ANTIDERIVATIVES FOR COMPLEX FUNCTIONS

Marc Culler - Spring 2010

1. Derivatives

Most of the antiderivatives that we construct for real-valued functions in Calculus II are found by trial and error. That is, whenever a function appears as the result of differentiating, we turn the equation around to obtain an antiderivative formula.

Since we have the same product rule, quotient rule, sum rule, chain rule etc. available to us for differentiating complex functions, we already know many antiderivatives.

For example, by differentiating $f(z) = z^n$ one obtains $f'(z) = nz^{n-1}$, and from this one sees that the antiderivative of z^n is $\frac{1}{n+1}z^{n+1}$ – *except for the very important case where* $n = -1$. Of course that special case is very important in real analysis as it leads to the natural logarithm function. It will turn out to be very important in complex analysis as well, and will lead us to the complex logarithm. But this foreshadows some major differences from the real case, due to the fact that the complex exponential is not an injective function and therefore does not have an inverse function.

Besides the differentiation formulas, we have another tool available as well – the Cauchy Riemann equations.

Theorem 1.1. *Suppose $f(z) = u(z) + iv(z)$, where $u(z)$ and $v(z)$ are real-valued functions of the complex variable z.*

- *If $f'(z)$ exists then the partial derivatives of u and v exist at the point z and satisfy $u_x(z) = v_y(z)$ and $u_y(z) = -v_x(z)$.*
- *If u and v are defined and have continuous partial derivatives in an open set Ω which satisfy $u_x = v_y$ and $u_y = -v_x$, then f is analytic in Ω and $f'(z) = a(z) + ib(z)$ where a and b are the real-valued functions defined on Ω such that*

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix}.$$

In particular, in this case we have

$$f' = u_x - iu_y = u_x + iv_x = v_y - iu_y = v_y + iv_x.$$

Date: October 20, 2010.

We can use this to check that the function $\exp(z)$ is equal to its derivative. Since we have $\exp(x+iy) = e^x \cos y + ie^x \sin y$, we take $u(x+iy) = e^x \cos y$ and $v(x+iy) = e^x \sin y$ so, as expected, we have

$$\exp'(x+iy) = u_x - iu_y = e^x \cos y + e^x \sin y = \exp(x+iy).$$

2. Antiderivatives

In Calculus III we learned how to find a function with specified partial derivatives, or determine that such a function does not exist. We can apply this idea to find antiderivatives of complex functions. If we are given a complex function $f(z) = u(z) + iv(z)$ then we could attempt to find real-valued functions U and V such that $U_x = u$, $U_y = -v$, $V_x = v$ and $V_y = u$. Then, according to Theorem 1.1, the function $F(z) = U(z) + iV(z)$ would be an antiderivative for f.

Should we expect to be able to find functions U and V like that? We know from Calculus III that this is not always possible. Suppose we are given a vector field

$$\mathbf{V} = P(x,y)\mathbf{e}_1 + Q(x,y)\mathbf{e}_2$$

(where P and Q have continuous partials). We learned in Calculus III that there cannot exist a function $f(x,y)$ with $\nabla f = \mathbf{V}$ *unless* $P_y = Q_x$. The reason is that if such a function f did exist then we would know $P_y = f_{yx} = f_{xy} = Q_x$. (Recall that we have to know that f has continuous second partials to conclude that the mixed partials are equal.)

Assume that $f(z) = u(z)+iv(z)$ is an analytic function such that u and v have continuous partial derivatives. (Remember that this continuity is an extra assumption. We have not shown that the real and imaginary parts of an analytic function necessarily have continuous partial derivatives.) What does our easy test tell us? We are looking for functions functions U and V such that $U_x = u$, $U_y = -v$, $V_x = v$ and $V_y = u$. If these functions existed then their mixed partials would be equal, so the easy test says we need to check that $-v_x = U_{xy} = U_{yx} = u_y$, and $u_x = V_{xy} = V_{yx} = v_y$. But this is exactly what the Cauchy-Riemann equations guarantee! So the easy test will not cause us any trouble (at least not when u and v have continuous partials). This might encourage us to imagine that *any* analytic function might have an antiderivative. After all, this does happen in Calculus II: if $f(x)$ is continous on the interval $[a,b]$ then $F(x) = \int_a^x f(t)dt$ satisfies $F'(x) = f(x)$ for $x \in [a,b]$. Could something similar be true for analytic functions? The answer turns out to be surprising.

Example 2.1. Let's use this idea to find an antiderivative for the function $f(z) = z^2$.

We write $f(x+iy) = (x^2 - y^2) + i(2xy)$. We are looking for a function $F(x+iy) = U(x+iy) + iV(x+iy)$ such that $F'(z) = f(z)$. So, by the Cauchy-Riemann equations, we need $U_x = x^2 - y^2$ and $U_y = -2xy$, or $\nabla U = (x^2 + y^2)\mathbf{e}_1 - 2xy\mathbf{e}_2$.

Recall from Calculus III that the "Fundamental Theorem of Calculus for vector fields" tells us how to find U, if it exists. We arbitrarily decide that $U(0) = 0$ (since U is only determined up to a constant anyway).

If we choose a piecewise smooth path γ starting at 0 and ending at the point $a+ib$ then the fundamental theorem says:

$$\int_\gamma (x^2-y^2)dx - 2xydy = \int_\gamma \nabla U \cdot ds = U(a+ib) - U(0) = U(a+ib).$$

Let's choose the path $\gamma(t) = ta\mathbf{e}_1 + tb\mathbf{e}_2$, $t \in [0,1]$. Using this path we have

$$\int_\gamma (x^2-y^2)dx - 2xydy = \int_0^1 (a^2-b^2)t^2 adt - 2abt^2 bdt = \frac{1}{3}a^3 - ab^2.$$

In other words, $U(x,y) = \frac{1}{3}x^3 - xy^2$.

A very similar calculation shows that $V(x,y) = x^2y - \frac{1}{3}y^3$. Thus we have

$$F(x+iy) = \frac{x^3}{3} + ix^2y - xy^2 - i\frac{y^3}{3} = \frac{z^3}{3}.$$

It works!

We summarize this discussion in the following proposition.

Proposition 2.2. *Suppose that $f(z) = u(z) + iv(z)$ is an analytic function in an open set Ω. Let $z_0 \in \Omega$. Assume there exists a function $F(z)$, analytic on Ω, such that $F'(z) = f(z)$ and $F(z_0) = 0$. If γ is any piecewise smooth path in Ω starting at z_o and ending at z then we have*

$$F(z) = \int_\gamma udx - vdy + i\int_\gamma vdx + udy$$

3. The function $f(z) = \frac{1}{z}$

Now for the first part of the surprise. It turns out to be easy to see, by looking at the function $f(z) = 1/z$, that it cannot be true that every function which is analytic on an open set Ω has an antiderivative defined on all of Ω.

Proposition 3.1. *There is no function $F(z)$ which is analytic on $\mathbb{C} - \{0\}$ and is an antiderivative of $\frac{1}{z}$.*

Proof. Assume that F were such a function. By subtracting a constant from F we can assume that $F(1) = 0$.

We know that every number $z \in \mathbb{C} - \{0\}$ can be written as $z = \exp(w)$ for some $w \in \mathbb{C}$. Now consider the composite function $F \circ \exp$. By the chain rule we have

$$(F \circ \exp)'(w) = F'(\exp(w))\exp'(w) = \frac{1}{\exp w}\exp(w) = 1.$$

We also know that two analytic functions with the same derivative differ by a constant. Thus we see that $F \circ \exp(w) = w + c$ for some constant c, and in fact we must have $c = 0$ since $c = F(\exp(0)) = F(1) = 0$.

But this is certainly impossible, since it implies that

$$2\pi i = F \circ \exp(2\pi i) = F(1) = 0.$$

In other words, if such a function F existed then it would be an inverse function to the function exp, which does not have an inverse since it is not injective. $\square$

Is this a paradox? What goes wrong if we try to use line integrals to find an antiderivative for $1/z$, as we did for the function z^2?

If we define $f(z) = 1/z$ then we have

$$f(x + iy) = \frac{1}{x + iy} = \frac{x}{x^2 - y^2} - \frac{iy}{x^2 - y^2}.$$

If f had an antiderivative $F(z) = U(z) + iV(z)$ then the imaginary part $V(z)$ would satisfy

$$V_x = \frac{-y}{x^2 + y^2} \text{ and } V_x = \frac{x}{x^2 + y^2},$$

or equivalently,

$$\nabla V = \frac{-y}{x^2 + y^2}\mathbf{e}_1 + \frac{x}{x^2 + y^2}\mathbf{e}_2.$$

Let's try using our line integral method to find a formula for $V(z)$ (assuming such a function exists). We can't start our paths at 0 since it is not in the domain of f. So we will start at 1. We may as well require that $V(1) = 0$.

Proposition 2.2 says that if V exists then to find $V(a + ib)$ we can choose *any* path γ from 1 to $a + ib$ in Ω and we will have

$$V(a + ib) = \int_\gamma \frac{-y\,dx}{x^2 + y^2} + \frac{x\,dy}{x^2 + y^2}.$$

What happens if we evaluate this integral using two different paths from 1 to -1 in Ω, such as the upper and lower halves of the unit circle? Let's set $\gamma_+(t) = \cos t + i \sin t$ and $\gamma_-(t) = \cos t - i \sin t$. Then we have

$$\int_{\gamma_+} \frac{-y\,dx}{x^2 + y^2} + \frac{x\,dy}{x^2 + y^2} = \int_0^\pi (-\sin t)(-\sin t) + (\cos t)(\cos t)\,dt = \int_0^\pi dt = \pi,$$

and

$$\int_{\gamma_-} \frac{-y\,dx}{x^2 + y^2} + \frac{x\,dy}{x^2 + y^2} = \int_0^\pi (\sin t)(-\sin t) + (\cos t)(-\cos t)\,dt = -\int_0^\pi dt = -\pi.$$

So choosing these two different paths we get two different values for the integral. It is not a coincidence that the two values differ by a multiple of 2π. In fact, you should be able to convince yourself that the "value" of $F(-1)$ obtained this way could be any number of the form $i(\pi + 2n\pi)$ for an integer n. That is, the numbers that arise from computing integrals like this, using various choices of path, include every number z such that $\exp(z) = -1$.

4. Contour integrals

It should be clear by now that line integrals will play an important role in computing antiderivatives of complex functions, in those cases where an antiderivative exists. Hopefully it is also clear that we would rather not be rewriting these integrals in terms of real and imaginary parts. So it is time to introduce a complex version of the line integral from Calculus III. The complex version is called a *contour integral*.

First we introduce some terminology.

Definition 4.1. By a *path* we mean a continuous function $\gamma : [a, b] \to \mathbb{C}$, where $[a, b]$ is an interval in $\mathbb{R}$. If γ is a path we may write $\gamma(t) = x(t) + iy(t)$ where x and y are continuous functions of $t \in [a, b]$. If $x'(t)$ and $y'(t)$ exist for some t in $[a, b]$, then we say that $\gamma'(t)$ exists and we define $\gamma'(t) = x(t) + iy(t)$. We say that γ is *smooth* if γ' is defined and continuous on the entire interval $[a, b]$ and $\gamma'(t) \neq 0$ for $t \in [a, b]$. (Here we are taking $\gamma'(a)$ to mean the right derivative at a and $\gamma'(b)$ to mean the left derivative at b.) We say that γ is a *contour* if $\gamma : [a, b] \to \mathbb{C}$ is *piecewise smooth* in the sense that there are numbers $x_0 = a < x_1 < \cdots x_n = b$ such that each for $i = 1, \ldots n$ the restriction of γ to the interval $[x_{i-1}, x_n]$ is smooth.

If $\gamma(a) = \gamma(b)$ we will say that γ is a *closed contour*. Otherwise we say that it starts at $\gamma(a)$ and ends at $\gamma(b)$.

Of course this is nothing new. Viewing $\mathbb{C}$ as a 2-dimensional vector space over $\mathbb{R}$, with standard basis $(1, i)$, we see that contours correspond exactly to the types of paths that we used for line integrals in Calculus III. The complex number $\gamma'(t)$, viewed as a vector, is just the velocity of the path γ, and the real number $|\gamma'(t)|$ is its speed. Recall that we can compute the arclength of a path by integrating the speed of a parametrization *provided that the speed is never* 0. (If the speed is allowed to be 0 then the parametrization might stop and retrace portions of the path, making the integral of speed larger than the arclength.) This is the main reason why we require $\gamma'(t) \neq 0$ for a smooth path. Another reason is that a curve with corners could be given a smooth parametrization if the velocity were allowed to be 0 at the corners, and we would rather think of a corner as being a non-smooth point on the curve.

Definition 4.2. We will write $\text{Length}(C)$ to denote the arclength of a contour C. Thus, if C is smooth, with smooth parametrization $\gamma : [a, b] \to \mathbb{C}$ then

$$\text{Length}(C) = \int_a^b |\gamma'(t)|\, dt.$$

If C is piecewise smooth then Length(C) is the sum of the lengths of the smooth paths that make up C.

To integrate a continuous complex-valued function of a real variable, we just integrate the real and imaginary parts separately. (Such a function could be viewed as a continuous path, but we are not thinking of it that way at the moment.)

Definition 4.3. Suppose $g : [a, b] \to \mathbb{C}$ is a continuous complex-valued function of a real variable. Write $g(t) = x(t) + iy(t)$ where x and y are real-valued functions. Then

$$\int_a^b g(t)dt = \int_a^b x(t)dt + i\int_a^b y(t)dt.$$

Recall from Calculus that the value of this integral does not change if the path is reparametrized (as long as the direction of travel remains the same; reversing the direction changes the sign). For this reason, people often specify a contour by just describing the curve in the plane that the path follows, and the direction along that curve. You may then choose any parametrization you wish in order to compute the integral.

Definition 4.4. Suppose that $\gamma : [a, b] \to \mathbb{C}$ is a smooth path, and that $f(z)$ is a continuous function defined on an open set which contains the image curve of γ. Then we define

$$\int_\gamma f(z)dz = \int_a^b f(\gamma(t))\gamma'(t)dt.$$

If γ is a contour, so there are numbers $x_0 = a < x_1 < \cdots x_n = b$ such that each for $i = 1, \ldots n$ the restriction of γ to the interval $[x_{i-1}, x_n]$ is smooth, then we define

$$\int_\gamma f(z)dz = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(\gamma(t))\gamma'(t)dt.$$

Of course, if $f(z) = u(z) + iv(z)$, then

$$\int_\gamma f(z)dz = \int_\gamma udx - vdy + i\int_\gamma vdx + udy.$$

This means that we can reformulate Proposition 2.2 in a way which makes it clear that it is really the fundamental theorem of calculus. We do this below, and give a self-contained proof, which involves nothing more than the following chain rule for a composition of a complex functions with a smooth path.

Proposition 4.5. *Suppose that $f(z)$ is analytic in an open set Ω and that $\gamma : [a, b] \to \Omega$ is a smooth path. Then*

$$(f \circ \gamma)'(t) = f'(\gamma(t))\gamma'(t).$$

Proof. Express f and γ in terms of real and imaginary parts as $f(z) = u(z) + iv(z)$ and $\gamma(t) = x(t) + iy(t)$. We will apply the chain rule from real 2-variable calculus to u and v. We obtain

$$(u \circ \gamma)'(t) = u_x(\gamma(t))x'(t) + u_y(\gamma(t))y'(t)$$

and

$$(v \circ \gamma)'(t) = v_x(\gamma(t))x'(t) + v_y(\gamma(t))y'(t)$$

But, since $f'(z) = u_x(z) - iu_y(z) = v_y(z) + iv_x(z)$, we have

$$f'(\gamma(t))\gamma'(t) = \Big(u_x(\gamma(t))x'(t) + u_y(\gamma(t))y'(t)\Big) + i\Big(v_x(\gamma(t))x'(t) + v_y(\gamma(t))y'(t)\Big),$$

where we used the first form of $f'(z)$ to expand the real part and the second form to expand the imaginary part. □

Theorem 4.6 (Fundamental Theorem of Calculus)**.** *Suppose that $f(z)$ is analytic in an open set Ω and that $\gamma : [a, b] \to \Omega$ is a contour from $A = \gamma(a)$ to $B = \gamma(b)$. Then*

$$\int_\gamma f'(z)dz = f(B) - f(A).$$

Proof. In the case where γ is a smooth path we apply the ordinary fundamental theorem of calculus to the real and imaginary parts of $f \circ \gamma$, obtaining

$$\int_\gamma f'(z)dz = \int_a^b f'(\gamma(t))\gamma'(t)dt = \int_a^b (f \circ \gamma)'(t)dt = f(\gamma(b)) - f(\gamma(a)) = f(B) - f(A).$$

In the more general case of a contour, the right hand side will be a telescoping sum:

$$F(X_n) - F(X_{n-1}) + F(X_{n-1}) - F(X_{n-2}) + \cdots + F(X_1) - F(X_0) = F(X_n) - F(X_0),$$

where $X_k = \gamma(x_k)$. The sum collapses to again give $F(B) - F(A)$. □

5. Estimating contour integrals

The following basic estimate for the size of a contour integral will be used repeatedly.

Proposition 5.1. *Suppose that $f(z)$ is analytic in an open set Ω and that C is a contour in Ω. If $f(w)| < M$ for each point w lying on the contour C then*

$$\left|\int_C f(z)dz\right| \leq M\,\text{Length}(C).$$

The Proposition follows easily from the following lemma, which is an integral version of the triangle inequality.

Lemma 5.2. *If $v : [a, b] \to \mathbb{C}$ is a continuous function then*

$$\left|\int_a^b v(t)\,dt\right| \leq \int_a^b |v(t)|dt.$$

Remark 5.3. We can think of v as the velocity of a path in the complex plane. Since continuous functions of a real variable have antiderivatives, we may write $v(t) = \gamma'(t)$ for some path γ. Then the left side of the conclusion of the Lemma is the arclength of γ, while the right side is the displacement from the first point to the last point. Thus the lemma says that "the shortest path between two points is a straight line." We do need to supply a proof, though. Actually, we will give two.

First proof of 5.2. We consider Riemann sums for the two integrals. Let $\epsilon > 0$ and suppose that $a = t_0 < t_1 < \cdots < t_n = b$ is a partition such that both integrals are approximated within ϵ by their right sums based on this partition. Then

$$\left|\int_a^b v(t)\,dt\right| - \epsilon \le |\sum_{i=1}^n v(t_i)\Delta t_i| \le \sum_{i=1}^n |v(t_i)|\Delta t_i \le \int_a^b |v(t)|dt + \epsilon.$$

Since ϵ was arbitrary, this implies the lemma. □

Second proof of 5.2. Here we reduce the proof to the analogous result for real-valued functions (which can be proved using Riemann sums and the triangle inequality as above). Let $\int_a^b v(t)dt = re^{i\theta}$. Then $\int_a^b e^{-i\theta} v(t)dt = r$. Since the latter integral is real, it is equal to the integral of the real part of the integrand. Thus we have

$$\left|\int_a^b v(t)\,dt\right| = r = \left|\int_a^b \Re(e^{-i\theta}v(t))dt\right| \le \int_a^b |\Re(e^{-i\theta}v(t))|dt$$
$$\le \int_a^b |e^{-i\theta}v(t)|dt = \int_a^b |v(t)|dt.$$

□

Proof of Proposition 5.1. We just apply Lemma 5.2 to the definition of the contour integral. Choose a parametrization $\gamma : [a, b] \to \mathbb{C}$ for C. Then

$$\left|\int_C f(z)dz\right| = \left|\int_a^b f(\gamma(t))\gamma'(t)dt\right| \le \int_a^b |f(\gamma(t))||\gamma'(t)|dt$$
$$\le M \int_a^b |\gamma'(t)dt| = M\,\text{Length}(C).$$

□

6. Path independence

So far we have seen that an analytic function $f(z)$ on an open set Ω may fail to have an antiderivative defined everywhere in Ω, but may have an antiderivative defined on some smaller open set. The basic example is $f(z) = \frac{1}{z}$, which does not have an antiderivative on $\mathbb{C} - \{0\}$, but does have an antiderivative, for example, on the set $\{z \in \mathbb{C} \mid \Re z > 0\}$.

We have also seen, from the Fundamental Theorem of Calculus, that if $f(z)$ does have an antiderivative on the domain Ω then it can be computed using contour integrals. Given a point $w_0 \in \Omega$, if f has an antiderivative on Ω then there is a unique antiderivative F on Ω such that $F(w_0) = 0$, and we have

$$F(w) = \int_C f(z)dz$$

where C is *any* contour from w_0 to w. The next theorem shows that the property that the contour integral above does not depend on the choice of contour characterizes the functions which have antiderivatives.

Definition 6.1. We will call a contour L is a *polygonal path* if it is made of finitely many parametrized line segments, and a *zigzag* if all of these segments are either horizontal or vertical. That is, L is a zigzag if it can be parametrized as $\gamma(t) = x(t) + iy(t)$ where either $x'(t) = 0$ or $y'(t) = 0$ for any t at which $\gamma'(t)$ exists and is non-zero (and this occurs at all but finitely many values of t).

Definition 6.2. We say that an open set Ω is a *domain* if for any $z_1, z_2 \in \Omega$ there is a polygonal path from z_1 to z_2 contained in Ω.

Theorem 6.3. *Let $f(z)$ be analytic on a domain Ω. The following are equivalent*

(1) $f(z)$ has an antiderivative defined on all of Ω;

(2) $\oint_C f(z)dz = 0$ for any closed contour C in Ω;

(3) $\oint_P f(z)dz = 0$ for any closed polygon P in Ω.

(4) $\oint_L f(z)dz = 0$ for any closed zigzag L in Ω.

Remark 6.4. Condition (2) is equivalent to saying that $\int_C f(z)dz = \int_{C'} f(z)dz$ whenever C and C' are contours in Ω with the same starting and ending points. Conditions (3) and (4) say the same thing, but for more special types of contours.

Proof of Theorem 6.3. We have (1) $\Rightarrow$ (2) by the Fundamental Theorem of Calculus, and (2) $\Rightarrow$ (3) since every polygon is a contour, and (3) implies (4) since every zigzag is a polygon. So it remains to show that (4) $\Rightarrow$ (1).

We assume that condition (4) holds. Let w_0 be an arbitrary point in the domain Ω. We will construct a function F defined on Ω such that $F(w_0) = 0$ and $F'(w) = f(w)$.

By Remark 6.4 it makes sense to define

$$F(w) = \int_L f(z)dz$$

where L is any zigzag from w_0 to w. (This is well-defined because different choices for L will produce the same value for $F(w)$. Moreover, the Fundamental Theorem of Calculus guarantees that this will give a formula for F if such a function F exists.) We must show that

$$\lim_{h\to 0}\left|\frac{F(w+h)-F(w)}{h}-f(w)\right|=0.$$

Let ϵ be given. Since Ω is open and $f'(w)$ exists, there exists $\delta > 0$ such that

- the disk $B(w,\delta)=\{z \mid |z-w|<\delta\}$ is contained in Ω; and
- $|h|<\delta \Rightarrow \left|\dfrac{f(z+h)-f(z)}{h}-f'(w)\right|<\dfrac{\epsilon}{2}$.

Suppose that $0<|h|<\delta$ and and let K be a zigzag from z_0 to w. There is a zigzag L from w to $w+h$ which is contained in the disk $B(w,h)$, for example consisting of one horizontal and one vertical line segment. Since

$$F(z+h)=\int_K f(z)dz+\int_L f(z)dz$$

we have $F(z+h)-F(z)=\int_L f(z)dz$. By the Fundamental Theorem of Calculus, since the constant function 1 has antiderivative z, we know that

$$\int_L f(w)dz=f(w)\int_L dz=f(w)h.$$

Since $|f(z)-f(w)|<\frac{\epsilon}{2}$ for z on the zigzag L, and the zigzag L has length $<2|h|$, we have

$$\begin{aligned}\left|\frac{F(z+h)-F(z)}{h}-f(w)\right| &= \left|\frac{1}{h}\int_L f(z)dz-\frac{1}{h}\int_L f(w)dz\right|\\ &= \frac{1}{|h|}\left|\int_L (f(z)-f(w))dz\right|\\ &\le \frac{\epsilon\,\mathrm{Length}(L)}{2|h|}\le \epsilon\end{aligned}$$

Since $\epsilon>0$ was arbitrary, this shows that

$$\lim_{h\to 0}\left|\frac{F(w+h)-F(w)}{h}-f(w)\right|=0,$$

and hence that $F'(w)=f(w)$ for all $w\in\Omega$. Thus (4) $\Rightarrow$ (1), and the proof is complete. □

7. A topological interlude

Now we will take a break from analysis and introduce the basic concepts and theorems about the topology of $\mathbb{R}$ and $\mathbb{C}$ that we will need later on. It turns out that everything we have to say here is valid for $\mathbb{R}^m$ for any $m>0$, so we will give the statements in this general form. Of course we know that $\mathbb{C}$ is the same as $\mathbb{R}^2$, topologically speaking, if we identify (x,y) with $x+iy$.

This discussion starts at the very beginning, so parts of this section will be review. First we recall a crucial property of the real number system, often taken as an axiom in real analysis classes.

Definition 7.1. A set $X \subset \mathbb{R}$ is said to be *bounded above* if there is a number $M \in \mathbb{R}$ such that $x \leq M$ for all $x \in X$.

7.2 (Least Upper Bound Axiom)**.** *If a nonempty set $X \subset \mathbb{R}$ is bounded above then there exists a number $\sup X$ such that*

- *$x \leq \sup X$ for all $x \in X$; and*
- *if $x \leq M$ for all $x \in X$ then $\sup X \leq M$.*

The number $\sup x$ is called the least upper bound for X.

If $X \subset \mathbb{R}$ is nonempty and bounded below then, applying the least upper bound property to $\{-x \mid x \in X\}$ it is easy to see that X has a greatest lower bound $\inf X$. Moreover, if $X \subset \mathbb{R}$ and $Y \subset \mathbb{R}$ are both nonempty and $x \leq y$ for all $x \in X$ and all $y \in Y$, then $\sup X \leq \inf Y$.

Notation 7.3. We will use the following notation.

- If $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ then $|x| = \sqrt{x_1^2 + \cdots x_m^2}$, so $|x - y|$ is the the distance from x to y.
- For $x \in \mathbb{R}^m$, the *ball of radius r about x* is the set $B_x(r) = \{y \in \mathbb{R}^m \mid |y - x| < r\}$.

Definition 7.4. We say a sequence (x_n) in $\mathbb{R}^m$ *converges* to x when

- for every $\epsilon > 0$ there exists an integer N such that $|x - x_n| < \epsilon$ for all $n \geq N$.

In other words, we say $x_n \to x_\infty$ as $n \to \infty$ if the ball $B_{x_\infty}(\epsilon)$ contains a tail of the sequence (x_n) for every $\epsilon > 0$.

Definition 7.5. By a *box* in $\mathbb{R}^m$ we mean a set of the form $[a_1, b_1] \times \cdots \times [a_m, b_m]$. Thus a box in $\mathbb{R}$ is a closed interval and a box in $\mathbb{R}^2$ is a rectangle. Two boxes $[a_1, b_1] \times \cdots \times [a_m, b_m]$ and $[c_1, d_1] \times \cdots \times [d_m, d_m]$ in $\mathbb{R}^m$ are *similar* if there is a scale factor α such that $|b_i - a_i| = \alpha|d_i - c_i|$ for $i = 1, \ldots m$. The *diameter* of the box $B = [a_1, b_1] \times \cdots \times [a_m, b_m]$ is the number $\mathrm{Diam}(B) = |(b_1 - a_1, \ldots, b_n - a_n)|$. (Note that if x and y are two points in B then $|x - y| \leq \mathrm{Diam}(B)$.) By the *interior* of the box $[a_1, b_1] \times \cdots \times [a_m, b_m]$ we will mean the open set $(a_1, b_1) \times \cdots \times (a_m, b_m)$, which might be empty.

Proposition 7.6. *A nested sequence of boxes has a nonempty intersection. That is, if (B_n) is a sequence of boxes in $\mathbb{R}^m$ with $B_1 \supset B_2 \supset \cdots$ then $\bigcap_{n=1}^{\infty} B_n$ is a nonempty box.*

Proof. First we prove this for $\mathbb{R} = \mathbb{R}^1$. Here each B_n is a closed interval; let $B_n = [a_n, b_n]$. Then the nestedness condition translates to $a_1 \leq a_2 \leq \cdots$ and $b_1 \geq b_2 \geq \cdots$. Let $a = \sup\{a_n \mid n \in \mathbb{N}\}$ and $b = \inf\{b_n \mid n \in \mathbb{N}\}$. Then we have $a_n \leq a \leq b \leq b_n$ for all $n \in \mathbb{N}$, and hence $[a, b] \subset [a_n, b_n]$ for all n. Moreover, if $x \in \bigcap_1^\infty [a_n, b_n]$ then $x \geq a$ and $x \leq b$. Thus $\bigcap_{n=1}^\infty [a_n, b_n] = [a, b]$, which is a nonempty interval (possibly just a point).

The argument for $\mathbb{R}^m$ is the same for all $m > 1$. To keep the notation simple we will just give it for $m = 2$. Suppose that $B_n = [a_n, b_n] \times [c_n, d_n]$. Then $([a_n, b_n])$ and $([c_n, d_n])$ are nested sequences of intervals. Let $\bigcap_{n=1}^\infty [a_n, b_n] = [a, b]$ and $\bigcap_{n=1}^\infty [c_n, d_n] = [c, d]$. It is easy to check that $\bigcap_{n=1}^\infty B_n = [a, b] \times [c, d]$, which is nonempty. □

Definition 7.7. A set $U \subset \mathbb{R}^m$ is said to be *open* if it has the following property:

- For every point x of U there exists $\epsilon > 0$ such that $B_x(\epsilon) \subset U$.

A set $C \subset \mathbb{R}^m$ is said to be *closed* if its complement $\mathbb{R}^n - C$ is open. Notice that $\emptyset$ and $\mathbb{R}^m$ are both open and closed.

Lemma 7.8. *Suppose $X \subset \mathbb{R}^m$ and $f : X \to \mathbb{R}^n$ is continuous. Let U be an open set in $\mathbb{R}^n$. Then there exists an open set ${}^fU$ in $\mathbb{R}^m$ such that $f^{-1}(U) = X \cap {}^fU$.*

Proof. First fix $x \in X$. There exists $\epsilon > 0$ such that $B_\epsilon(f(x)) \subset U$. By the definition of continuity there exists $\delta > 0$ such that $f(X \cap B_\delta(x)) \subset B_\epsilon(f(x)) \subset U$. Define $B_x = B_\delta(x)$. We then have $B_x \cap X \subset f^{-1}(U)$.

Now define ${}^fU = \bigcup\{B_x \mid x \in f^{-1}(U)\}$. We have

$$f^{-1}(U) \subset X \cap {}^fU = X \cap \bigcup_{x \in f^{-1}(U)} B_x = \bigcup_{x \in f^{-1}(U)} X \cap B_x \subset f^{-1}(U).$$

Thus $f^{-1}(U) = X \cap {}^fU$. □

Connectedness. At the beginning of the course we defined an open set to be connected if any two points of the set can be joined by a polygonal path. Of course this definition does not make much sense for sets which are not open. For example, the circle does not meet this condition.

Now we will replace our older definition with a more abstract one that can be used for arbitrary subsets of $\mathbb{R}^m$. Eventually we will see that the two definitions are equivalent for open sets.

From now on, we use the following.

Definition 7.9. A subset $X \subset \mathbb{R}^m$ is *disconnected* if there exist two open sets U and V such that

- $X \subset U \cup V$,

- $X \cap U \neq \emptyset$,
- $X \cap V \neq \emptyset$,
- $X \cap U \cap V = \emptyset$.

The pair of sets U and V are called a *separation* of X. If X is not disconnected then we say X is *connected*. That is, X is connected if it has no separation.

It is easy to check that an open set is disconnected if and only if it can be written as the union of two disjoint nonempty open subsets.

The next proposition shows that an open set which satisfies our new definition also satisfies the old one. As a point of logic, notice that this proposition does not show that the definitions are the same. In fact, it allows the possibility that no open sets are connected under the new definition. (Of course this is not true, but we haven't ruled it out yet.)

Proposition 7.10. *If O is a connected open set in $\mathbb{R}^m$ then any two points of O can be joined by a polygonal path.*

Proof. We may as well assume that O is nonempty since there is nothing to prove if $O = \emptyset$. Let $x \in O$. Define U to be the collection of all points of O which can be joined to x by a polygonal path. If $y \in U$ then $B_\epsilon(y) \subset O$ for some $\epsilon > 0$. Every point z of $B_\epsilon(y)$ can be joined to y by a line segment L which is contained in $B_\epsilon(y)$. By adding L on the end of a polygonal path from $x$to y we get a polygonal path from x to z. This shows that $B_\epsilon(y) \subset U$, so U is open. Now define V to be the collection of all points of O which cannot be joined to x by a polygonal path, so $X = U \cup V$ and $U \cap V = \emptyset$. Suppose $y \in V$ and $B_\epsilon(y) \subset O$. If there exists a point z of $B_\epsilon(y)$ which can be joined to X by a polygonal path then we could append to this path a line segment from z to y inside $B_\epsilon(y)$. This would contradict $y \in V$. Therefore V is open. Now since O is connected, one of sets U and V must be empty, and we know that $x \in U$. Therefore we conclude that $V = \emptyset$, so $O = U$. This completes the proof. □

We should start by showing that there really are some sets which are connected with our new definition.

Proposition 7.11. *Any interval in $\mathbb{R}$ (open, closed, half-open, or infinite) is connected.*

Proof. Let I be an interval in $\mathbb{R}$ and suppose, for contradiction, that U and V form a separation of I. Let $x \in I \cap U$ and $y \in I \cap V$. We may assume $x < y$. (Otherwise rename U and V.) Thus $\{t \in I \mid [x, t] \subset U\}$ is bounded above, and we may define $z = \sup\{t \in I \mid [x, t] \subset U\}$.

If $z \in U$ then there exists $\epsilon > 0$ such that $B_\epsilon(z) \subset U$. Also, there exists $t \in (z - \epsilon, z]$ such that $[x, t] \subset U$. Thus $[x, z + \epsilon] = [x, t] \cup [t, z + \epsilon] \subset U$, which is a contradiction since $z = \sup\{t \in I \mid [x, t] \subset U\}$.

If $z \in V$ then there exists $\epsilon > 0$ such that $B_\epsilon(z) \subset V$. This implies that $\sup\{t \in I \mid [x, t] \subset U\} \leq z - \epsilon < z$, which is again a contradiction.

Since $z \in I \subset U \cup V$, we have a contradiction, and there can not be a separation of I. □

Proposition 7.12. *If $X \subset \mathbb{R}^m$ is connected and $f : X \to \mathbb{R}^n$ is continuous then $f(X)$ is connected.*

Proof. We will prove the contrapositive. Suppose that $f(X)$ is disconnected. Let U and V be open sets in $\mathbb{R}^n$ which form a separation of $f(X)$. According to Lemma 7.8 there are open sets ${}^fU$ and ${}^fV$ such that $f^{-1}(U) = X \cap {}^fU$ and $f^{-1}(V) = X \cap {}^fV$.

We have $X = f^{-1}(U) \cup f^{-1}(V) \subset {}^fU \cup {}^fV$. Since $\emptyset \neq f^{-1}(U) \subset {}^fU$ and $\emptyset \neq f^{-1}(V) \subset {}^fV$, we have $X \cap {}^fU \neq \emptyset$ and $X \cap {}^fV \neq \emptyset$. Finally, $X \cap {}^fU \cap {}^fV \subset f^{-1}(U) \cap f^{-1}(V) = \emptyset$. Thus ${}^fU$ and ${}^fV$ form a separation of X. □

Corollary 7.13. *A line segment in $\mathbb{R}^m$ is connected.*

Proposition 7.14. *Suppose that $\mathcal{C}$ is a collection of connected sets in $\mathbb{R}^m$ such that $\bigcap \mathcal{C} \neq \emptyset$. Then $\bigcup \mathcal{C}$ is connected.*

Proof. Let us write $X = \bigcup \mathcal{C}$ and choose $x \in \bigcap \mathcal{C}$. Suppose, for contradiction, that U and V form a separation of X. We may assume that $x \in U$.

Since $V \cap X \neq \emptyset$, there is an element $y \in V \cap X$. Since $X = \bigcup \mathcal{C}$, we have $y \in C$ for some $C \in \mathcal{C}$. We also have $x \in \bigcap \mathcal{C} \subset C$. Thus $x \in U \cap C$ and $y \in V \cap C$, which shows that $C \cap U$ and $C \cap V$ are nonempty.

Since U and V form a separation of X we also have $C \subset X \subset U \cup V$ and $C \cap U \cap V \subset X \cap U \cap V = \emptyset$. But this shows that U and V form a separation of C, which is a contradiction since C is connected. □

Corollary 7.15. *Balls and boxes are connected subsets of $\mathbb{R}^m$.*

Finally, the next corollary shows that our two definitions of connected open sets are equivalent.

Corollary 7.16. *If $X \subset \mathbb{R}^m$ is a subset such that any two points of X can be joined by a polygonal path, then X is connected.*

The next Proposition implies that any open set is a union of domains.

Proposition 7.17. *Let $X \subset \mathbb{R}^m$. Define a relation on X by $x \sim y$ if x and y are contained in a connected subset of X. Then $\sim$ is an equivalence relation and the equivalence classes are connected.*

Proof. The transitive property of $\sim$ is the only one which is not obvious. But this follows from Proposition 7.14: if $x \sim y$ and $y \sim z$ then there are connected subsets A and B of X such that $x, y \in A$ and $y, z \in B$. By Proposition 7.14, $A \cup B$ is a connected set containing all three points, so $x \sim z$. Since the equivalence class $[x]$ of $x \in X$ is the union of all connected subsets of X which contain x, it is connected by Proposition 7.14. $\square$

The equivalence classes for the relation above are called the *connected components* of X. They are disjoint connected sets whose union is X.

Corollary 7.18. *Any open set in $\mathbb{R}^n$ is a disjoint union of domains.*

Compactness. We start with two definitions.

Definition 7.19. Suppose $X \subset \mathbb{R}^n$. A collection $\mathcal{U}$ of open sets is called *an open cover* of X if $X \subset \bigcup \mathcal{U}$. If $\mathcal{U}$ is an open cover of X and $\mathcal{U}' \subset \mathcal{U}$ is also an open cover of X, then $\mathcal{U}'$ is called a *subcover* of $\mathcal{U}$.

Definition 7.20. A set $K \subset \mathbb{R}^m$ is said to be *compact* if every open cover of K has a finite subcover. That is, given any collection $\mathcal{U}$ with $K \subset \bigcup \mathcal{U}$ there are finitely many open sets $U_1, \cdots, U_n \in \mathcal{U}$ such that $K \subset \bigcup_{i=1}^n U_i$.

Proposition 7.21. *All compact sets in $\mathbb{R}^m$ are closed.*

Proof. Suppose K is compact and $y \in \mathbb{R}^m$. For every $x \in K$ set $r_x = |x - y|/2$ and $U_x = B_x(r_x)$. Then $\mathcal{U} = \{U_x \mid x \in K\}$ is an open cover of K such that $U_x \cap B_{r_x}(y) = \emptyset$ for all $U_x \in \mathcal{U}$. Since K is compact, there is a finite subcover $\{U_{x_1}, \ldots, U_{x_N}\}$. Set $r = \min\{r_{x_1}, \ldots, r_{x_N}\}$. Then $B_r(y) \cap U_{x_n} = \emptyset$ for $n = 1, \ldots, N$. Since $K \subset \bigcup_{n=1}^N U_{x_n}$ we have

$$B_r(y) \cap K \subset B_r(y) \cap \left(\bigcup_{n=1}^N U_{x_n} \right) \subset \bigcup_{n=1}^N (K \cap U_{x_n}) = \emptyset.$$

Thus there is a ball centered at y which is contained in the complement of K. This shows that $\mathbb{R}^m - K$ is open, so K is closed. $\square$

Definition 7.22. A set $X \subset \mathbb{R}^n$ is said to be *bounded* if $X \subset B_R(0)$ for some $R > 0$. A sequence (x_n) of points in $\mathbb{R}^n$ is said to be *bounded* if there is a ball $B_R(0)$ which contains each term x_n. Equivalently, X is bounded if X is contained in a box, and (x_n) is bounded if there is a box containing all of the terms x_n.

Proposition 7.23. *All compact sets in $\mathbb{R}^m$ are bounded.*

Proof. Suppose $K \subset \mathbb{R}^m$ is compact. We will show that there is a positive integer R such that K is contained in the ball of radius R centered at 0. We let $\mathcal{U} = \{B_n(0) \mid n \in \mathbb{N}\}$. Since $\bigcup \mathcal{U} = \mathbb{R}^m$, $\mathcal{U}$ is certainly an open cover of K. Since K is compact there is a finite subcover $\{B_{n_1}(0), \dots B_{n_N}(0)\}$. Let $R = \max\{n_1, \dots, n_N\}$. Then we have

$$K \subset B_{n_1}(0) \cup \cdots \cup B_{n_N}(0) = B_R(0),$$

as claimed. □

Our next goal is to show that every closed and bounded subset of $\mathbb{R}^m$ is compact. For this we formulate a lemma about boxes that we will be able to use in several different contexts. The lemma encapsulates an argument known as the "method of bisection."

Suppose that B is a box in $\mathbb{R}^m$. By a *subdivision* of B we mean a collection of boxes $B_1, \dots B_N$ such that $B = B_1 \cup \cdots \cup B_N$ and such that the interior of B_i is disjoint from the interior of B_j whenever $i \neq j$. For any subdivision the volume of B is the sum of the volumes of the boxes B_n for $n = 1, \dots N$.

One simple method for subdividing a box is as follows. Suppose we are given $B = [a_1, b_1] \times \cdots [a_m, b_m]$. For each $n = 1, \dots, m$, let $c_n = (a_n + b_n)/2$ denote the midpoint of the interval $[a_m, b_m]$. Then there are 2^m boxes of the form $I_1 \times \cdots \times I_m$ where, for $n = 1, \dots, m$, I_n is either the interval $[a_n, c_n]$ or the interval $[c_n, b_n]$. Each of these smaller boxes is contained in B, each is similar to B, their union is equal to B, and the sum of their volumes is equal to the volume of B. We will say that this family of 2^m boxes is the *midpoint subdivision* of B.

Lemma 7.24 (Bisection Lemma)**.** *Let B be a box in $\mathbb{R}^m$. Suppose that $\mathcal{G}$ is a collection of boxes which are contained in B and are similar to B, with the following property:*

($\star$) *for any $x \in B$ there exists $\epsilon > 0$ such that if X is a box similar to B with $x \in X \subset B$ and $\mathrm{Diam}(X) < \epsilon$ then X is an element of $\mathcal{G}$.*

Then there is a subdivision of B into boxes which are all in the family $\mathcal{G}$.

Proof. Let's say that a box contained in B is good if it is in the collection $\mathcal{G}$, and bad otherwise.

We will construct a sequence $(\mathcal{S}_n)$ of subdivisions of B inductively. We start with $\mathcal{S}_1 = \{B\}$. For the inductive step, if $\mathcal{S}_n$ has been defined then we define $\mathcal{S}_{n+1}$ to be the subdivision obtained from $\mathcal{S}_n$ by replacing each bad box X in $\mathcal{S}_n$ by the 2^m boxes obtained by midpoint subdivision of X. Since midpoint subdivision of a box produces similar boxes, all of the boxes in each $\mathcal{S}_n$ are similar to B.

If, for some n, all of the boxes in $\mathcal{S}_n$ are good then the lemma is proved. Therefore we assume for a contradiction that each $\mathcal{S}_n$ contains a bad box.

Since any bad box in $\mathcal{S}_n$ must have been produced by midpoint subdivision of a bad box in $\mathcal{S}_{n-1}$, there is a sequence of bad boxes

$$B = X_1 \supset X_2 \supset X_3 \cdots$$

such that $\operatorname{Diam}(X_{n+1}) = \frac{1}{2}\operatorname{Diam}(X_n)$. In particular, this means that $\operatorname{Diam}(X_n) \to 0$ as $n \to \infty$.

Now Proposition 7.6 tells us that there is a point $x \in B$ which is contained in the intersection of all of the bad boxes X_n. By our hypothesis there exists $\epsilon > 0$ such that any box which contains x and has diameter less than ϵ is good. But we can easily find N so that $\operatorname{Diam}(X_N) = 2^{-N}\operatorname{Diam}(B) < \epsilon$. The box X_N must therefore be good, which is a contradiction.

The contradiction shows that for some n the subdivision $\mathcal{S}_n$ contains only good boxes, which completes the proof. ☐

Theorem 7.25 (Heine-Borel)**.** *A closed and bounded subset of $\mathbb{R}^m$ is compact.*

Proof. Let $K \subset \mathbb{R}^m$ be closed and bounded, and let $\mathcal{U}$ be an open cover of K.

Since K is bounded, there is a box B containing K. Let $\mathcal{G}$ be the collection of all boxes $X \subset B$ which are similar to G and satisfy $X \cap K \subset U$ for some $U \in \mathcal{U}$. We will check that condition $(\star)$ of the Bisection Lemma 7.24 holds for this family $\mathcal{G}$. Let $x \in B$ be given. There are two cases, according to whether $x \in K$ or not.

If $x \in K$ then, since $\mathcal{U}$ is an open cover of K, there exists $U \in \mathcal{U}$ which contains x. Since U is open, there exists $\epsilon > 0$ such that $B_\epsilon(x) \subset U$. If X is a box with $x \in X$ and $\operatorname{Diam}(X) < \epsilon$ then $|x - y| < \epsilon$ for all $y \in X$, so $X \cap K \subset B_\epsilon(x) \subset U$.

If $x \notin K$ then, since K is closed, there exists $\epsilon > 0$ such that $B_\epsilon(x) \cap K = \emptyset$. If X is a box with $x \in X$ and $\operatorname{Diam}(X) < \epsilon$ then $X \cap K \subset B_\epsilon(x) \cap K = \emptyset$, so $X \cap K$ is contained in every set U in $\mathcal{U}$.

This shows that $(\star)$ holds. Now by the Bisection Lemma we may subdivide B into boxes $X_1, \ldots, X_N$ where each X_n is in $\mathcal{G}$, which means that $K \cap X_n \subset U_n$ for some $U_n \in \mathcal{U}$. Thus

$$K = K \cap B = K \cap \left(\bigcup_{n=1}^{N} X_n \right) = \bigcup_{n=1}^{N} (K \cap X_n) \subset \bigcup_{n=1}^{N} U_n.$$

This shows that $\{U_1, \ldots, U_N\}$ is a finite subcover of $\mathcal{U}$. ☐

Next we will show that every bounded sequence of points in $\mathbb{R}^m$ has a subsequence which converges. The proof uses the following lemma.

Lemma 7.26. *Let (x_n) be a sequence of points of $\mathbb{R}^n$. Suppose $x \in \mathbb{R}^n$ has the following property:*

- *for every $\epsilon > 0$ there are infinitely many $n \in \mathbb{N}$ such that $x_n \in B_\epsilon(x)$.*

Then there is a subsequence (x_{n_k}) which converges to x.

Proof. We will define the subsequence by induction in such a way that $x_{n_k} \in B_{1/k}(x)$. Then it is clear that $x_{n_k} \to x$ as $k \to \infty$.

Since $B_1(x)$ contains x_n for infinitely many values of n, we may choose n_1 so $x_{n_1} \in B_1(x)$. For the induction step, suppose $x_{n_1}, \dots, x_{n_N}$ have been defined so that $x_{n_k} \in B_{1/k}(x)$ for $k = 1, \dots, N$. Since $B_{1/(N+1)}(x)$ contains x_n for infinitely many n we may choose n_{N+1} so that $n_{N+1} > n_N$ and $x_{n_{N+1}} \in B_{1/(N+1)}(x)$. This completes the proof. □

Theorem 7.27 (Bolzano-Weierstrass)**.** *Any bounded sequence of points in $\mathbb{R}^m$ has a convergent subsequence.*

Proof. Suppose for a contradiction that (x_n) is a bounded sequence with no convergent subsequence. Since (x_n) is bounded there is a box B containing all of the terms x_n. According to Lemma 7.26, for each $x \in B$ there exists $\epsilon > 0$ such that $B_\epsilon(x)$ contains x_n for at most a finite number of $n \in \mathbb{N}$. This means that if X is a box with $x \in X \subset B$ and $\mathrm{Diam}(X) < \epsilon$ then there are only finitely many $n \in \mathbb{N}$ such that $x_n \in X$. Thus condition $(\star)$ of the Bisection Lemma 7.24 holds if we define $\mathcal{G}$ to be the collection of all boxes $X \subset B$ such that X is similar to B and there are only finitely many $n \in \mathbb{N}$ such that $x_n \in X$. According to the Bisection Lemma we may therefore subdivide B into finitely many boxes in $\mathcal{G}$. Each of these boxes contains x_n for only finitely many $n \in \mathbb{N}$, yet their union is all of B, which contains x_n for all $n \in \mathbb{N}$. That absurd statement gives us the contradiction. □

Proposition 7.28. *If $K \subset \mathbb{R}^m$ is compact and $f : K \to \mathbb{R}^n$ is continuous then $f(K)$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $f(K)$. By Lemma 7.8, for each $U \in \mathcal{U}$ there exists an open set ${}^fU \subset \mathbb{R}^m$ such that $f^{-1}(U) = K \cap {}^fU$.

Consider the family of open sets ${}^f\mathcal{U} = \{{}^fU \mid U \in \mathcal{U}\}$. We have

$$K = f^{-1}(f(K)) = f^{-1}(\cup\,\mathcal{U}) = \bigcup_{U \in \mathcal{U}} f^{-1}(U) = \bigcup_{U \in \mathcal{U}} (K \cap {}^fU) \subset \bigcup_{U \in \mathcal{U}} {}^fU.$$

Thus ${}^f\mathcal{U}$ is an open cover of K. Since K is compact, there exist $U_1, \dots, U_N \in \mathcal{U}$ such that $\{{}^fU_1, \dots, {}^fU_N\}$ is a subcover of ${}^f\mathcal{U}$.

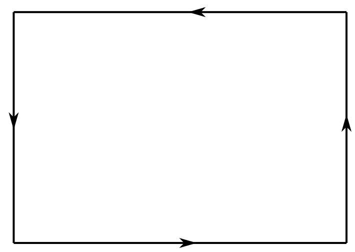
Now we have

$$f(K) = \bigcup_{i=1}^{N} f(K \cap {}^fU_i) = \bigcup_{i=1}^{N} f(f^{-1}(U_i) \subset \bigcup_{i=1}^{N} U_i.$$

Thus $\{U_1, \dots, U_N\}$ is a finite subcover of $\mathcal{U}$. □

8. Cauchy's Theorem for a rectangle

Next we will consider a contour integral around the simplest possible closed zigzag, namely the boundary of a rectangle. If R is a rectangle, we will denote by ∂R the zigzag consisting of the four segments on the boundary of R, oriented counter-clockwise.



Theorem 8.1 (Cauchy-Goursat)**.** *Suppose that $f(z)$ is analytic in a domain Ω containing a rectangle R. Then*

$$\int_{\partial R} f(z)dz = 0.$$

Proof. We begin by describing the overall strategy of our proof. Rather than show directly that $\int_{\partial R} f(z)dz = 0$, we will instead prove the equivalent statement that, for any $\epsilon > 0$,

$$(\star\star) \qquad \left|\int_{\partial R} f(z)dz\right| < \epsilon\,\text{Area}(R).$$

(Since Area(R) is a fixed number, this statement implies that the integral is arbitrarily close to 0, and hence cannot equal any number different from 0.) Suppose now that we are given $\epsilon > 0$.

To prove inequality $(\star\star)$ we will use the Bisection Lemma 7.24 to subdvide R into finitely many rectangles $R_1, \dots R_N$, all similar to R, such that the following three properties hold:

(1) $\displaystyle\sum_{i=1}^{N} \text{Area}(R_i) = \text{Area}(R)$;

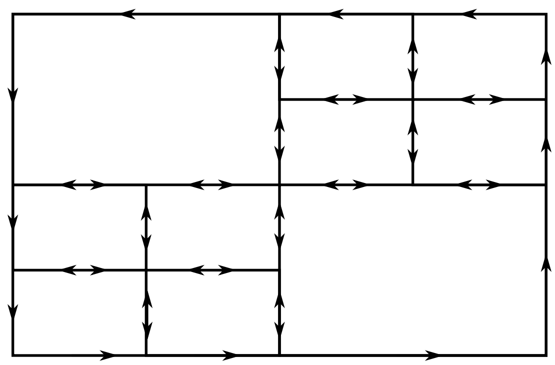
(2) $\displaystyle\sum_{i=1}^{N} \int_{\partial R_i} f(z)dz = \int_{\partial R} f(z)dz$;

(3) $\displaystyle\int_{\partial R_i} f(z)dz \leq \epsilon\,\text{Area}(R_i)$ for each $i = 1, \dots, N$.

It is clear that these three properties imply $(\star\star)$.

When we apply the Bisection Lemma we will take the family $\mathcal{G}$ of "good" rectangles to contain all rectangles $R' \subset R$, similar to R, such that $\int_{\partial R'} f(z)dz \leq \epsilon\,\text{Area}(R')$. Of course, to apply the lemma we must verify that property $(\star)$ in the statement of Lemma 7.24 holds for this family $\mathcal{G}$.

However, before checking property $(\star)$, we pause to observe that if it does hold then inequality $(\star\star)$ holds. In other words, we are asserting that Properties (1)-(3) must hold for the family of rectangles which are produced by the Bisection Lemma. Property (1) is clear for any subdivision of the rectangle R. Property (3) is also clear, given our definition of $\mathcal{G}$. But Property (2) requires some work. We may rewrite $\sum_{i=1}^{N} \int_{\partial R_i} f(z)dz$ as a sum of contour integrals whose contour runs along one side of one of the subrectangles. Observe that for each side which is contained in the interior of R, there are exactly two integrals in the sum, and the contours used for these two integrals run across that side in opposite directions. These two contour integrals therefore cancel. (In the figure below, each segment with two arrowheads corresponds to one such cancellation.) This leaves only the contours corresponding to line segments contained in the boundary of R. Since the union of these segments is the entire boundary of R, and since each of these contours is oriented consistently with the orientation of ∂R, this implies (2). Thus all that remains is to verify that $\mathcal{G}$ satisfies the property $(\star)$.



Next we will use Proposition 5.1 to to estimate $\int_{\partial R'} f(z)dz$ for a rectangle R' which is contained in R and similar to R. The purpose of this estimate is to find a condition on f that will guarantee that the rectangle R' is in $\mathcal{G}$. This is the point in the proof where we use the fact that f is analytic, and this will be the first application of Lemma 7.24 where the similarity condition is important. Let $\alpha \leq 1$ be the *aspect ratio* of R, that is, the length of the two shorter sides of R divided by the length of the two longer sides. Since every rectangle R' is similar to R, the aspect ratio of R' is also equal to α. Let us suppose that our rectangle R' has sides of length s and αs. Then the length of the contour $\partial R'$ is $(2 + 2\alpha)s$, and the area of R' is αs^2.

If we apply Proposition 5.1 in a naive way we will not be able to show that R' is in $\mathcal{G}$, since that estimate only gives a bound which is proportional to the length of $\partial R'$ and we need one which is proportional to the area of R'. But the Fundamental Theorem of Calculus comes to our rescue here. Choose any point $w \in R'$. Since $f'(w)$ exists, we may write, for all $z \in R'$,

$$f(z) = f(w) + f'(w)(z - w) + \mathcal{E}_w(z)(z - w)$$

where $\mathcal{E}_w(z) \to 0$ as $z \to w$. Since the linear function $L(z) = f(w) + f'(w)(z - w)$ obviously has an antiderivative, the Fundamental Theorem of Calculus tells us that its integral around any closed contour must be 0. Therefore we have

$$\begin{aligned} \int_{\partial R'} f(z)dz &= \int_{\partial R'} f(w) + f'(w)(z-w)dz + \int_{\partial R'} \mathcal{E}_w(z)(z-w)dz \\ &= \int_{\partial R'} \mathcal{E}_w(z)(z-w)dz. \end{aligned}$$

Now we can apply Proposition 5.1 to the integral of $\mathcal{E}_w(z)(z-w)$. We know that $|z-w|$ can not be larger than the diameter of R', which in turn is at most $\sqrt{2}s$. Thus if we knew that $|\mathcal{E}_w(z)| < M$ for all $z \in R'$ we could conclude that

$$\left|\int_{\partial R'} f(z)dz\right| \leq M(\sqrt{2}s)(2s + 2\alpha s) = M\left(\frac{\sqrt{2}(2+2\alpha)}{\alpha}\right) \text{Area}(R') = kM\,\text{Area}(R'),$$

where we define k to be the constant $\frac{\sqrt{2}(2+2\alpha)}{\alpha}$. In particular if $M = \epsilon/k$ we could conclude $|\int_{\partial R'} f(z)dz| \leq \epsilon\,\text{Area}(R')$. Thus we have shown:

- For any $w \in R$, if R' is a rectangle similar to R with $w \in R' \subset R$ and if $|\mathcal{E}_w(z)| \leq \dfrac{\epsilon}{k}$ for all $z \in R'$, then R' is in $\mathcal{G}$.

Since $\mathcal{E}_w(z) \to 0$ as $z \to w$, there exists $\delta > 0$ such that $|z - w| < \delta \Rightarrow \mathcal{E}_w(z) < \epsilon/k$. Therefore, if $w \in R' \subset R$, where R' is similar to R and $\text{Diam}(R') < \delta$, then it follows from the statement above that R' is in $\mathcal{G}$. This verifies that property ($\star$) in Lemma 7.24 holds for $\mathcal{G}$, and completes the proof. □