SEQUENCES AND SERIES OF FUNCTIONS II: POWER SERIES AND TAYLOR SERIES

PETE L. CLARK

1. Power Series as Functions

We are now in a position to draw some spectacular consequences of our hard-earned theorems concerning uniform convergence. Namely, we will use the theory we have developed to show that the function defined by a power series on the interior of its interval of convergence is a wonderful thing: it is continuous and indeed, infinitely differentiable. Moreover, its derivatives (and antiderivatives) can be computed term-by-term. That is to say, in many important ways, functions defined by convergent power series can be manipulated as easily as polynomial functions.

Here is the main result.

Theorem 1. *(Wonderful Properties of Power Series) Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R > 0$. Consider $f(x) = \sum_{n=0}^{\infty} a_n x^n$ as a function on the interval $I = (-R, R)$.*
a) f is a continuous function.
b) Indeed, f is differentiable. Moreover, its derivative can be computed "term-by-term":

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

c) Since f' is again a power series which has, as we have seen, the same radius of convergence $R > 0$, it follows that f is in fact infinitely differentiable. Moreover we have the formula $f^{(n)}(0) = (n!)a_n$.

Proof: We begin with the proof of part a). Let $0 < A < R$ be any positive number less than R, so f in particular is a function on the closed interval $[-A, A]$. We claim that the series $\sum_n a_n x^n$ converges *uniformly* to f on the subinterval $[-A, A]$. If so, then since the Nth partial sum, $\sum_{n=0}^{N} a_n x^n$ is just a polynomial, we have a uniformly convergent sequence of continuous functions, so the limit function f is also continuous. Moreover, since any $x \in I = (-R, R)$ lies in $[-A, A]$ for some $A < R$, and to say that f is continuous on I just means that it is continuous at every point of I, it will follow that f is continuous on I.

To see that the convergence is uniform is (literally) a textbook application of the Weierstrass M-Test.[1] Namely, on the interval $[-A, A]$, the maximum value of $|a_n x^n$ is $||a_n x^n|| = |a_n| A^n$. Since we know that power series converge absolutely

[1] Indeed this theorem is also due to Weierstrass, and he certainly must have propounded the M-Test with this application in mind.

on the interior of the interval of convergence and $A < R$, we have

$$\sum_n ||a_n x^n|| = \sum_n |a_n| A^n < \infty,$$

so the M-Test shows the uniform convergence on $[-A, A]$.

For part b), recall first that the uniform limit of a sequence of differentiable functions need *not* be differentiable, and even if it is, the derivative of the limit need not be equal to the limit of the derivative. Rather, we proved a more complicated result (Theorem 4 from the first handout on sequences and series of functions): if $f_n' g$, then f is differentiable and $f' = g$. But since $S_n'(x) = \sum_{n=1}^N n a_n x^{n-1}$ gives the partial sums of a power series $g = \sum_{n=0}^\infty n a_n x^{n-1}$ with the same radius of convergence $R > 0$, we may apply part a) to g to conclude that $S_N'(x)$ converges uniformly to $g(x)$ on $[-A, A]$ (for any $A > 0$) and then the aformentioned Theorem 4 to conclude that $f'(x) = g(x)$. As in part a), because we can do this for any $0 < A < R$, this shows that f is differentiable on all of I with derivative equal to $\sum_n n a_n x^{n-1}$.

For part c), as stated in the theorem, since differentiation preserves the radius of convergence, we get that f is infinitely differentiable. That the given formula holds for the nth derivative at zero we leave to the reader to check. This completes the proof of the theorem.

Remark: Note of course that we did *not* show that the partial sums of the power series converge uniformly on all of $I = (-R, R)$. When $R = \infty$ this is never the case (except in the trivial case where all but finitely many coefficients are equal to zero), as you are asked to show in Exercise 76.

In the case that $R < \infty$, we can ask what happens at the boundary. Suppose for the sake of simplicity that $R = 1$ and the power series converges at $R = 1$, i.e., $\sum_{n=0}^\infty a_n$ converges. Then we can in fact view f as a function with domain (at least) $(-1, 1]$. If f were discontinuous at 1, we could conclude that the convergence is not uniform. However, Abel's Theorem says exactly that f always *is* continuous at $x = 1$ (from the left, of course). In fact it turns out to be true that if $f(1)$ converges, then the convergence is necessarily uniform on $(-1, 1]$, and Abel's Theorem follows from this fact because the uniform limit of continuous functions must be continuous. Since this is far from a key point for us, we omit the proof (which is in fact similar to, and if anything easier than, the proof of Abel's Theorem). The special case when the a_n's are non-negative is easier and is pursued in Exercise 77.

Exercise 76: Let $\sum_n a_n x^n$ be a power series with infinite radius of convergence, so it defines a function $f(x)$ on all of $\mathbb{R}$. Show that the following are equivalent:
a) The partial sums $\sum_{n=0}^N a_n x^n$ converge uniformly to f on all of $\mathbb{R}$.
b) $a_n = 0$ for all sufficiently large n (i.e., $f(x)$ is a polynomial).
Hint: Use the Cauchy criterion to show that if $\sum_n f_n(x)$ is a series of functions converging uniformly on the entire real line, then $||f_n|| \to 0$.

Exercise 77:
Suppose that $\sum_{n=0}^\infty a_n x^n$ is a power series with each $a_n \geq 0$. Suppose that the radius of convergence is 1, so that the power series defines a function $f(x)$ at least on $(-1, 1)$. Show that the following are equivalent:

a) $f(1) = \sum_n a_n$ converges.
b) The power series converges uniformly on $[0, 1]$.
c) $f(x)$ is bounded on $[0, 1)$.
(Hint for a) implies b): suppose $f_n : (a, b] \to \mathbb{R}$ is any sequence of functions converging uniformly on (a, b) and pointwise on $(a, b]$. Show that the convergence is uniform on $(a, b]$.)

The part of Theorem 1 that asserts $a_n = \frac{f^{(n)}(0)}{n!}$ for any power series with positive radius of convergence has the following important consequence:

Corollary 2. *(Uniqueness Theorem) Let $\sum_n a_n x^n$ and $\sum_n b_n x^n$ be two power series with positive radii of convergence R_1 and R_2, so they define functions $f(x)$ on $(-R_1, R_1)$ and $g(x)$ on $(-R_2, R_2)$. Suppose that $f(x) = g(x)$ for all x in $(-\delta, \delta)$ for some $\delta \leq \min(R_1, R_2)$. Then $a_n = b_n$ for all n.*

Exercise 78: Prove Corollary 2.

In other words, if two power series define the same function – even in just a small interval around zero – then all of their coefficients must be equal. This is not at all obvious, since of course in general $\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} b_n$ does not imply $a_n = b_n$ for all n! Another way of saying it is that the only power series that a function can be equal to in a small interval about zero is its Taylor series, which brings us to the next section.

2. Taylor series

The following is a review of topics that ought to be familiar (possibly with a less sophisticated presentation) from both calculus courses and from Math 242.

Let f be an infinitely differentiable function defined on some interval I containing 0. Recall that for any positive integer N, there is a unique degree N polynomial $T_N(x)$ satisfying the $N + 1$ conditions:

$$T_N(0) = f(0),\ T_N'(0) = f'(0), \ldots,\ T_N^{(N)}(0) = f^{(N)}(0),$$

namely $T_N(x) = \sum_{n=0}^{N} \frac{f^n(0)}{n!} x^n$. $T_N(x)$ is called the **degree N Taylor polynomial of f centered at $x = 0$.** [2]

Recall that T_N is the best approximation to f near 0 by a degree N polynomial. The meaning of this (which is not always made clear in beginning courses) is: it is the unique polynomial P of degree at most N for which

$$\lim_{x \to 0} \frac{f(x) - P}{x^N} = 0.$$

This is an immediate consequence of **Taylor's theorem**, which bounds the error between f and T_N in terms of the $(N + 1)$st derivative of f: [3]

$$f(x) = T_N(x) + \frac{1}{N!} \int_0^x (x - t)^N f^{(N+1)}(t) dt.$$

[2]Many would call it the MacLaurin polynomial, and reserve the name Taylor polynomial for the analogous polynomial defined with x replaced by $x - c$. For my part, I find this quite inane.

[3]For the (quite straightforward) proof of Taylor's Theorem, see p. 267 of your text.

Note that this is not really an error estimate, in that it tells us an exact difference between $f(x)$ and $T_N(x)$. However, it can be used to derive a very useful estimate for the error $E_N(x) := |f(x) - T_N(x)|$, namely: fix a closed interval $[-A, A]$ inside the domain of f and for any positive integer n, let $M_n := M_{n(A)}$ be $||f^{(n)}||$, i.e., the maximum value of the nth derivative on the interval $[-A, A]$. (Since f has derivatives of all orders, all derivatives are continuous, so each derivative really does have a maximum value on any interval of the form $[-A, A]$.) Then:

Corollary 3. *(Taylor Error Estimate) With hypotheses as above,*

$$E_N(x) = |f(x) - T_N(x)| \leq \frac{M_{N+1}|x|^{N+1}}{(N+1)!}.$$

For the proof, just plug in the estimate for $|f^{N+1}(x)| \leq M_{N+1}$ and compute the integral.

If we now put $a_n = \frac{\max\{|f^n(x)|\ x \in [-A,A]\}}{n!} = \frac{M_n}{n!}$, we see the close relationship between the theory we developed in the last section and the apparently much more concrete question of whether the Taylor polynomials can be used to approximate f to arbirary accuracy on the interval $[-A, A]$. Indeed, to say that we can approximate f with arbitrary accuracy by Taylor polynomials at a point x in the domain means that the error $E_N(x) \to 0$ as $N \to \infty$. That $E_N \to 0$ is equivalent to $f(x) = \lim_{N\to\infty} T_N(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}}{n!} x^n$, i.e., $f(X)$ is equal to its **Taylor series**. But in turn the estimates we have made show the following:

Theorem 4. *Let $f : [-A, A] \to \mathbb{R}$ be an infinitely differentiable function. For any $n \geq 0$, let M_n be the maximum value of $|f^{(n)}(x)|$ on $[-A, A]$. Suppose that $\lim_{n\to\infty} \frac{M_n}{n!} A^n = 0$. Then the radius of convergence of the Taylor series is at least A, and the partial sums of the Taylor series converge uniformly to f on $[-B, B]$ for any $0 < B < A$.*

In other words, the theorem is giving conditions for an infinitely differentiable function to be equal to the function defined by a convergent power series.

Proof: If $\lim_{n\to\infty} \frac{M_n}{n!} A^n = 0$, then in particular $\lim_{n\to\infty}(f^{(n)}(0)/n!)A^n = 0$, and this implies that the radius of convergence of $\sum_n (f^{(n)}(0)/n!)x^n$ is at least A (since it is a property of power series that if we plug in any value strictly larger than the radius of convergence, the general term diverges to infinity in absolute value). Thus the partial sums of the Taylor series converge uniformly to $T(x)$ on $[-B, B]$. By Corollary 3 $E_N(x) \to 0$ pointwise on $[-A, A]$ so that $T_N(x) \to f(x)$ pointwise on $[-A, A]$. It follows that the Taylor series actually converges on $[-A, A]$, that $f(x) = T(x)$ on $[-A, A]$ and the convergence of T_n to f is uniform on closed subintervals $[-B, B]$.

The converse of the theorem is also true: if for any $A > 0$, $\limsup_n \frac{M_n}{n!} A^n > 0$, then there is no interval about zero on which f is equal to a power series with positive radius of convergence. This is also easily seen, just because if f were given by a power series on some interval $(-\delta, \delta)$, then taking any $A < \delta$, the series in the statement of the theorem would converge.

Remark: Of course all of this would work just as well with power series centered at a point $c \neq 0$.

Definition: We say that an infinitely differentiable function f is **analytic** at $x = 0$ if it satisfies the hypothesis of Theorem 4, or equivalently, if in some interval about zero it is given by a convergent power series expansion. More generally, a function f defined on some open interval I (possibly all of $\mathbb{R}$) is said to be **analytic** if near any point c in its domain it is equal to a power series in $(x - c)$ with positive radius of convergence. As their name implies, analytic functions are the favored class of functions to study in analysis: they have all kinds of nice properties that merely infinitely differentiable functions need not have. For instance, it follows from the Uniqueness Theorem that if two analytic functions are equal in a small interval then they are equal on their entire domain of definition.

Example of an infinitely differentiable function $f : \mathbb{R} \to \mathbb{R}$ which is *not* analytic: let $f(x)$ be defined by 0 if $x \leq 0$ and $e^{\frac{-1}{x}}$ when $x > 0$. The point is that f decays so rapidly as $x \to 0^+$ that $f^{(n)}(0) = 0$ for all n. Thus the Taylor series for f at $x = 0$ is just $\sum_n 0x^n = 0$, which clearly is not equal to f in any interval about 0. Note that we could equall well have taken f to be $e^{\frac{1}{|x|}}$ (i.e., define f for negative values by reflection through the y-axis), and still all the derivatives will be equal to zero at $x = 0$. We chose the example we did to illustrate the point that two infinitely differentiable functions can agree on an interval – in this case, $f(x)$ is equal to the zero function on all of $(-\infty, 0]$ – without being equal for all values in their domain. Analytic functions cannot do this: they are determined by their values on any arbitrarily small interval in their domain.

In this example the Taylor series has infinite radius of convergence; the function just happens not to be equal to its Taylor series for any positive value of x. On the other hand, there is no guarantee that the Taylor series must have positive radius of convergence. Indeed, we state without proof the following remarkable result:

Theorem 5. *(Borel) Let $\{a_n\}_{n=0}^{\infty}$ be* any *sequence of real numbers. Then there exists an infinitely differentiable function $f : \mathbb{R} \to \mathbb{R}$ such that for all n,*
$f^{(n)}(0) = (n!)a_n$, so that $\sum_{n=0} a_n x^n$ is the Taylor series of f centered at $x = 0$.

Applying Borel's theorem to, say, the sequence $a_n = n^n$, it asserts the existence of a function whose Taylor series converges only at $x = 0$.

After all this, let us give some applications where Theorem 4 can be used to show that functions are equal to their Taylor series. In fact, we will give some examples that are familiar from calculus. However, we will then explain why the approach thus far is not really satisfactory and suggest a way to proceed directly from the power series expansions themselves.

Example: In calculus class, one learns (or is told) that the exponential function $f(x) = e^x$ is a differentiable function with $f(0) = 1$ and $f'(x) = f(x)$. If f is any such function, it is clearly infinitely differentiable (because $f = f' = f'' = \ldots$); moreover, if we fix any $A > 0$, take $M = ||f||_{[-A,A]}$, i.e., the maximum value of f on $[-A, A]$. This is finite since f is continuous, and in fact it is the same bound for any derivative of f. It follows that the error $|f(x) - T_N(x)|$ is at most

$E_N(x) \leq M\frac{A^{N+1}}{(N+1)!}$. But since we know that the series $\sum_n x^n/n!$ has infinite radius of convergence, it follows that $\lim_{n\to\infty} \frac{A^{N+1}}{(N+1)} \to 0$, so $E_N(x) \to 0$ uniformly on $[-A, A]$, and $f(x)$ is equal to its Taylor series expansion at zero, namely to $\sum_{n=0}^{\infty} \frac{x^n}{n!}$.

Example: In calculus class, one learns (or is told) that the functions $f(x) = \sin x$ and $g(x) = \cos x$ satisfy the following properties: f and g are both differentiable, and $f' = g, g' = f$; moreover $f(0) = 0$ and $g(0) = 1$. Again this determines all the derivatives of f and g, and the Taylor series expansion for f about zero must be

$$T_f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!},$$

while the Taylor expansion for g must be

$$T_g(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

Because of the periodicity of the derivatives, it follows again that $E_N(x) \to 0$ uniformly on $[-A, A]$ for any $A > 0$. You are asked to prove a more general result in Exercise 78. Thus it follows that $f(x)$ and $g(x)$ are both given by their Taylor expansions.

Exercise 79: a) Let $f : \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function with the following property: for any $A > 0$, there exists a number $M = M(A)$ such that for all n and all $x \in [-A, A]$, $|f^{(n)}(x)| \leq M$. Show that f is equal to a power series with infinite radius of convergence.
b) Suppose that f is a function which is equal to its own nth derivative for some $n \geq 1$. Show that the hypothesis of part a) applies.
c)* Show that, for any fixed $n \geq 1$, the set of all infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$ which are equal to their own nth derivatives, forms a subspace of the $\mathbb{R}$-vector space of all functions from $\mathbb{R}$ to $\mathbb{R}$. What is its dimension? Can you find a basis?

However, there is something a bit fishy going on here. How do we know that the functions e^x, $\sin x$ and $\cos x$ are differentiable? A little reflection (possibly with a calculus text in hand) shows that this comes down to showing the following limits:

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1;$$

$$\lim_{h\to 0} \frac{\sin h}{h} = 1;$$

$$\lim_{h\to 0} \frac{1 - \cos h}{h} = 0.$$

But if you then look and see what the calculus books say for why these limits exist, you find arguments which are at best unconvincing and at worst completely circular: for instance, my calculus textbook (the one I am teaching from right now) defines e as the unique number a such that the slope of the tangent line to $y = a^x$ at $x = 0$ is equal to 1, notes that numerically and geometrically it seems that such an a exists and that it is somewhere between 2 and 3, and then when it comes time to differentiate e^x it reminds us that e has been defined so that this limit is equal to 1: no good. For the sine function, the argument that $\frac{\sin x}{x} \to 1$ uses geometric considerations about arc lengths of curves; after a very careful reading of all the details we find that a certain process of computing arclengths via inscribed polygons is assumed to converge: no good.

Let us try instead a different approach: we certainly have available the convergent power series that we want e^x, $\sin x$ and $\cos x$ to turn out to be. Instead of trying to show that the functions, defined rather imprecisely, are equal to their power series expansions, can we work directly with the power series expansions and show that they have all the properties that e^x, $\sin x$ and $\cos x$ are supposed to have? We can, and that will be our approach to the elementary functions.