sum $\sum_{i=1}^N f^{(i)} : X \to \mathbf{R}$ by

$$(\sum_{i=1}^N f^{(i)})(x) := \sum_{i=1}^N f^{(i)}(x).$$

Example 14.5.1. If $f^{(1)} : \mathbf{R} \to \mathbf{R}$ is the function $f^{(1)}(x) := x$, $f^{(2)} : \mathbf{R} \to \mathbf{R}$ is the function $f^{(2)}(x) := x^2$, and $f^{(3)} : \mathbf{R} \to \mathbf{R}$ is the function $f^{(3)}(x) := x^3$, then $f := \sum_{i=1}^3 f^{(i)}$ is the function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) := x + x^2 + x^3$.

It is easy to show that finite sums of bounded functions are bounded, and finite sums of continuous functions are continuous (Exercise 14.5.1).

Now to add infinite series.

Definition 14.5.2 (Infinite series). Let (X, d_X) be a metric space. Let $(f^{(n)})_{n=1}^\infty$ be a sequence of functions from X to $\mathbf{R}$, and let f be another function from X to $\mathbf{R}$. If the partial sums $\sum_{n=1}^N f^{(n)}$ converge pointwise to f on X as $N \to \infty$, we say that the infinite series $\sum_{n=1}^\infty f^{(n)}$ *converges pointwise* to f, and write $f = \sum_{n=1}^\infty f^{(n)}$. If the partial sums $\sum_{n=1}^N f^{(n)}$ converge uniformly to f on X as $N \to \infty$, we say that the infinite series $\sum_{n=1}^\infty f^{(n)}$ *converges uniformly* to f, and again write $f = \sum_{n=1}^\infty f^{(n)}$. (Thus when one sees an expression such as $f = \sum_{n=1}^\infty f^{(n)}$, one should look at the context to see in what sense this infinite series converges.)

Remark 14.5.3. A series $\sum_{n=1}^\infty f^{(n)}$ converges pointwise to f on X if and only if $\sum_{n=1}^\infty f^{(n)}(x)$ converges to $f(x)$ for *every* $x \in X$. (Thus if $\sum_{n=1}^\infty f^{(n)}$ does not converge pointwise to f, this does not mean that it diverges pointwise; it may just be that it converges for some points x but diverges at other points x.)

If a series $\sum_{n=1}^\infty f^{(n)}$ converges uniformly to f, then it also converges pointwise to f; but not vice versa, as the following example shows.

Example 14.5.4. Let $f^{(n)} : (-1, 1) \to \mathbf{R}$ be the sequence of functions $f^{(n)}(x) := x^n$. Then $\sum_{n=1}^\infty f^{(n)}$ converges pointwise, but not uniformly, to the function $x/(1-x)$ (Exercise 14.5.2).

It is not always clear when a series $\sum_{n=1}^{\infty} f^{(n)}$ converges or not. However, there is a very useful test that gives at least one test for uniform convergence.

Definition 14.5.5 (Sup norm). If $f : X \to \mathbf{R}$ is a bounded real-valued function, we define the *sup norm* $\|f\|_\infty$ of f to be the number

$$\|f\|_\infty := \sup\{|f(x)| : x \in X\}.$$

In other words, $\|f\|_\infty = d_\infty(f, 0)$, where $0 : X \to \mathbf{R}$ is the zero function $0(x) := 0$, and d_∞ was defined in Definition 14.4.2. (Why is this the case?)

Example 14.5.6. Thus, for instance, if $f : (-2, 1) \to \mathbf{R}$ is the function $f(x) := 2x$, then $\|f\|_\infty = \sup\{|2x| : x \in (-2, 1)\} = 4$ (why?). Notice that when f is bounded then $\|f\|_\infty$ will always be a non-negative real number.

Theorem 14.5.7 (Weierstrass M-test). *Let (X, d) be a metric space, and let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of bounded real-valued continuous functions on X such that the series $\sum_{n=1}^{\infty} \|f^{(n)}\|_\infty$ is convergent. (Note that this is a series of plain old real numbers, not of functions.) Then the series $\sum_{n=1}^{\infty} f^{(n)}$ converges uniformly to some function f on X, and that function f is also continuous.*

Proof. See Exercise 14.5.3. □

To put the Weierstrass M-test succinctly: absolute convergence of sup norms implies uniform convergence of functions.

Example 14.5.8. Let $0 < r < 1$ be a real number, and let $f^{(n)} : [-r, r] \to \mathbf{R}$ be the series of functions $f^{(n)}(x) := x^n$. Then each $f^{(n)}$ is continuous and bounded, and $\|f^{(n)}\|_\infty = r^n$ (why?). Since the series $\sum_{n=1}^{\infty} r^n$ is absolutely convergent (e.g., by the ratio test, Theorem 7.5.1), we thus see that $f^{(n)}$ converges uniformly in $[-r, r]$ to some continuous function; in Exercise 14.5.2 we see that this function must in fact be the function $f : [-r, r] \to \mathbf{R}$ defined by $f(x) := x/(1 - x)$. In other words, the series $\sum_{n=1}^{\infty} x^n$ is pointwise convergent, but not uniformly convergent, on $(-1, 1)$, but is uniformly convergent on the smaller interval $[-r, r]$ for any $0 < r < 1$.

The Weierstrass M-test is especially useful in relation to *power series*, which we will encounter in the next chapter.

Exercise 14.5.1. Let $f^{(1)}, \ldots, f^{(N)}$ be a finite sequence of bounded functions from a metric space (X, d_X) to $\mathbf{R}$. Show that $\sum_{i=1}^N f^{(i)}$ is also bounded. Prove a similar claim when "bounded" is replaced by "continuous". What if "continuous" was replaced by "uniformly continuous"?

Exercise 14.5.2. Verify the claim in Example 14.5.4.

Exercise 14.5.3. Prove Theorem 14.5.7. (Hint: first show that the sequence $\sum_{i=1}^N f^{(i)}$ is a Cauchy sequence in $C(X \to \mathbf{R})$. Then use Theorem 14.4.5.)

14.6 Uniform convergence and integration

We now connect uniform convergence with Riemann integration (which was discussed in Chapter 11), by showing that uniform limits can be safely interchanged with integrals.

Theorem 14.6.1. *Let $[a, b]$ be an interval, and for each integer $n \geq 1$, let $f^{(n)} : [a, b] \to \mathbf{R}$ be a Riemann-integrable function. Suppose $f^{(n)}$ converges uniformly on $[a, b]$ to a function $f : [a, b] \to \mathbf{R}$. Then f is also Riemann integrable, and*

$$\lim_{n\to\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} f.$$

Proof. We first show that f is Riemann integrable on $[a, b]$. This is the same as showing that the upper and lower Riemann integrals of f match: $\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f$.

Let $\varepsilon > 0$. Since $f^{(n)}$ converges uniformly to f, we see that there exists an $N > 0$ such that $|f^{(n)}(x) - f(x)| < \varepsilon$ for all $n > N$ and $x \in [a, b]$. In particular we have

$$f^{(n)}(x) - \varepsilon < f(x) < f^{(n)}(x) + \varepsilon$$

for all $x \in [a, b]$. Integrating this on $[a, b]$ we obtain

$$\underline{\int}_{[a,b]} (f^{(n)} - \varepsilon) \leq \underline{\int}_{[a,b]} f \leq \overline{\int}_{[a,b]} f \leq \overline{\int}_{[a,b]} (f^{(n)} + \varepsilon).$$

Since $f^{(n)}$ is assumed to be Riemann integrable, we thus see

$$(\int_{[a,b]} f^{(n)}) - \varepsilon(b-a) \leq \underline{\int}_{[a,b]} f \leq \overline{\int}_{[a,b]} f \leq (\int_{[a,b]} f^{(n)}) + \varepsilon(b-a).$$

In particular, we see that

$$0 \leq \overline{\int}_{[a,b]} f - \underline{\int}_{[a,b]} f \leq 2\varepsilon(b-a).$$

Since this is true for every $\varepsilon > 0$, we obtain $\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f$ as desired.

The above argument also shows that for every $\varepsilon > 0$ there exists an $N > 0$ such that

$$|\int_{[a,b]} f^{(n)} - \int_{[a,b]} f| \leq 2\varepsilon(b-a)$$

for all $n \geq N$. Since ε is arbitrary, we see that $\int_{[a,b]} f^{(n)}$ converges to $\int_{[a,b]} f$ as desired. □

To rephrase Theorem 14.6.1: we can rearrange limits and integrals (on compact intervals $[a, b]$),

$$\lim_{n\to\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} \lim_{n\to\infty} f^{(n)},$$

provided that the convergence is uniform. This should be contrasted with Example 14.2.5 and Example 1.2.9.

There is an analogue of this theorem for series:

Corollary 14.6.2. *Let $[a, b]$ be an interval, and let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of Riemann integrable functions on $[a, b]$ such that the series $\sum_{n=1}^{\infty} f^{(n)}$ is uniformly convergent. Then we have*

$$\sum_{n=1}^{\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} \sum_{n=1}^{\infty} f^{(n)}.$$

Proof. See Exercise 14.6.1. □

This Corollary works particularly well in conjunction with the Weierstrass M-test (Theorem 14.5.7):

Example 14.6.3. (Informal) From Lemma 7.3.3 we have the geometric series identity

$$\sum_{n=1}^{\infty} x^n = \frac{x}{1-x}$$

for $x \in (-1, 1)$, and the convergence is uniform (by the Weierstrass M-test) on $[-r, r]$ for any $0 < r < 1$. By adding 1 to both sides we obtain

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

and the converge is again uniform. We can thus integrate on $[0, r]$ and use Corollary 14.6.2 to obtain

$$\sum_{n=0}^{\infty} \int_{[0,r]} x^n \, dx = \int_{[0,r]} \frac{1}{1-x} \, dx.$$

The left-hand side is $\sum_{n=0}^{\infty} r^{n+1}/(n+1)$. If we accept for now the use of logarithms (we will justify this use in Section 15.5), the anti-derivative of $1/(1-x)$ is $-\log(1-x)$, and so the right-hand side is $-\log(1-r)$. We thus obtain the formula

$$-\log(1-r) = \sum_{n=0}^{\infty} r^{n+1}/(n+1)$$

for all $0 < r < 1$.

Exercise 14.6.1. Use Theorem 14.6.1 to prove Corollary 14.6.2.

14.7 Uniform convergence and derivatives

We have already seen how uniform convergence interacts well with continuity, with limits, and with integrals. Now we investigate how it interacts with derivatives.

The first question we can ask is: if f_n converges uniformly to f, and the functions f_n are differentiable, does this imply that f is also differentiable? And does f_n' also converge to f'?

The answer to the second question is, unfortunately, no. To see a counterexample, we will use without proof some basic facts about trigonometric functions (which we will make rigourous in Section 15.7). Consider the functions $f_n : [0, 2\pi] \to \mathbf{R}$ defined by $f_n(x) := n^{-1/2}\sin(nx)$, and let $f : [0, 2\pi] \to \mathbf{R}$ be the zero function $f(x) := 0$. Then, since sin takes values between -1 and 1, we have $d_\infty(f_n, f) \leq n^{-1/2}$, where we use the uniform metric $d_\infty(f, g) := \sup_{x\in[0,2\pi]} |f(x) - g(x)|$ introduced in Defintion 14.4.2. Since $n^{-1/2}$ converges to 0, we thus see by the squeeze test that f_n converges uniformly to f. On the other hand, $f_n'(x) = n^{1/2}\cos(nx)$, and so in particular $|f_n'(0) - f'(0)| = n^{1/2}$. Thus f_n' does not converge pointwise to f', and so in particular does not converge uniformly either. In particular we have

$$\frac{d}{dx}\lim_{n\to\infty} f_n(x) \neq \lim_{n\to\infty}\frac{d}{dx} f_n(x).$$

The answer to the first question is also no. An example is the sequence of functions $f_n : [-1, 1] \to \mathbf{R}$ defined by $f_n(x) := \sqrt{\frac{1}{n^2} + x^2}$. These functions are differentiable (why?). Also, one can easily check that

$$|x| \leq f_n(x) \leq |x| + \frac{1}{n}$$

for all $x \in [-1, 1]$ (why? square both sides), and so by the squeeze test f_n converges uniformly to the absolute value function $f(x) := |x|$. But this function is not differentiable at 0 (why?). Thus, the uniform limit of differentiable functions need not be differentiable. (See also Example 1.2.10).

So, in summary, uniform convergence of the functions f_n says nothing about the convergence of the derivatives f_n'. However, the converse is true, as long as f_n converges at at least one point:

Theorem 14.7.1. *Let $[a, b]$ be an interval, and for every integer $n \geq 1$, let $f_n : [a, b] \to \mathbf{R}$ be a differentiable function whose derivative $f_n' : [a, b] \to \mathbf{R}$ is continuous. Suppose that the derivatives f_n' converge uniformly to a function $g : [a, b] \to \mathbf{R}$. Suppose also that there exists a point x_0 such that the limit $\lim_{n\to\infty} f_n(x_0)$ exists. Then the functions f_n converge uniformly to a differentiable function f, and the derivative of f equals g.*

Informally, the above theorem says that if f_n' converges uniformly, and $f_n(x_0)$ converges for some x_0, then f_n also converges uniformly, and $\frac{d}{dx} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{d}{dx} f_n(x)$.

Proof. We only give the beginning of the proof here; the remainder of the proof will be an exercise (Exercise 14.7.1).

Since f_n' is continuous, we see from the fundamental theorem of calculus (Theorem 11.9.4) that

$$f_n(x) - f_n(x_0) = \int_{[x_0,x]} f_n'$$

when $x \in [x_0, b]$, and

$$f_n(x) - f_n(x_0) = -\int_{[x,x_0]} f_n'$$

when $x \in [a, x_0]$.

Let L be the limit of $f_n(x_0)$ as $n \to \infty$:

$$L := \lim_{n\to\infty} f_n(x_0).$$

By hypothesis, L exists. Now, since each f_n' is continuous on $[a, b]$, and f_n' converges uniformly to g, we see by Corollary 14.3.2 that g is also continuous. Now define the function $f : [a, b] \to \mathbf{R}$ by setting

$$f(x) := L - \int_{[a,x_0]} g + \int_{[a,x]} g$$

for all $x \in [a, b]$. To finish the proof, we have to show that f_n converges uniformly to f, and that f is differentiable with derivative g; this shall be done in Exercise 14.7.1. □

Remark 14.7.2. It turns out that Theorem 14.7.1 is still true when the functions f_n' are not assumed to be continuous, but the proof is more difficult; see Exercise 14.7.2.

By combining this theorem with the Weierstrass M-test, we obtain

Corollary 14.7.3. *Let $[a, b]$ be an interval, and for every integer $n \geq 1$, let $f_n : [a, b] \to \mathbf{R}$ be a differentiable function whose derivative $f_n' : [a, b] \to \mathbf{R}$ is continuous. Suppose that the series $\sum_{n=1}^{\infty} \|f_n'\|_\infty$ is absolutely convergent, where*

$$\|f_n'\|_\infty := \sup_{x \in [a,b]} |f_n'(x)|$$

is the sup norm of f_n', as defined in Definition 14.5.5. Suppose also that the series $\sum_{n=1}^{\infty} f_n(x_0)$ is convergent for some $x_0 \in [a, b]$. Then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[a, b]$ to a differentiable function, and in fact

$$\frac{d}{dx} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{d}{dx} f_n(x)$$

for all $x \in [a, b]$.

Proof. See Exercise 14.7.3. □

We now pause to give an example of a function which is continuous everywhere, but differentiable nowhere (this particular example was discovered by Weierstrass). Again, we will presume knowledge of the trigonometric functions, which will be covered rigourously in Section 15.7.

Example 14.7.4. Let $f : \mathbf{R} \to \mathbf{R}$ be the function

$$f(x) := \sum_{n=1}^{\infty} 4^{-n} \cos(32^n \pi x).$$

Note that this series is uniformly convergent, thanks to the Weierstrass M-test, and since each individual function $4^{-n}\cos(32^n\pi x)$ is continuous, the function f is also continuous. However, it is not differentiable (Exercise 15.7.10); in fact it is a *nowhere differentiable function*, one which is not differentiable at *any* point, despite being continuous everywhere!

Exercise 14.7.1. Complete the proof of Theorem 14.7.1. Compare this theorem with Example 1.2.10, and explain why this example does not contradict the theorem.

Exercise 14.7.2. Prove Theorem 14.7.1 without assuming that f_n' is continuous. (This means that you cannot use the fundamental theorem of calculus. However, the mean value theorem (Corollary 10.2.9) is still available. Use this to show that if $d_\infty(f_n', f_m') \leq \varepsilon$, then $|(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| \leq \varepsilon|x - x_0|$ for all $x \in [a, b]$, and then use this to complete the proof of Theorem 14.7.1.

14.8 Uniform approximation by polynomials

As we have just seen, continuous functions can be very badly behaved, for instance they can be nowhere differentiable (Example 14.7.4). On the other hand, functions such as polynomials are always very well behaved, in particular being always differentiable. Fortunately, while most continuous functions are not as well behaved as polynomials, they can always be *uniformly approximated* by polynomials; this important (but difficult) result is known as the *Weierstrass approximation theorem*, and is the subject of this section.

Definition 14.8.1. Let $[a, b]$ be an interval. A *polynomial on* $[a, b]$ is a function $f : [a, b] \to \mathbf{R}$ of the form $f(x) := \sum_{j=0}^{n} c_j x^j$, where $n \geq 0$ is an integer and $c_0, \ldots, c_n$ are real numbers. If $c_n \neq 0$, then n is called the *degree* of f.

Example 14.8.2. The function $f : [1, 2] \to \mathbf{R}$ defined by $f(x) := 3x^4 + 2x^3 - 4x + 5$ is a polynomial on $[1, 2]$ of degree 4.

Theorem 14.8.3 (Weierstrass approximation theorem). *If $[a, b]$ is an interval, $f : [a, b] \to \mathbf{R}$ is a continuous function, and $\varepsilon > 0$, then there exists a polynomial P on $[a, b]$ such that $d_\infty(P, f) \leq \varepsilon$ (i.e., $|P(x) - f(x)| \leq \varepsilon$ for all $x \in [a, b]$).*

Another way of stating this theorem is as follows. Recall that $C([a, b] \to \mathbf{R})$ was the space of continuous functions from $[a, b]$ to $\mathbf{R}$, with the uniform metric d_∞. Let $P([a, b] \to \mathbf{R})$ be the space of all polynomials on $[a, b]$; this is a subspace of $C([a, b] \to \mathbf{R})$, since all polynomials are continuous (Exercise 9.4.7). The Weierstrass approximation theorem then asserts that every continuous function is an adherent point of $P([a, b] \to \mathbf{R})$; or in other words, that the closure of the space of polynomials is the space of continuous functions:

$$\overline{P([a, b] \to \mathbf{R})} = C([a, b] \to \mathbf{R}).$$

In particular, every continuous function on $[a, b]$ is the uniform limit of polynomials. Another way of saying this is that the space of polynomials is *dense* in the space of continuous functions, in the *uniform topology*.

The proof of the Weierstrass approximation theorem is somewhat complicated and will be done in stages. We first need the notion of an *approximation to the identity*.

Definition 14.8.4 (Compactly supported functions). Let $[a, b]$ be an interval. A function $f : \mathbf{R} \to \mathbf{R}$ is said to be *supported* on $[a, b]$ iff $f(x) = 0$ for all $x \notin [a, b]$. We say that f is *compactly supported* iff it is supported on some interval $[a, b]$. If f is continuous and supported on $[a, b]$, we define the improper integral $\int_{-\infty}^{\infty} f$ to be $\int_{-\infty}^{\infty} f := \int_{[a,b]} f$.

Note that a function can be supported on more than one interval, for instance a function which is supported on $[3, 4]$ is also automatically supported on $[2, 5]$ (why?). In principle, this might mean that our definition of $\int_{-\infty}^{\infty} f$ is not well defined, however this is not the case:

Lemma 14.8.5. *If $f : \mathbf{R} \to \mathbf{R}$ is continuous and supported on an interval $[a, b]$, and is also supported on another interval $[c, d]$, then $\int_{[a,b]} f = \int_{[c,d]} f$.*

Proof. See Exercise 14.8.1. □

Definition 14.8.6 (Approximation to the identity). Let $\varepsilon > 0$ and $0 < \delta < 1$. A function $f : \mathbf{R} \to \mathbf{R}$ is said to be an (ε, δ)-*approximation to the identity* if it obeys the following three properties:

(a) f is supported on $[-1, 1]$, and $f(x) \geq 0$ for all $-1 \leq x \leq 1$.

(b) f is continuous, and $\int_{-\infty}^{\infty} f = 1$.

(c) $|f(x)| \leq \varepsilon$ for all $\delta \leq |x| \leq 1$.

Remark 14.8.7. For those of you who are familiar with the Dirac delta function, approximations to the identity are ways to approximate this (very discontinuous) delta function by a continuous function (which is easier to analyze). We will not however discuss the Dirac delta function in this text.

Our proof of the Weierstrass approximation theorem relies on three key facts. The first fact is that polynomials can be approximations to the identity:

Lemma 14.8.8 (Polynomials can approximate the identity). *For every $\varepsilon > 0$ and $0 < \delta < 1$ there exists an (ε, δ)-approximation to the identity which is a polynomial P on $[-1, 1]$.*

Proof. See Exercise 14.8.8. □

We will use these polynomial approximations to the identity to approximate continuous functions by polynomials. We will need the following important notion of a *convolution*.

Definition 14.8.9 (Convolution). Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be continuous, compactly supported functions. We define the *convolution* $f * g : \mathbf{R} \to \mathbf{R}$ of f and g to be the function

$$(f * g)(x) := \int_{-\infty}^{\infty} f(y)g(x - y)\, dy.$$

Note that if f and g are continuous and compactly supported, then for each x the function $f(y)g(x-y)$ (thought of as a function of y) is also continuous and compactly supported, so the above definition makes sense.

Remark 14.8.10. Convolutions play an important rôle in Fourier analysis and in partial differential equations (PDE), and are also important in physics, engineering, and signal processing. An in-depth study of convolution is beyond the scope of this text; only a brief treatment will be given here.

Proposition 14.8.11 (Basic properties of convolution). *Let $f : \mathbf{R} \to \mathbf{R}$, $g : \mathbf{R} \to \mathbf{R}$, and $h : \mathbf{R} \to \mathbf{R}$ be continuous, compactly supported functions. Then the following statements are true.*

(a) *The convolution $f * g$ is also a continuous, compactly supported function.*

(b) *(Convolution is commutative) We have $f*g = g*f$; in other words*

$$\begin{aligned} f * g(x) &= \int_{-\infty}^{\infty} f(y)g(x-y)\, dy \\ &= \int_{-\infty}^{\infty} g(y)f(x-y)\, dy \\ &= g * f(x). \end{aligned}$$

(c) *(Convolution is linear) We have $f * (g + h) = f * g + f * h$. Also, for any real number c, we have $f * (cg) = (cf) * g = c(f * g)$.*

Proof. See Exercise 14.8.11. □

Remark 14.8.12. There are many other important properties of convolution, for instance it is associative, $(f * g) * h = f * (g * h)$, and it commutes with derivatives, $(f * g)' = f' * g = f * g'$, when f and g are differentiable. The Dirac delta function δ mentioned earlier is an identity for convolution: $f * \delta = \delta * f = f$. These results are slightly harder to prove than the ones in Proposition 14.8.11, however, and we will not need them in this text.

As mentioned earlier, the proof of the Weierstrass approximation theorem relies on three facts. The second key fact is that convolution with polynomials produces another polynomial:

Lemma 14.8.13. *Let $f : \mathbf{R} \to \mathbf{R}$ be a continuous function supported on $[0,1]$, and let $g : \mathbf{R} \to \mathbf{R}$ be a continuous function supported on $[-1,1]$ which is a polynomial on $[-1,1]$. Then $f * g$ is a polynomial on $[0,1]$. (Note however that it may be non-polynomial outside of $[0,1]$.)*

Proof. Since g is polynomial on $[-1,1]$, we may find an integer $n \geq 0$ and real numbers $c_0, c_1, \ldots, c_n$ such that

$$g(x) = \sum_{j=0}^{n} c_j x^j \text{ for all } x \in [-1,1].$$

On the other hand, for all $x \in [0,1]$, we have

$$f * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\ dy = \int_{[0,1]} f(y)g(x-y)\ dy$$

since f is supported on $[0,1]$. Since $x \in [0,1]$ and the variable of integration y is also in $[0,1]$, we have $x - y \in [-1,1]$. Thus we may substitute in our formula for g to obtain

$$f * g(x) = \int_{[0,1]} f(y) \sum_{j=0}^{n} c_j (x-y)^j\ dy.$$

We expand this using the binomial formula (Exercise 7.1.4) to obtain

$$f * g(x) = \int_{[0,1]} f(y) \sum_{j=0}^{n} c_j \sum_{k=0}^{j} \frac{j!}{k!(j-k)!} x^k (-y)^{j-k}\ dy.$$

We can interchange the two summations (by Corollary 7.1.14) to obtain

$$f * g(x) = \int_{[0,1]} f(y) \sum_{k=0}^{n} \sum_{j=k}^{n} c_j \frac{j!}{k!(j-k)!} x^k (-y)^{j-k}\ dy$$

(why did the limits of integration change? It may help to plot j and k on a graph). Now we interchange the k summation with the integral, and observe that x^k is independent of y, to obtain

$$f * g(x) = \sum_{k=0}^{n} x^k \int_{[0,1]} f(y) \sum_{j=k}^{n} c_j \frac{j!}{k!(j-k)!}(-y)^{j-k}\, dy.$$

If we thus define

$$C_k := \int_{[0,1]} f(y) \sum_{j=k}^{n} c_j \frac{j!}{k!(j-k)!}(-y)^{j-k}\, dy$$

for each $k = 0, \ldots, n$, then C_k is a number which is independent of x, and we have

$$f * g(x) = \sum_{k=0}^{n} C_k x^k$$

for all $x \in [0, 1]$. Thus $f * g$ is a polynomial on $[0, 1]$. □

The third key fact is that if one convolves a uniformly continuous function with an approximation to the identity, we obtain a new function which is close to the original function (which explains the terminology "approximation to the identity"):

Lemma 14.8.14. *Let $f : \mathbf{R} \to \mathbf{R}$ be a continuous function supported on $[0, 1]$, which is bounded by some $M > 0$ (i.e., $|f(x)| \leq M$ for all $x \in \mathbf{R}$), and let $\varepsilon > 0$ and $0 < \delta < 1$ be such that one has $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in \mathbf{R}$ and $|x - y| < \delta$. Let g be any (ε, δ)-approximation to the identity. Then we have*

$$|f * g(x) - f(x)| \leq (3M + 2\delta)\varepsilon$$

for all $x \in [0, 1]$.

Proof. See Exercise 14.8.14. □

Combining these together, we obtain a preliminary version of the Weierstrass approximation theorem: