false, and so it must be that *all* the coefficients a_k vanish. This proves that $h(z) = 0$, which is what we wanted to show.

The above proof shows that $h(z)$ must vanish within the circle of convergence, centered on $z = z_0$, of the Taylor series (1.293). By repeating the discussion as necessary, we can extend this region gradually until the whole of the domain D has been covered. Thus we have established that $f(z) = g(z)$ everywhere in D, if they agree on an infinite set of points with limit point z_0.

By this means, we may eventually seek to analytically extend the function to the whole complex plane. There may well be singularities at certain places, but provided we don't run into a solid "wall" of singularities, we can get around them and extend the definition of the function as far as we wish. Of course if the function has branch points, then we will encounter all the usual multi-valuedness issues as we seek to extend the function.

Let us go back for a moment to our example with the function $g(z)$ that was originally defined by the power series (1.288). We can now immediately invoke this theorem. It is easily established that the series (1.288) sums to give $1/(1-z)$ within the unit circle. Thus we have two analytic functions, namely $g(z)$ defined by (1.288) and $f(z)$ defined by (1.289) that agree in the entire unit circle. (Much more than just an arc with a limit point, in fact!) Therefore, it follows that there is a unique way to extend analytically outside the unit circle. Since $f(z) = 1/(1-z)$ is certainly analytic outside the unit circle, it follows that the function $1/(1-z)$ is the unique analytic extension of $g(z)$ defined by the power series (1.288).

Let us now consider a less trivial example, to show the power of analytic continuation.

1.10 The Gamma Function

The Gamma function $\Gamma(z)$ can be represented by the integral

$$\Gamma(z) = \int_0^\infty e^{-t}\, t^{z-1}\, dt\,, \tag{1.295}$$

which converges if $\mathrm{Re}(z) > 0$. It is easy to see that if $\mathrm{Re}(z) > 1$ then we can perform an integration by parts to obtain

$$\Gamma(z) = (z-1)\int_0^\infty e^{-t}\,t^{z-2}\,dt - \Big[e^{-t}\,t^{z-1}\Big]_0^\infty = (z-1)\,\Gamma(z-1)\,, \tag{1.296}$$

since the boundary term then gives no contribution. Shifting by 1 for convenience, we have

$$\Gamma(z+1) = z\,\Gamma(z)\,. \tag{1.297}$$

One easily sees that if z is a positive integer k, the solution to this recursion relation is $\Gamma(k) = (k-1)!$, since it is easily established by elementary integration that $\Gamma(1) = 1$. The responsibility for the rather tiresome shift by 1 in the relation $\Gamma(k) = (k-1)!$ lies with Leonhard Euler.

Of course the definition (1.295) is valid only when the integral converges. It's clear that the e^{-t} factor ensures that there is no trouble from the upper limit of integration, but from $t = 0$ there will be a divergence unless $\mathrm{Re}(z) > 0$. Furthermore, for $\mathrm{Re}(z) > 0$ it is clear that we can differentiate (1.295) with respect to z as many times as we wish, and the integrals will still converge.[13] Thus $\Gamma(z)$ defined by (1.295) is *finite* and analytic for all points with $\mathrm{Re}(z) > 0$.

We can now use (1.297) in order to give an analytic contiuation of $\Gamma(z)$ into the region where $\mathrm{Re}(z) \le 0$. Specifically, if we write (1.297) as

$$\Gamma(z) = \frac{\Gamma(z+1)}{z}\,, \tag{1.299}$$

then this gives a way of evaluating $\Gamma(z)$ for points in the strip $-1+\epsilon < \mathrm{Re}(z) < \epsilon$ (ϵ a small positive quantity) in terms of $\Gamma(z+1)$ at points with $\mathrm{Re}(z+1) > 0$, where $\Gamma(z+1)$ is known to be analytic. The function so defined, and the original Gamma function, have an overlapping region of convergence, and so we can make an analytic continuation into the strip $-1+\epsilon < \mathrm{Re}(z) < \epsilon$. The process can then be applied iteratively, to cover more and more strips over to the left-hand side of the complex plane, until the whole plane has been covered by the analytic extension. Thus by sending $z \to z+1$ in (1.299) we may write

$$\Gamma(z+1) = \frac{\Gamma(z+2)}{z+1}\,, \tag{1.300}$$

[13]Write $t^z = e^{z\,\log t}$, and so, for example,

$$\Gamma'(z) = \int_0^\infty dt\,t^{z-1}\,\log t\,e^{-t}\,. \tag{1.298}$$

Now matter how many powers of $\log t$ are brought down by repeated differentiation, the factor of t^{z-1} will ensure convergence at $t = 0$.

and plugging this into (1.299) itself we get

$$\Gamma(z) = \frac{\Gamma(z+2)}{(z+1)\,z}\,. \tag{1.301}$$

The right-hand side is analytic for $\mathrm{Re}(z) > -2$, save for the two manifest poles at $z = 0$ and $z = -1$, and so this has provided us with an analytic continuation of $\Gamma(z)$ into the region $\mathrm{Re}(z) > -2$. In the next iteration we use (1.299) with $z \to z+2$ to express $\Gamma(z+2)$ as $\Gamma(z+3)/(z+2)$, hence giving

$$\Gamma(z) = \frac{\Gamma(z+3)}{(z+2)(z+1)\,z}\,, \tag{1.302}$$

valid for $\mathrm{Re}(z) > -3$, and so on.

Of course the analytically continued function $\Gamma(z)$ is not necessarily analytic at every point in the complex plane, and indeed, we are already seeing, it has isolated poles. To explore the behaviour of $\Gamma(z)$ in the region of some point z with $\mathrm{Re}(z) \le 0$, we first iterate (1.297) just as many times n as are necessary in order to express $\Gamma(z)$ in terms of $\Gamma(z+n+1)$:

$$\Gamma(z) = \frac{\Gamma(z+n+1)}{(z+n)(z+n-1)(z+n-2)\cdots z}\,, \tag{1.303}$$

where we choose n so that $\mathrm{Re}(z+n+1) > 0$ but $\mathrm{Re}(z+n) < 0$. Since we have already established that $\Gamma(z+n+1)$ will therefore be finite, it follows that the only singularities of $\Gamma(z)$ can come from places where the denominator in (1.303) vanishes. This will therefore happen when $z = 0$ or z is a negative integer.

To study the precise behaviour near the point $z = -n$, we may set $z = -n + \epsilon$, where $|\epsilon| << 1$, and use (1.303) to give

$$\Gamma(-n+\epsilon) = \frac{(-1)^n\,\Gamma(1+\epsilon)}{(n-\epsilon)(n-\epsilon-1)\cdots(1-\epsilon)\,\epsilon} = \frac{(-1)^n}{n\,(n-1)\cdots 2\cdot 1\,\epsilon} + \cdots\,, \tag{1.304}$$

where the terms represented by $\cdots$ are analytic in ϵ. Thus there is a simple pole at $\epsilon = 0$. Its residue is calulated by multiplying (1.304) by ϵ and taking the limit $\epsilon \longrightarrow 0$. Thus we conclude that $\Gamma(z)$ is meromorphic in the whole finite complex plane, with simple poles at the points $z = 0,\ -1,\ -2,\ -3, \ldots$, with the residue at $z = -n$ being $(-1)^n/n!$. (Since $\Gamma(1) = 1$.)

The regular spacing of the poles of $\Gamma(z)$ is reminiscent of the poles of the functions $\mathrm{cosec}\,\pi z$ or $\cot \pi z$. Of course in these cases, they have simple poles at *all* the integers; zero negative and positive. We can in fact make a function with precisely this property out of $\Gamma(z)$, by writing the product

$$\Gamma(z)\,\Gamma(1-z)\,. \tag{1.305}$$

From what we saw above, it is clear that this function will have simple poles at precisely all the integers. Might it be that this function is related to $\operatorname{cosec}\pi z$ or $\cot\pi z$?

To answer this, consider again the original integral representation (1.295) for $\Gamma(z)$, and now make the change of variables $t \longrightarrow t^2$. This implies $dt/t \longrightarrow 2dt/t$, and so we shall have

$$\Gamma(z) = 2\int_0^\infty e^{-t^2}\, t^{2z-1}\, dt\,. \tag{1.306}$$

Thus we may write

$$\Gamma(a)\,\Gamma(1-a) = 4\int_0^\infty dx \int_0^\infty dy\, e^{-(x^2+y^2)}\, x^{2a-1}\, y^{-2a+1}\,. \tag{1.307}$$

Introducing polar coordinates *via* $x = r\,\cos\theta$, $y = r\,\sin\theta$, we therefore get

$$\Gamma(a)\,\Gamma(1-a) = 4\int_0^{\frac{1}{2}\pi} (\cot\theta)^{2a-1}\, d\theta \int_0^\infty r\, e^{-r^2}\, dr\,. \tag{1.308}$$

The r integration is trivially performed, giving a factor of $\frac{1}{2}$, and so we have

$$\Gamma(a)\,\Gamma(1-a) = 2\int_0^{\frac{1}{2}\pi} (\cot\theta)^{2a-1}\, d\theta\,. \tag{1.309}$$

Now, we let $s = \cot\theta$. This gives

$$\Gamma(a)\,\Gamma(1-a) = 2\int_0^\infty \frac{s^{2a-1}\, ds}{1+s^2}\,. \tag{1.310}$$

If we restrict a such that $0 < \mathrm{Re}(a) < 1$, this integral falls into the category of type 3 that we discussed a couple of sections ago. Thus we have

$$\Gamma(a)\,\Gamma(1-a) = \frac{2\pi}{\sin(2\pi\, a)} \sum_c \mathcal{R}_c\,, \tag{1.311}$$

where $\mathcal{R}_c$ are the residues at the poles of $(-z)^{2a-1}/(1+z^2)$. These poles lie at $z = \pm\mathrm{i}$, and the residues are easily seen to be $\frac{1}{2}e^{\pm\mathrm{i}\,\pi\, a}$. Thus we get

$$\begin{aligned}\Gamma(a)\,\Gamma(1-a) &= \frac{2\pi}{\sin(2\pi\, a)}\, \cos(\pi\, a) = \frac{2\pi\, \cos(\pi\, a)}{2\sin(\pi\, a)\, \cos(\pi\, a)}\,,\\ &= \frac{\pi}{\sin\pi\, a}\,. \end{aligned} \tag{1.312}$$

Although we derived this by restricting a such that $0 < \mathrm{Re}(a) < 1$ in order to ensure convergence in the integration, we can use the now-familiar technique of analytic continuation and conclude that

$$\Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\sin\pi\, z}\,, \tag{1.313}$$

in the whole complex plane. This result, known as the reflection formula, is one that will be useful in the next section, when we shall discuss the Riemann Zeta function.

Before moving on to the Riemann Zeta function, let us first use (1.313) to uncover a couple more properties of the Gamma function. The first of these is a simple fact, namely that

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}\,. \tag{1.314}$$

We see this by setting $z = \frac{1}{2}$ in (1.313).

The second, more significant, property of $\Gamma(z)$ that we can deduce from (1.313) is that $\Gamma(z)^{-1}$ an *entire* function. That is to say, $\Gamma(z)^{-1}$ is analytic everywhere in the finite complex plane. Since we have already seen that the only singularities of $\Gamma(z)$ are poles, this means that we need only show that $\Gamma(z)$ has no zeros in the finite complex plane. Looking at (1.313) we see that if it were to be the case that $\Gamma(z) = 0$ for some value of z, then it would have to be that $\Gamma(1-z)$ were infinite there.[14] But we know precisely where $\Gamma(1-z)$ is infinite, namely the poles at $z = 1, 2, 3\ldots$, and $\Gamma(z)$ is certainly not zero there.[15] Therefore $\Gamma(z)$ is everywhere non-zero in the finite complex plane. Consequently, $\Gamma(z)^{-1}$ is analytic everywhere in the finite complex plane, thus proving the contention that $\Gamma(z)^{-1}$ is an entire function.

Before closing this section, we may observe that we can also give a contour integral representation for the Gamma function. This will have the nice feature that it will provide us directly with an expression for $\Gamma(z)$ that is valid in the whole complex plane. Consider first the *Hankel integral*

$$\Gamma(z) = -\frac{1}{2\,\mathrm{i}\,\sin\pi z}\int_C e^{-t}\,(-t)^{z-1}\,dt\,, \tag{1.315}$$

where we integrate in the complex t-plane around the so-called *Hankel Contour* depicted in Figure 7 below. This starts at $+\infty$ just above the real axis, swings around the origin, and goes out to $+\infty$ again just below the real axis. As usual, we shall run the branch cut for the multi-valued function $(-t)^{z-1}$ along the positive real axis in the complex t plane, and $(-t)^{z-1}$ will be taken to be real and positive when t lies on the negative real axis.

We see can deform the contour in Figure 7 into the contour depicted in Figure 8, since no singularities are crossed in the process. If $\mathrm{Re}(z) > 0$, there will be no contribution from integrating around the small circle surrounding the origin, in the limit where its radius is sent to zero. Hence the contour integral is re-expressible simply in terms of the two semi-infinite line integrals just above and below the real axis.

[14] Recall that $\sin\pi z$ is an entire function, and it therefore has no singularity in the finite complex plane. Consequently, $1/(\sin\pi z)$ must be non-vanishing for all finite z.

[15] Instead, the poles of $\Gamma(1-z)$ at $z = 1, 2, 3\ldots$ are balanced in (1.313) by the poles in $1/\sin(\pi\,z)$.

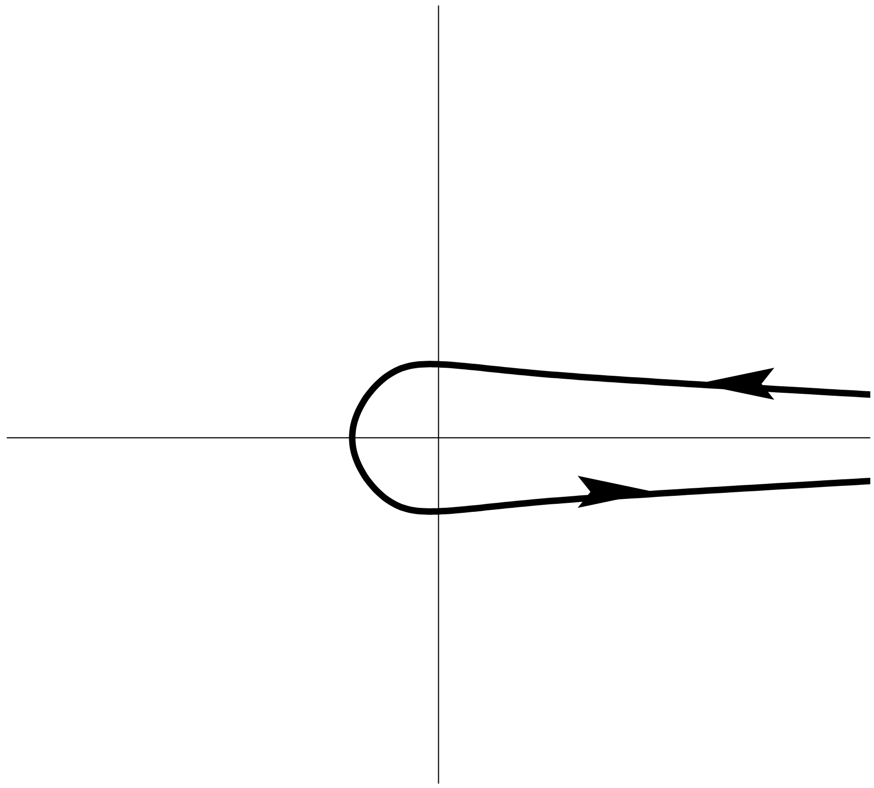

Figure 7: The Hankel contour

The integrals along the lower and upper causeways in Figure 8, we follow the same procedure that we have used before. We define the phase of $(-t)^z$ to be zero when t lies on the negative real t axis, and run the branch cut along the positive real t axis. For the integral below the real axis, we therefore have $(-t) = e^{\pi\,\mathrm{i}}\,x$, with x running from 0 to $+\infty$. For the integral above the real axis, we have $(-t) = e^{-\mathrm{i}\,\pi}\,x$, with x running from $+\infty$ to 0. Consequently, we get

$$\begin{aligned} \int_C e^{-t}\,(-t)^{z-1}\,dt &= (e^{\mathrm{i}\,\pi\,(z-1)} - e^{-\mathrm{i}\,\pi\,(z-1)}) \int_0^\infty e^{-t}\,t^{z-1}\,dt\,, \\ &= -2\,\mathrm{i}\,\sin(\pi z) \int_0^\infty e^{-t}\,t^{z-1}\,dt\,, \end{aligned} \tag{1.316}$$

and hence we see that (1.315) has reduced to the original real integral expression (1.295) when $\mathrm{Re}(z) > 0$. However, the integral in the expression (1.315) has a much wider applicability; it is actually single-valued and analytic for *all* z. (Recall that we are integrating around the Hankel contour, which does not pass through the point $t = 0$, and so there is no reason for any singularity to arise, for any value of z.) The poles in $\Gamma(z)$ (which we know from our earlier discussion to occur at $z = 0, -1, -2, \ldots$) must therefore be due entirely to the $1/\sin(\pi z)$ prefactor in (1.316. Indeed, as we saw a while ago, $1/\sin(\pi z)$ has simple

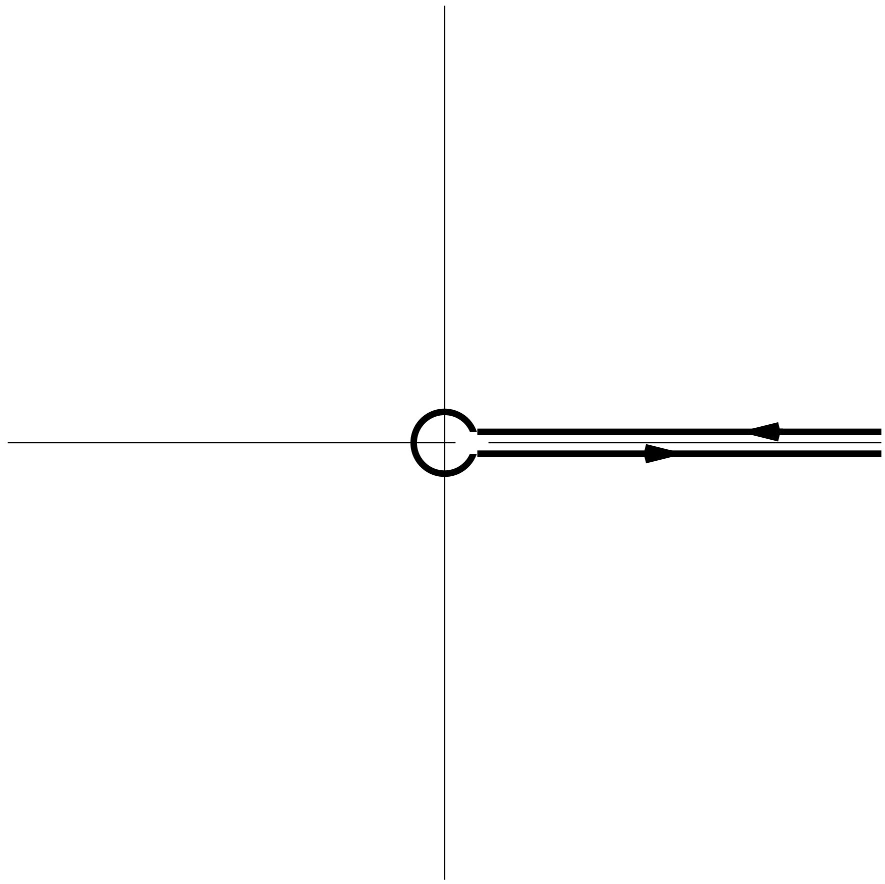

Figure 8: The deformation of the Hankel contour

poles when z is an integer.[16]

Combining (1.315) with (1.313), we can give another contour integral expression for $\Gamma(z)$, namely

$$\frac{1}{\Gamma(z)} = -\frac{1}{2\pi\,\mathrm{i}} \int_C e^{-t}\,(-t)^{-z}\,dt\,, \tag{1.317}$$

where we again integrate around the Hankel contour of Figure 7, in the complex t plane. Again, this integral is valid for all z. Indeed with this expression we see again the result that we previously deduced from (1.313), that $\Gamma(z)^{-1}$ is an *entire* function, having no singularities anywhere in the finite complex plane.

A pause for reflection is appropriate here. What we have shown is that $\Gamma(z)$ defined by (1.315) or (1.317) gives the analytic continuation of our original Gamma function (1.295) to

[16]The reason why (1.316) doesn't also imply that $\Gamma(z)$ has simple poles when z is a positive integer is that the integral itself vanishes when z is a positive integer, and this cancels out the pole from $1/\sin(\pi z)$. This vanishing can be seen from the fact that when z is a positive integer, the integrand is analytic (there is no longer a branch cut), the contour can be closed off at infinity to make a closed contour encircling the origin, and hence Cauchy's theorem implies the integral vanishes.

the entire complex plane, where it is analytic except for simple poles at $z = 0, -1, -2, \ldots$. How is it that these contour integrals do better than the previous real integral (1.295), which only converged when $\mathrm{Re}(z) > 0$? The crucial point is that in our derivation, when we related the real integral in (1.295) to the contour integral (1.315), we noted that the contribution from the little circle as the contour swung around the origin would go to zero provided that the real part of z was greater than 0.

So what has happened is that we have re-expressed the real integral in (1.295) in terms of a contour integral of the form (1.315), which gives the same answer when the real part of z is greater than 0, but it disagrees when the real part of z is ≤ 0. In fact it disagrees by the having the rather nice feature of being convergent and analytic when $\mathrm{Re}(z) \le 0$, unlike the real integral that diverges. So as we wander off westwards in the complex z plane we wave a fond farewell to the real integral, with its divergent result, and adopt instead the result from the contour integral, which happily provides us with analytic answers even when $\mathrm{Re}(z) \le 0$. We should not be worried by the fact that the integrals are disagreeing there; quite the contrary, in fact. The whole point of the exercise was to find a better way of representing the function, to cover a wider region in the complex plane. If we had merely reproduced the bad behaviour of the original integral in (1.295), we would have achieved nothing by introducing the contour integrals (1.315) and (1.317).

Now we turn to the Riemman Zeta function, as a slightly more intricate example of the analytic continuation of a function of a complex variable.

1.11 The Riemann Zeta Function

Consider the *Riemman Zeta Function*, $\zeta(s)$. This is originally defined by

$$\zeta(s) \equiv \sum_{n=1}^{\infty} \frac{1}{n^s}\,. \tag{1.318}$$

This sum converges whenever the real part of s is greater than 1. (For example, $\zeta(2) = \sum_{n\ge1} n^{-2}$ can be shown to equal $\pi^2/6$, whereas $\zeta(1) = \sum_{n\ge1} n^{-1}$ is logarithmically divergent. The sum is more and more divergent as $\mathrm{Re}(s)$ becomes less than 1.)

Since the series (1.318) defining $\zeta(s)$ is convergent everywhere to the right of the line $\mathrm{Re}(s) = 1$ in the complex plane, it follows that $\zeta(s)$ is analytic in that region. It is reasonable to ask what is its analytic continuation over to the left of $\mathrm{Re}(s) = 1$. As we have already seen from the simple example of $f(z) = 1/(1-z)$, the mere fact that our original power series diverges in the region with $\mathrm{Re}(s) \le 0$ does not in any way imply that the "actual" function $\zeta(s)$ will behave badly there. It is just our power series that is inadequate.

How do we do better? To begin, recall that we define the Gamma function $\Gamma(s)$ by

$$\Gamma(s) = \int_0^\infty e^{-u}\, u^{s-1}\, du \tag{1.319}$$

We saw in the previous section that if $s = k$, where k is an integer, then $\Gamma(k)$ is nothing but the factorial function $(k-1)!$. If we now let $u = n\, t$, then we see that

$$\Gamma(s) = n^s \int_0^\infty e^{-n\, t}\, t^{s-1}\, dt\,. \tag{1.320}$$

We can turn this around, to get an expression for n^{-s}.

Plugging into the definition (1.318) of the zeta function, we therefore have

$$\zeta(s) = \frac{1}{\Gamma(s)} \sum_{n=1}^\infty \int_0^\infty e^{-n\, t}\, t^{s-1}\, dt\,. \tag{1.321}$$

Taking the summation through the integral, we see that we have a simple geometric series, which can be summed explicitly:

$$\sum_{n=1}^\infty e^{-n\, t} = \frac{1}{1-e^{-t}} - 1 = \frac{1}{e^t - 1}\,, \tag{1.322}$$

and hence we arrive at the following integral representation for the zeta function:

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{t^{s-1}\, dt}{e^t - 1}\,. \tag{1.323}$$

So far so good, but actually we haven't yet managed to cross the barrier of the $\mathrm{Re}(s) = 1$ line in the complex plane. The denominator in the integrand goes to zero like t as t tends to zero, so to avoid a divergence from the integration at the lower limit $t = 0$, we must insist that the real part of s should be greater than 1. This is the same restriction that we encountered for the original power series (1.318). What we do now is to turn our real integral (1.323) into a complex contour integral, using the same sort of ideas that we used in the previous section.

To do this, consider the integral

$$\int_C \frac{(-z)^{s-1}\, dz}{e^z - 1}\,, \tag{1.324}$$

where C is the same Hankel contour, depicted in Figure 7, that we used in the discussion of the Gamma function in the previous section. Since the integrand we are considering here clearly has poles at $z = 2\pi\, \mathrm{i}\, n$ for all the integers n, we must make sure that as it circles round the origin, the Hankel contour keeps close enough to the origin (with passing through it) so that it does not encompass any of the poles at $z = \pm 2\pi\, \mathrm{i}, \pm 4\pi\, \mathrm{i}, \ldots$.

By methods analogous to those we used previously, we see that we can again deform this into the contour depicted in Figure 8, where the small circle around the origin will be sent to zero radius. It is clear that there is no contribution from the little circle, provided that the real part of s is greater than 1. Hence the contour integral is re-expressible simply in terms of the two semi-infinite line integrals just above and below the real axis.

As usual, we choose to run the branch cut of the function $(-z)^{s-1}$ along the positive real axis, and take $(-z)^{s-1}$ to be real and positive when z lies on the negative real axis. For the integral below the real axis, we shall then have $(-z) = e^{\pi\,\mathrm{i}}\,t$, with t running from 0 to $+\infty$. For the integral above the real axis, we shall have $(-z) = e^{-\mathrm{i}\,\pi}\,t$, with t running from $+\infty$ to 0. Consequently, we get

$$\int_C \frac{(-z)^{s-1}\,dz}{e^z - 1} = \Big(e^{\mathrm{i}\,\pi\,(s-1)} - e^{-\mathrm{i}\,\pi\,(s-1)}\Big) \int_0^\infty \frac{t^{s-1}\,dt}{e^t - 1} = -2\mathrm{i}\,\sin\pi s \int_0^\infty \frac{t^{s-1}\,dt}{e^t - 1}\,, \tag{1.325}$$

From (1.323), this means that we have a new expression for the zeta function, as

$$\zeta(s) = -\frac{1}{2\mathrm{i}\,\Gamma(s)\,\sin\pi s} \int_C \frac{(-z)^{s-1}\,dz}{e^z - 1}\,. \tag{1.326}$$

We can neaten this result up a bit more, if we make use of the reflection formula (1.313) satisfied by the Gamma function, which we proved in the previous section:

$$\Gamma(s)\,\Gamma(1-s) = \frac{\pi}{\sin\pi s}\,. \tag{1.327}$$

Using this in (1.326), we arrive at the final result

$$\zeta(s) = -\frac{\Gamma(1-s)}{2\pi\,\mathrm{i}} \int_C \frac{(-z)^{s-1}\,dz}{e^z - 1}\,. \tag{1.328}$$

Now comes the punch-line. The integral in (1.328) is a single-valued and analytic function of s *for all values of* s. (Recall that it is evaluated using the Hankel contour in Figure 7, which does not pass through $z = 0$. And far out at the right-hand side of the Hankel contour, the e^z factor in the denominator will ensure rapid convergence. Thus there is no reason for any singular behaviour.) Consequently, the only possible non-analyticity of the zeta function can come from the $\Gamma(1-s)$ prefactor. Now, we studied the singularities of the Gamma function in the previous section. The answer is that $\Gamma(1-s)$ has simple poles at $s = 1, 2, 3, \ldots$, and no other singularities. So these are the only possible points in the finite complex plane where $\zeta(s)$ might have poles. But we already know that $\zeta(s)$ is analytic whenever the real part of s is greater than 1. So it must in fact be the case that the poles of $\Gamma(1-s)$ at $s = 2, 3, \ldots$ are exactly cancelled by zeros coming from the integral in (1.328). Only the pole at $s = 1$ might survive, since we have no independent argument that tells us that $\zeta(s)$ is analytic there. And in fact there *is* a pole in $\zeta(s)$ there.

To see this, we need only to evaluate the integral in (1.328) at $s = 1$. This is an easy task. It is given by

$$\frac{1}{2\pi\,\mathrm{i}} \int_C \frac{dz}{e^z - 1}\,. \tag{1.329}$$

Now, since we no longer have a multi-valued function in the integrand we don't have to worry about a branch cut along the positive real axis. The integrand has become infinitesimally small out at the right-hand ends of the Hankel contour, and so we can simply join the two ends together without affecting the value of the integral. We now have a closed contour encircling the origin, and so we can evaluate it using the residue theorem; we just need to know the residue of the integrand at $z = 0$. Doing the series expansion, one finds

$$\frac{1}{e^z - 1} = \frac{1}{z} - \tfrac{1}{2} + \tfrac{1}{12} z - \tfrac{1}{720}\, z^3 + \cdots \tag{1.330}$$

so the residue is 1. From (1.328), this means that near to $s = 1$ we shall have

$$\zeta(s) \sim -\Gamma(1 - s)\,. \tag{1.331}$$

In fact $\Gamma(1 - s)$ has a simple pole of residue -1 at $s = 1$, as we saw in the previous section, and so the upshot is that $\zeta(s)$ has a simple pole of residue $+1$ at $s = 1$, but it is otherwise analytic everywhere.

It is interesting to try working out $\zeta(s)$ for some values of s that were inaccessible in the original series definition (1.318). For example, let us consider $\zeta(0)$. From (1.328) we therefore have

$$\zeta(0) = \frac{1}{2\pi\,\mathrm{i}} \int_C \frac{dz}{z\,(e^z - 1)}\,, \tag{1.332}$$

where we have used that $\Gamma(1) = 1$. Again, we can close off the Hankel contour of Figure 7 out near $+\infty$, since there is no branch cut, and the e^z in the denominator means that the integrand is vanishly small there. We therefore just need to use the calculus of residues to evaluate (1.332), for a closed contour encircling the second-order pole at $z = 0$. For this, we have

$$\frac{1}{z\,(e^z - 1)} = \frac{1}{z^2} - \frac{1}{2z} + \frac{1}{12} + \cdots\,, \tag{1.333}$$

showing that the residue is $-\frac{1}{2}$. Thus we obtain the result

$$\zeta(0) = -\tfrac{1}{2}\,. \tag{1.334}$$

One can view this result rather whimsically as a "regularisation" of the divergent expression that one would obtain from the original series definition of $\zeta(s)$ in (1.318):

$$\zeta(0) = \sum_{n\ge 1} n^0 = \sum_{n\ge 1} 1 = 1 + 1 + 1 + 1 + \cdots = -\tfrac{1}{2}\,. \tag{1.335}$$

Actually, this strange-looking formula is not entirely whimsical. It is precisely the sort of divergent sum that arises in a typical Feynman diagram loop calculation in quantum field theory (corresponding, for example, to summing the zero-point energies of an infinite number of harmonic oscillators). The whole subtlety of handling the infinities in quantum field theory is concerned with how to recognise and subtract out unphysical divergences associated, for example, with the infinite zero-point energy of the vacuum. This process of renormalisation and regularisation can actually, remarkably, be made respectable, and in particular, it can be shown that the final results are independent of the regularisation scheme that one uses. One scheme that has been developed is known as "Zeta Function Regularisation," and it consists precisely of introducing regularisation parameters that cause a divergent sum such as (1.335) to be replaced by $\sum_{n\ge1} n^{-s}$. The regularisation scheme (whose rigour can be proved up to the "industry standards" of the subject) then consists of replacing the infinite result for $\sum_{n\ge1} 1$ by the expression $\zeta(0)$, where $\zeta(s)$ is the analytically-continued function defined in (1.328).

The Riemann zeta function is very important also in number theory. This goes beyond the scope of this course, but a couple of remarks on the subject may be of interest. First, we may make the following manipulation, valid for $\mathrm{Re}(s) > 1$:

$$\begin{aligned}\zeta(s) &= \sum_{n\ge1} n^{-s} = 1^{-s} + 2^{-s} + 3^{-s} + 4^{-s} + 5^{-s} + 6^{-s} + 7^{-s} + \cdots \\ &= 1^{-s} + 3^{-s} + 5^{-s} + \cdots + 2^{-s}\,(1^{-s} + 2^{-s} + 3^{-s} + \cdots) \\ &= 1^{-s} + 3^{-s} + 5^{-s} + \cdots + 2^{-s}\,\zeta(s)\,,\end{aligned} \tag{1.336}$$

whence

$$(1-2^{-s})\,\zeta(s) = 1^{-s} + 3^{-s} + 5^{-s} + \cdots\,. \tag{1.337}$$

So all the terms where n is a multiple of 2 are now omitted in the sum. Now, repeat this excercise but pulling out a factor of 3^{-s}:

$$\begin{aligned}(1-2^{-s})\,\zeta(s) &= 1^{-s} + 5^{-s} + 7^{-s} + 11^{-s} + \cdots + 3^{-s}\,(1^{-s} + 3^{-s} + 5^{-s} + 7^{-s} + \cdots)\,, \\ &= 1^{-s} + 5^{-s} + 7^{-s} + 11^{-s} + \cdots + 3^{-s}\,(1-2^{-s})\,\zeta(s)\,,\end{aligned} \tag{1.338}$$

whence

$$(1-2^{-s})\,(1-3^{-s})\,\zeta(s) = 1^{-s} + 5^{-s} + 7^{-s} + 11^{-s} + \cdots\,. \tag{1.339}$$

We have now have a sum where all the terms where n is a multiple of 2 or 3 are omitted. Next, we do the same for factors of 5, then 7, then 11, and so on. If $2, 3, 5, 7, \ldots, p$ denote all the prime numbers up to p, we shall have

$$(1-2^{-s})\,(1-3^{-s})\cdots(1-p^{-s})\,\zeta(s) = 1 + {\sum}' n^{-s}\,, \tag{1.340}$$

where $\sum'$ indicates that only those values of n that are prime to $2, 3, 5, 7, \ldots, p$ occur in the summation. It is now straightforward to show that if we send p to infinity, this summation goes to zero, since the "first" term in the sum is the lowest integer that is prime to all the primes, i.e. $n = \infty$. Since $\mathrm{Re}(s) > 1$, the "sum" is therefore zero. Hence we arrive at the result, known as *Euler's product* for the zeta function:

$$\frac{1}{\zeta(s)} = \prod_p \Big(1 - \frac{1}{p^s}\Big)\,, \qquad \mathrm{Re}(s) > 1\,, \tag{1.341}$$

where the product is over all the prime numbers. This indicates that the Riemann zeta function can play an important rôle in the study of prime numbers.

We conclude this section with an application of the technique we discussed in section 1.8, for summing infinite series by contour integral methods. It is relevant to the discussion of the zeros of the Riemann zeta function. Recall that we showed previously that the zeta function could be represented by the integral (1.328), which we repeat here:

$$\zeta(s) = -\frac{\Gamma(1-s)}{2\pi\,\mathrm{i}} \int_C \frac{(-z)^{s-1}\,dz}{e^z - 1}\,, \tag{1.342}$$

where C is the Hankel contour. Now, imagine making a closed contour C', consisting of a large outer circle, centred on the origin, and with radius $(2N+1)\,\pi$, which joins onto the Hankel contour way out to the east in the complex plane. See Figure 9 below. As we observed previously, the integrand in (1.342) has poles at $z = 2\pi\,\mathrm{i}\,n$ for all the integers n. In fact, of course, it is very similar to the cosec and cot functions that we have been considering in our discussion in this section, since

$$\frac{1}{e^z - 1} = \frac{e^{-\frac12 z}}{e^{\frac12 z} - e^{-\frac12 z}} = \tfrac12 e^{-\frac12 z}\,\mathrm{cosech}\,(\tfrac12 z)\,. \tag{1.343}$$

The only difference is that because we now have the hyperbolic function cosech rather than the trigonometric function cosec, the poles lie along the imaginary axis rather than the real axis.

In fact it is easy to find the nature of the poles at $z = 2\pi\,\mathrm{i}\,n$, by just writing $z = 2\pi\,\mathrm{i}\,n + w$. Noting that

$$e^z = e^{2\pi\,\mathrm{i}\,n + w} = e^{2\pi\,\mathrm{i}\,n}\,e^w = e^w\,, \tag{1.344}$$

we see that

$$\begin{aligned}\frac{1}{e^z-1} &= \frac{1}{e^w-1} = \frac{1}{(1+w+\frac{1}{2}w^2+\cdots)-1} = \frac{1}{w(1+\frac{1}{2}w+\cdots)} \\ &= \frac{1}{w}+\tfrac{1}{2}+\cdots = \frac{1}{z-2\pi\,\mathrm{i}\,n}+\tfrac{1}{2}+\cdots\,. \end{aligned} \tag{1.345}$$

Thus the pole at $z = 2\pi\,\mathrm{i}\,n$ is a simple pole, with residue 1.

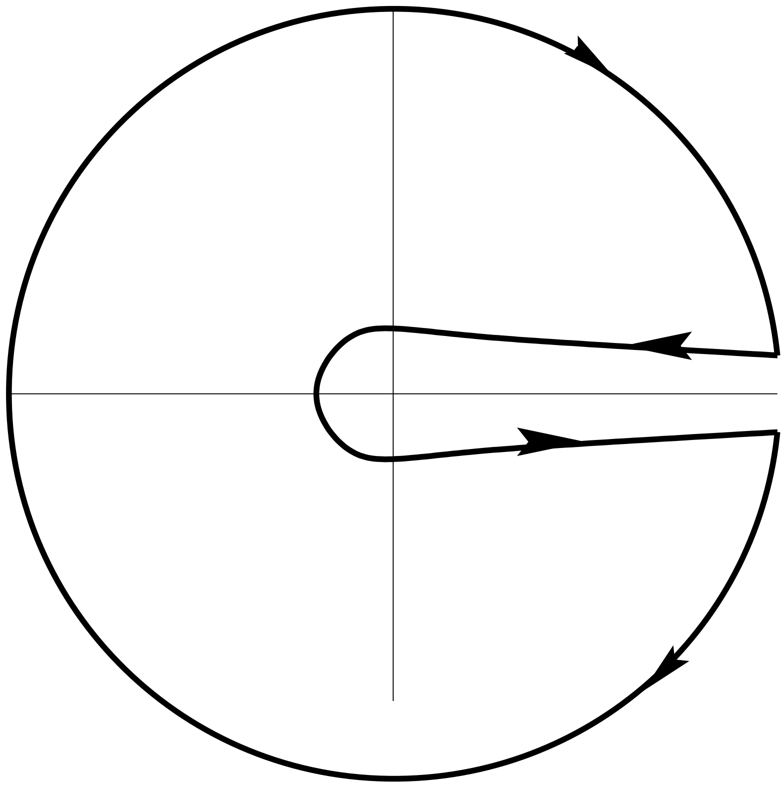

Figure 9: The contour C' composed of the Hankel contour plus a large circle

Since the Hankel contour itself was arranged so as to sneak around the origin without encompassing the poles at $z = \pm 2\pi\,\mathrm{i}, \pm 4\pi\,\mathrm{i}, \ldots$, it follows that the closed contour C' will precisely enclose the poles at $z = 2\pi\,\mathrm{i}\,n$, for all non-vanishing positive and negative integers n. For some given positive integer m, consider the pole at

$$z = 2\pi\,\mathrm{i}\,m = 2\pi\,e^{\frac{1}{2}\pi\,\mathrm{i}}\,m\,. \tag{1.346}$$

When we evaluate the residue $\mathcal{R}_m$ here, we therefore have

$$\mathcal{R}_m = (2\pi\,e^{-\frac{1}{2}\pi\,\mathrm{i}}\,m)^{s-1}\,, \tag{1.347}$$

since $(e^z-1)^{-1}$ itself clearly has a simple pole with residue 1 there. (We have used the fact that (1.346) implies $-z = 2\pi\,m\,e^{-\frac{1}{2}\pi\,\mathrm{i}}$, since we have to be careful when dealing with the multiply-valued function $(-z)^{s-1}$.) There is also a pole at $z = -2\pi\, e^{\frac{1}{2}\pi\,\mathrm{i}}\, m$, which by similar reasoning will have the residue $\mathcal{R}_{-m}$ given by

$$\mathcal{R}_{-m} = (2\pi\, e^{\frac{1}{2}\pi\,\mathrm{i}}\, m)^{s-1}\,, \tag{1.348}$$

Putting the two together, we therefore get

$$\mathcal{R}_m + \mathcal{R}_{-m} = 2\,(2\pi\, m)^{s-1}\, \sin(\tfrac{1}{2}\pi\, s)\,. \tag{1.349}$$

By the theorem of residues, it follows that if we evaluate

$$\int_{C'} \frac{(-z)^{s-1}\, dz}{e^z - 1}\,, \tag{1.350}$$

where C' is the closed contour defined above, and then we send the radius $(2N+1)\pi$ of the outer circle to infinity, we shall get

$$\begin{aligned}
\int_{C'} \frac{(-z)^{s-1}\, dz}{e^z - 1} &= -2\pi\,\mathrm{i} \sum_{m\ge 1} (\mathcal{R}_m + \mathcal{R}_{-m})\,,\\
&= -4\pi\,\mathrm{i} \sum_{m\ge 1} (2\pi\, m)^{s-1}\, \sin(\tfrac{1}{2}\pi\, s)\\
&= -2\,(2\pi)^s\,\mathrm{i}\, \sin(\tfrac{1}{2}\pi\, s) \sum_{m\ge 1} m^{s-1}\,,\\
&= -2\,(2\pi)^s\,\mathrm{i}\, \sin(\tfrac{1}{2}\pi\, s)\, \zeta(1-s)\,.
\end{aligned} \tag{1.351}$$

It is clear from the final step that we should require $\mathrm{Re}(s) < 0$ here. (Note that the direction of the integration around large circle is *clockwise*, which is the direction of decreasing phase, so we pick up the extra -1 factor when using the theorem of residues.)

Now, if we consider the closed contour C' in detail, we find the following. It is comprised of the sum of the Hankel contour, plus the circle at large radius $R = (2N+1)\,\pi$, with N sent to infinity. On the large circle we shall have

$$|(-z)^{s-1}| = R^{s-1}\,, \tag{1.352}$$

which falls off faster than $1/R$ since we are requiring $\mathrm{Re}(s) < 0$. This is enough to outweigh the factor of R that comes from writing $z = R\, e^{\mathrm{i}\,\theta}$ on the large circle. Since the $(e^z - 1)^{-1}$ factor cannot introduce any divergence (the radii $R = (2N+1)\,\pi$ are cleverly designed to avoid passing through the poles of $(e^z - 1)^{-1}$), it follows that the contribution from integrating around the large circle goes to zero as N is sent to infinity. Therefore when evaluating the contour integral on the left-hand side of (1.351), we are left only with the contribution from the Hankel contour C. But from (1.342), this means that we have

$$\int_{C'} \frac{(-z)^{s-1}\, dz}{e^z - 1} = \int_{C} \frac{(-z)^{s-1}\, dz}{e^z - 1} = -\frac{2\pi\,\mathrm{i}}{\Gamma(1-s)}\, \zeta(s)\,. \tag{1.353}$$

Comparing with (1.351), we therefore conclude that if $\mathrm{Re}(s) < 0$,

$$\zeta(s) = 2\,(2\pi)^{s-1}\,\Gamma(1-s)\,\sin(\tfrac{1}{2}\pi s)\,\zeta(1-s)\,. \tag{1.354}$$

This can be neatened up using the reflection formula (1.313) to write $\Gamma(1-s) = \pi/(\Gamma(s)\,\sin(\pi s))$, and then using the fact that $\sin(\pi s) = 2\sin(\frac{1}{2}\pi s)\,\cos(\frac{1}{2}\pi s)$. This gives us the final result

$$2^{s-1}\,\Gamma(s)\,\zeta(s)\,\cos(\tfrac{1}{2}\pi s) = \pi^s\,\zeta(1-s)\,, \tag{1.355}$$

Both sides are analytic functions, except at isolated poles, and so even though we derived the result under the restriction $\mathrm{Re}(s) < 0$, it immediately follows by analytic continuation that it is valid in the whole complex plane. This beautiful formula was discovered by Riemann.

There is a very important, and still unproven conjecture, known as *Riemann's Hypothesis*. This concerns the location of the zeros of the zeta function. One can easily see from Euler's product (1.341), or from the original series definition (1.318), that $\zeta(s)$ has no zeros for $\mathrm{Re}(s) > 1$. One can also rather easily show, using Riemann's formula that we derived above, that when $\mathrm{Re}(s) < 0$ the only zeros lie at the negative even integers, $s = -2, -4, \ldots$. This leaves the strip $0 \le \mathrm{Re}(s) \le 1$ unaccounted for. Riemann's Hypothesis, whose proof would have far-reaching consequences in number theory, is that in this strip, all the zeros of $\zeta(s)$ lie on the line $\mathrm{Re}(s) = \frac{1}{2}$.

Let us use Riemann's formula to prove the result stated above, namely that for $\mathrm{Re}(s) < 0$ the only zeros of $\zeta(s)$ lie at the negative even integers, $s = -2, -4\ldots$. To do this, we need only observe that taking $\mathrm{Re}(s) > 1$ in (1.355), the functions making up the left-hand side are non-singular. Furthermore, in this region the left-hand side is non-zero except at the zeros of $\cos(\frac{1}{2}\pi\, s)$. (Since $\Gamma(s)$ and $\zeta(s)$ are both, from their definitions, clearly non-vanishing in this region.) In this region, the zeros of $\cos(\frac{1}{2}\pi\, s)$ occur at $s = 2n+1$, where n is an integer with $n \ge 1$. They are simple zeros. Thus in this region the right-hand side of (1.355) has simple zeros at $s = 2n+1$. In other words, $\zeta(s)$ has simple zeros at $s = -2, -4, -6, \ldots$, and no other zeros when $\mathrm{Re}(s) < 0$.

Combined with the observation that the original series definition (1.318) makes clear that $\zeta(s)$ cannot vanish for $\mathrm{Re}(s) > 1$, we arrive at the conclusion that any possible additional zeros of $\zeta(s)$ must lie in the strip with $0 \le \mathrm{Re}(s) \le 1$. Riemann's formula does not help us in this strip, since it reflects it back onto itself. It is known that there are infinitely many zeros along the line $\mathrm{Re}(s) = \frac{1}{2}$. As we mentioned before, the still-unproven *Riemann Hypothesis* asserts that there are no zeros in this strip except along $\mathrm{Re}(s) = \frac{1}{2}$.