Notes on Analytic Functions

In these notes we define the notion of an *analytic* function. While this is not something we will spend a lot of time on, it becomes much more important in some other classes, in particular complex analysis.

First we recall the following facts, which are useful in their own right:

Theorem 1. *Let (f_n) be a sequence of functions on (a, b) and suppose that each f_n is differentiable. If (f_n) converges to a function f uniformly and the sequence of derivatives (f_n') converges uniformly to a function g, then f is differentiable and $f' = g$.*

So, the uniform limit of differentiable functions is itself differentiable *as long as* the sequence of derivatives also converges uniformly, in which case the derivative of the uniform limit is the uniform limit of the derivatives. We will not prove this here as the proof is not easy; however, note that the proof of Theorem 26.5 in the book can be modified to give a proof of the above theorem in the special case that each derivative f_n' is continuous.

Note that we need the extra assumption that the sequence of derivatives converges uniformly, as the following example shows:

Example 1. For each n, let $f_n : (-1, 1)$ be the function $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$. We have seen before that (f_n) converges uniformly to $f(x) = |x|$ on $(-1, 1)$. Each f_n is differentiable at 0 but f is not, so the uniform limit of differentiable functions need not be differentiable.

We shouldn't expect that the uniform limit of differentiable functions be differentiable for the following reason: differentiability is a *local* condition which measures how rapidly a function changes, while uniform convergence is a *global* condition which measures how close functions are to one another. The idea is that we could have two functions very close to one another (say in the sup metric sense) but so that one is changing much more rapidly in which case their derivative may no longer be close to one another. To guarantee this, we need some extra assumptions as in the theorem.

Using the above result, we can give a proof that power series are (infinitely) differentiable which is different than the one given in the book, using the fact that power series converge uniformly on certain intervals. We will also need Lemma 26.3 in the book, showing that a power series and the series obtained by differentiating term-by-term have the same radius of convergence.

Theorem 2. *Suppose that $\sum a_n x^n$ is a power series which radius of convergence $R > 0$. Then the function f to which $\sum a_n x^n$ converges is infinitely differentiable on $(-R, R)$.*

Proof. Let $0 < R_1 < R$. Then $\sum a_n x^n$ converges to f uniformly on $[-R_1, R_1]$. This means that the sequence (f_k) of partial sums:

$$f_k = a_0 + a_1 x + \cdots + a_k x^k$$

converges uniformly to f on $[-R_1, R_1]$. Now, each f_k is differentiable and

$$f_k' = a_1 + 2a_2 x + \cdots + k a_k x^{k-1}.$$

However, now we see that the functions f_k' are the partial sums of the series $\sum n a_n x^{n-1}$. Since this latter series also has radius of convergence R, it converges to some function g on $(-R, R)$, and moreover its sequence of partial sums (f_k') converges uniformly to g on $[-R_1, R_1]$. Since (f_k) converges uniformly to f on $[-R_1, R_1]$, each f_k is differentiable on $(-R_1, R_1)$, and (f_k') converges uniformly to g on $[-R_1, R_1]$, Theorem 1 says that f is differentiable and $f' = g$; i.e., the power series $\sum a_n x^n$ is differentiable on $(-R_1, R_1)$ and its derivative is $\sum n a_n x^{n-1}$. Taking $R_1 \to R$ shows that $\sum a_n x^n$ is differentiable on all of $(-R, R)$.

Applying the same result to the power series $\sum n a_n x^{n-1}$ shows that this is differentiable, so the original series was twice differentiable, and continuing in this manner shows that the original series is infinitely differentiable on $(-R, R)$. $\square$

The upshot is that power series are (infinitely) differentiable on their intervals of convergence, and their derivatives are obtained by differentiating term-by-term, so that you can manipulate them as if they were infinite polynomials. Note also that the same works for series $\sum a_n (x - x_0)^n$ which are not centered at 0.

Now we come to the notion of analyticity. Roughly, an analytic function is one which can be expressed as a power series, although the power series needed to do so may change from point to point:

Definition 1. A function $f : (a, b) \to \mathbb{R}$ is said to be *analytic* if for any $x_0 \in (a, b)$ there exists a power series $\sum a_n (x - x_0)^n$ and an open interval I_{x_0} centered at x_0 such that $\sum a_n (x - x_0)^n$ converges to f on I_{x_0}.

Example 2. Many of your favorite functions from calculus are analytic, such as

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \ \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \text{ and } \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Each of these series actually converge on all of $\mathbb{R}$.

Example 3. Here is an example of an analytic function were we need different series to express it. Let $f : \mathbb{R} \backslash \{1\} \to \mathbb{R}$ be the function $f(x) = \frac{1}{1-x}$. We know that

$$f(x) = \sum_{n=0}^{\infty} x^n \text{ for } x \in (-1, 1),$$

showing that f is analytic on $(-1, 1)$. We claim that f is analytic on all of $\mathbb{R} \backslash \{1\}$, but now we see that to express f as a convergent power series for $|x| > 1$, the above series won't work.

How do we express the above function as a power series for $|x| > 1$? Note the following. Suppose that f is analytic and fix x_0. According to the definition, there is some power series $\sum a_n (x - x_0)^n$ such that

$$f(x) = \sum a_n (x - x_0)^n = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + \cdots \text{ for } x \text{ close enough to } x_0.$$

The first thing to note is that since power series are always differentiable (in fact, infinitely differentiable), an analytic function must itself be infinitely differentiable. We can use this fact to determine what the coefficients a_n in the above expression must be.

Plugging in $x = x_0$ gives

$$f(x_0) = a_0 + a_1 (x_0 - x_0) + a_2 (x_0 - x_0)^2 + \cdots = a_0$$

since all terms with a positive power of $x - x_0$ vanish when plugging in x_0. Taking derivatives of $f(x) = \sum a_n (x - x_0)^n$ at $x = x_0$ gives

$$f'(x) = a_1 + 2a_2 (x - x_0) + \text{higher order terms},$$

so

$$f'(x_0) = a_1 + 2a_2(x_0 - x_0) + \cdots = a_1.$$

In general, we have

$$f^{(k)}(x) = k!a_k + (k+1)!a_{k+1}(x - x_0) + \text{higher order terms},$$

so

$$f^{(k)}(x_0) = k!a_k.$$

The upshot is that when expressing f as a power series of the form $\sum a_n(x-x_0)^n$, the coefficients a_n *must* be the so-called *Taylor coefficients* of f at x_0:

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

The series $\sum \frac{f^{(k)}(x_0)}{k!}(x-x_0)^n$ is called the *Taylor series* of f around x_0. Using this, we can reformulate the definition of an analytic function as follows:

Definition 2. An infinitely-differentiable function $f : (a, b) \to \mathbb{R}$ is *analytic* if for any $x_0 \in (a, b)$, the Taylor series of f around x_0 has positive radius of convergence and converges to f on its interval of convergence.

The point is that we have no choice as to what power series we need when trying to express a function as a power series, we must use the Taylor series. The only remaining questions that determine whether or not a function is analytic is whether its Taylor series around any point has positive radius of convergence, and, more importantly, whether its Taylor series converges to the function itself.

Example 4. Returning to the function $f : \mathbb{R}\backslash\{1\} \to \mathbb{R}$ of Example 3, to express $\frac{1}{1-x}$ as a power series at any $x \neq 1$, we compute its Taylor series. It will turn out that this Taylor series will have positive radius of convergence and will converge to $\frac{1}{1-x}$.

Example 5. Not all infinitely differentiable functions are analytic. For example, let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

Example 3 in Section 31 of the book shows that this function is infinitely differentiable, and in particular that $f^{(k)}(0) = 0$ for all k. Thus, the Taylor series of f around 0 is the zero series, which has infinite radius of convergence but does not equal f for $x > 0$. Thus f is not analytic.

So, there are functions which are differentiable but not twice differentiable, twice differentiable but not three times differentiable, k times differentiable but not $k+1$ times differentiable, and infinitely differentiable but not analytic. One of the amazing things you will see in complex analysis is that there are no such distinctions among *complex differentiable* functions.

Finally, let us give one important property of analytic functions, showing that they are in a sense *rigid*.

Theorem 3. *Suppose that f, g are two analytic functions on $\mathbb{R}$. If $f(x) = g(x)$ for all x in some interval (a, b), then $f = g$ on all of $\mathbb{R}$.*

The point is that the interval (a, b) can be incredibly small, and nevertheless having analytic functions equal on any small interval whatsoever forces them to be the same everywhere. Quite an amazing fact, given that it is easy to draw graphs of infinitely differentiable functions where something like this wouldn't be true.

First we prove a lemma, saying that nonzero analytic functions are *isolated* zeros:

Lemma 1. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a nonzero, analytic function and that $c \in \mathbb{R}$ satisfies $f(c) = 0$. Then there is an interval $(c-\delta, c+\delta)$, with $\delta > 0$, around c on which f is nonzero except at c itself.*

Proof. Since f is analytic, there is an interval I around c and a power series

$$\sum_{n=0}^{\infty} a_n (x-c)^n$$

which converges to $f(x)$ for any $x \in I$. (Actually, this will be the Taylor series of f around c and I will be the interval of convergence of this series, but we don't need that precise description here.) Since f is not the zero function, this power series is not the zero series, so at least one coefficient is nonzero—let k be the smallest positive integer so that $a_k \neq 0$. Note that $k \geq 1$ since $a_0 = f(c) = 0$.

Define the function g on I by setting

$$g(x) = \sum_{n=0}^{\infty} a_{n+k} (x-c)^n \text{ for } x \in I.$$

Note that g then is also analytic on I, so in particular it is continuous. Since $a_k = g(c) \neq 0$, there exists $\delta > 0$ so that for any $y \in (c-\delta, c+\delta)$, $g(y) \neq 0$. To see this, note that the set of points where g is nonzero is open in I since it is the preimage under the continuous function g of the open subset $\mathbb{R} \backslash \{0\}$ of $\mathbb{R}$; thus c is an interior point of this set, so there exists $\delta > 0$ such that $(c-\delta, c+\delta)$ is contained within the set of points where g is nonzero. Another way to see this is to note that since $|g(c)| > 0$, by the ϵ-δ definition of continuity at c, there exists $\delta > 0$ so that

$$|c - x| < \delta \text{ implies } |g(c) - g(x)| < |g(c)|.$$

Since $|g(c)| - |g(x)| \leq |g(c) - g(x)|$, from this it follows that for such x in the δ-interval around c, $|g(x)| > 0$, so $g(x) \neq 0$.

We then have

$$f(x) = (x-c)^k g(x) \text{ for any } x \in (c-\delta, c+\delta).$$

Since g is nonzero on this interval, this then can only be zero when $(x-c)^k$ is zero, and hence only at $x = c$. Thus $(c-\delta, c+\delta)$ is an interval around c on which f is nonzero except at c itself. □

Proof of Theorem. Consider the function $f - g$. This is analytic on $\mathbb{R}$ and has non-isolated zeros since $f(x) - g(x) = 0$ for $x \in (a, b)$. The lemma then implies that $f - g$ must be the zero function, so $f = g$ on all of $\mathbb{R}$ as claimed. □

Finally, we show that power series are always integrable, and that their integrals are given by term by term differentiation. First, we need the following fact, showing that integrability is preserved under uniform convergence:

Theorem 4. *Suppose that (f_n) is a sequence of integrable functions on $[a, b]$ converging uniformly to f. Then f is integrable and*

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

Proof. We show that f is integrable using the Riemann-Lebesgue theorem; on the final homework you will show this directly using the definition of integrability. We know that if each f_n is continuous at $x \in [a, b]$, then the uniform limit f is also continuous at x. Thus, if f is not continuous at some point, it must be that at least one of the f_n's is also discontinuous at that point. Hence, the discontinuity set of f is contained in the union of the discontinuity sets of the f_n's:

$$D(f) \subseteq \bigcup_n D(f_n).$$

Since each f_n is integrable, each $D(f_n)$ is a zero set. Hence so is their countable union, and thus so is the subset $D(f)$. Since the uniform limit of bounded functions is bounded, this shows that f is integrable by the Riemann-Lebesgue theorem.

Let $\epsilon > 0$ and choose N such that

$$|f_n(x) - f(x)| < \frac{\epsilon}{b-a} \text{ for } n \geq N.$$

Then for $n \geq N$, we have

$$\left| \int_a^b f - \int_a^b f_n \right| = \left| \int_a^b (f - f_n) \right| \leq \int_a^b |f(x) - f_n(x)| \, dx \leq \int_a^b \frac{\epsilon}{b-a} \, dx = \epsilon.$$

We conclude that the sequence of numbers $\int_a^b f_n$ converges to $\int_a^b f$, as claimed. □

Now we show that power series are always integrable; we will need Lemma 26.3 in the book, showing that a power series and the one obtained by term-by-term integration have the same radius of convergence.

Theorem 5. *Suppose that the power series $\sum a_n x^n$ has radius of convergence $R > 0$. Then for any $0 < R_1 < R$, the function to which $\sum a_n x^n$ is integrable on $[-R_1, R_1]$ and*

$$\int_{-R_1}^{R_1} \left(\sum a_n x^n \right) dx = \sum \left(\int_{-R_1}^{R_1} a_n x^n \, dx \right).$$

Proof. Let f be the function to which $\sum a_n x^n$ converges on $(-R, R)$. Then the sequence of partial sums:

$$f_n = a_0 + a_1 x + \cdots + a_n x^n$$

converges uniformly to f on $[-R_1, R_1]$. Since each f_n is integrable (since it is continuous), the previous theorem implies that f is integrable on $[-R_1, R_1]$ and

$$\int_{-R_1}^{R_1} f = \lim_{n \to \infty} \int_{-R_1}^{R_1} f_n.$$

But the sequence of numbers $\int_{-R_1}^{R_1} f_n$ are the partial sums of the series

$$\sum \left(\int_{-R_1}^{R_1} a_n x^n \, dx \right),$$

which also converges on $(-R, R)$. We conclude that

$$\int_{-R_1}^{R_1} \left(\sum a_n x^n \right) dx = \int_{-R_1}^{R_1} f = \lim_{n \to \infty} \int_{-R_1}^{R_1} f_n = \sum \left(\int_{-R_1}^{R_1} a_n x^n \, dx \right)$$

as claimed. □