Combinatorial Calculus: From Taylor Series to Feynman Diagrams

Theo Johnson-Freyd

July 7–12, 2008
Canada/USA Mathcamp
Reed College

Abstract

A Feynman diagram is many things (a picture, a process, an event, a morphism). For me, a Feynman diagram is a combinatorial integral. This class will explain some of the beautiful combinatorics that underlies calculus, beginning with derivatives and Taylor's theorem, and concluding with integrals and Feynman Diagrams. For example, the generalized Chain Rule ($d^n[f(u(x))]/dx^n$ in terms of df/du and du/dx) also generalizes the number of partitions of n objects. Along the way, we will develop some multi-variable calculus — certainly not a whole course, but whatever is needed to get at the full combinatorial elegance.

These lecture notes are from a class taught at Canada/USA Mathcamp, a summer program for gifted high school students. Definitions and exercises were handed out each day. These notes are available at http://math.berkeley.edu/∼theojf/CombinatorialCalculus.pdf.

Day 1: Taylor Series

Show of hands: how many of you have taken high school calculus? How many know Taylor series?

Well, then, I'll remind you about Taylor series. In math classes, words like "recall", "remember", and "remind" mean: "I know you don't really know this, but I'll generally pretend that you do. Still, here's the definition."

Let's say I have a nice function f. How nice?

Definition 1.1 *A function is* continuous at 0 *if it satisfies*

$$f(x) = f(0) + o(1)$$

In this class, I'll occasionally use little-O notation. ("o" is for "order".)

Definition 1.2 *We say that* $f(x) = o(g(x))$ *if* $\lim_{x\to 0} f(x)/g(x) = 0$

"When x is small, $f(x)$ is much smaller than $g(x)$." In any case, if $f(x)$ is continuous at 0, and x is near 0, then "$f(0)$ is a good approximation for $f(x)$". We'll call this a "zeroth-order approximation".

Exercise 1.1 *If* $f = o(g)$ *and* a *is any non-zero real number, then* $f = o(ag)$. *If* $f = o(g)$ *and* $g = o(h)$, *then* $f = o(h)$.

Exercise 1.2 *If* $f(x)$ *is any function, not necessarily continuous, then there's at most one real number* a_0 *such that*

$$f(x) = a_0 + o(1)$$

Pictorially, we have a graph of a function, and we're approximating it by a horizontal line.

Well, we can do a little better. Many a function is not particularly close to being constant, but its rate of change doesn't vary too much, and so we can approximate the function by a straight line that's not necessarily horizontal. How can we find this line? Well, $f(x) - f(0) = o(1)$ is roughly a straight line going through the origin. If $f(x) - f(0) \approx ax$, then $a \approx (f(x) - f(0))/x$, so we define

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x}$$

Then f is *differentiable at* 0 if $f(x) = f(0) + f'(0)\,x + o(x)$.

Given a function $f(x)$, we define the *derivative* $f'(x)$ of f by creating a new function $g_x(h) = f(x+h)$ and letting $f'(x) = g'_x(0)$, where the right-hand-side is as in the above definition. I.e. we shift the function over, take its derivative, and shift back. So

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x+0)}{h}$$

just as you learned in calculus.

Definition 1.3 *A function* $f(x)$ *is* differentiable *at* x *if the* derivative $f'(x)$ *exists:*

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x+0)}{h}$$

If f *is differentiable in a neighborhood of* x *(an open interval containing* x*), then* $f'(x)$ *is a function defined in that neighborhood, and its derivative at* x*, or in a whole neighborhood of* x*, might exist. If it does, the function* f *is called* twice differentiable at x *or* twice differentiable near (in a neighborhood of) x. *Similarly, we have functions that are* k-times differentiable near x, *etc.*

Exercise 1.3 *If $f(x)$ is any function such that $f(x) = a_0 + a_1 x + o(x)$, then $a_1 = f'(0)$.*

We might want to approximate functions more closely, e.g. by nth degree polynomials.

Exercise 1.4 *Given a function $f(x)$, there is at most one polynomial $a(x) = a_0 + a_1 x + \ldots + a_n x^n$ of degree n such that*

$$f(x) = a(x) + o(x^n)$$

Definition 1.4 *The nth* Taylor polynomial *of f is the degree-n polynomial $a(x)$, if it exists, such that*

$$f(x) = a(x) + o(x^n)$$

We talked about how to define the derivative $f'(x)$ of a function $f(x)$, which may or may not exist. Then we can define the *kth derivative* by induction: define $f^{(0)}(x) = f(x)$, and $f^{(k)}(x) = [f^{(k-1)}]'(x)$. Of course, some functions don't have derivatives, and so it's possible that $f^{(k)}(x)$ is not defined for some k (and hence not for any higher k). Indeed, for any given x, $f^{(k)}(x)$ is not defined unless $f^{(k-1)}$ exists in a neighborhood of x.

Exercise 1.5 *Let $f(x)$ be differentiable and $f(0) = 0$. Show that if $f'(x) = o(x^n)$, then $f(x) = o(x^{n+1})$. The converse is harder: if f' is continuous in a neighborhood of 0 and $f(x) = o(x^n)$, then $f'(x) = o(x^{n-1})$.*

Exercise 1.6 Taylor's theorem: *Let $f(x)$ be n-times differentiable. Then $f(x) = a_0 + a_1 x + \ldots + a_n x^n + o(x^n)$ where $a_k = f^{(k)}(0)/k!$. (Hint/Outline: If f is differentiable, by definition $f(x) = a_0 + a_1 x + o(x)$. Use the previous exercise to show that if $f'(x) = a(x) + o(x^k)$ for some polynomial $b(x)$ of degree k, then there is a polynomial $B(x)$ of degree $k+1$ such that $f(x) = b(x) + o(x^{k+1})$. Hence, by induction, show that if f is n-times differentiable, then $f(x) = a(x) + o(x^n)$ for some polynomial $a(x)$ of degree n. Either by finding a formula relating $b(x)$ and $B(x)$ above, or using the previous exercise, prove the formula for the Taylor coefficients.)*

Hence even if the inductive definition of the kth derivative fails (because, say, derivatives of f don't exists throughout a neighborhood of x), we can define $f^{(k)}(0)$ in terms of the kth Taylor polynomial of f.

Exercise 1.7 *Here's an alternate description of the coefficients of the Taylor polynomial:*

1. *Define $f^{[0]}(x) = f(x)$, and define $f^{[k]}(x)$ by induction:*

$$f^{[k]}(x) = \begin{cases} \dfrac{f^{[k-1]}(x) - f^{[k-1]}(0)}{x}, & x \neq 0 \\ \lim\limits_{x \to 0} \dfrac{f^{[k-1]}(x) - f^{[k-1]}(0)}{x}, & x = 0 \end{cases}$$

Show that if $f(x) = o(x^n)$ then $f^{[k]}(x)$ exists for $k \leq n$.

2. *Show that if* $f(x) = a_0 + a_1 x + \ldots + a_n x^n + o(x^n)$ *is continuous, then* $f^{[k]}(x)$ *exists for* $k \le n$*, and that* $a_k = f^{[k]}(0)$*.*

Let's say a function $f(x)$ has Taylor polynomials at all orders:

$$\begin{aligned} f(x) &= a_0 + o(1) \\ &= a_0 + a_1 x + o(x) \\ &= a_0 + a_1 x + a_2 x^2 + o(x^2) \\ &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + o(x^3) \\ &= \ldots \end{aligned}$$

Then we should think of f as a *series*:

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

Of course, the right-hand-side might not converge for any x.

Definition 1.5 *A* formal function *or* formal power series *is a sequence of numbers, thought of as a series.*

Given a function, its *Taylor series* is the formal power series encoded in its Taylor polynomials. Taylor series are unique if they exist.

Exercise 1.8 *Use Exercise 1.6 to find the Taylor series for* e^x*, using the fact that the derivative of* e^x *is* e^x*. Use Exercise 1.7 to find the Taylor series for* $1/(1-x)$*.*

Since we have the product and chain rules, derivatives and Exercise 1.6 are the best way to compute Taylor series. But there's a standard ambiguity in the naming of the coefficients. By Exercise 1.6, for an ∞-times differentiable function, the Taylor series of $f(x)$ is $\sum_{n=0}^{\infty} f^{(n)}(0) x^n/n!$. But by Exercise 1.7, this sum is also $\sum_{n=0}^{\infty} f^{[0]}(0)\, x^n$. So which of $f^{(n)}(0)$ and $f^{[n]}(0)$ deserves to be called "the nth Taylor coefficient"? In this class we will use both, adopting the following notation: if $f(x)$ is a formal function, then the coefficients of the power series are

$$f(x) = \sum_{n=0}^{\infty} f_{(n)} \frac{x^n}{n!} = \sum_{n=0}^{\infty} f_{[n]}\, x^n$$

Definition 1.6 *Let* $f(x) = \sum f_{(n)}\, x^n/n! = \sum f_{[n]}\, x^n$ *and* $g(x) = \sum g_{(n)}\, x^n/n! = \sum g_{[n]}\, x^n$ *be formal functions. The* sum *and* product *are given by*

$$(f+g)_{(n)} = f_{(n)} + g_{(n)} \tag{1}$$

$$(f+g)_{[n]} = f_{[n]} + g_{[n]} \tag{2}$$

$$(f \cdot g)_{(n)} = \sum_{k=0}^{n} \binom{n}{k} f_{(k)} \cdot g_{(n-k)} \tag{3}$$

$$(f \cdot g)_{[n]} = \sum_{k=0}^{n} f_{[k]} \cdot g_{[n-k]} \tag{4}$$

The derivative *of f is given by*

$$(f')_{(n)} = f_{(n+1)} \qquad\qquad (f')_{[n]} = (n+1) f_{[n+1]} \tag{5}$$

Exercise 1.9 *Justify the above formulas. (Be sure to show that they are consistent: we gave two formulas for each concept.)*

We can think of a formal function f as a collection of diagrams. Each diagram consists of a vertex labeled by x with a bunch of upward "incoming" strings and one downward "outgoing" string. Then $f_{(n)}x^n$, or maybe $f_{[n]}x^n$, corresponds to a vertex with n incoming string, each labeled by x — we could say that the "value" of the diagram with n upward edges is $f_n x^n$. With this picture, the multiplication rules have a nice combinatorial interpretation. $(f \cdot g)_{(n)}$ or $(f \cdot g)_{[n]}$ should measure how to take n strings labeled x, put some of them through an f vertex and the rest through a g vertex, and then multiply the outputs. To get $(f \cdot g)_{(n)}$ or $(f \cdot g)_{[n]}$, we add up all the ways of doing this, although just like the $\sum$ in f is replaced with "collection of", we can even think of $(f \cdot g)_{(n)}$ or $(f \cdot g)_{[n]}$ as the *collection* of all ways to combine n xs, an f vertex, and a g vertex. If we use [] convention, then we're taking only those ways in which the first many xs go to f and the last many go to g. If we use the () convention, then we're taking all ways, with xs potentially crossing. Given a diagram, its value in the () convention divided by its value in the [] convention is the total number of symmetries of the diagram.

Exercise 1.10 *Show that the formal power series* 0 *given by* $0_{[n]} = 0_{(n)} = 0$ *is an additive identity of formal power series. Explain how to subtract formal power series. Show that the formal power series* 1 *given by* $1_{(0)} = 1_{[0]} = 1$ *and* $1_{(n)} = 1_{[n]} = 0$ *for* $n \geq 1$ *is a multiplicative identity.*

Exercise 1.11 *Using Exercise 1.8, show that the Taylor series f for e^x satisfies $f' = f$, and that the Taylor series g for $1/(1-x)$ satisfies $(1-x) \cdot g = 1$.*

Exercise 1.12 *Let f be a formal function. Find a formula for the coefficients of the nth power f^n of f; find formulas using each notion of coefficients.*

Exercise 1.13 *If f is a formal power series with $f_{(0)} = f_{[0]} \neq 0$, find a formula for the coefficients of f^{-1}, the multiplicative inverse of f. (Hint: Long division works.)*

Day 2: Chain Rule

We saw last time that some functions can be made into *formal functions*. Recall that a formal function is a sequence of numbers, thought of as a power series, and we make no guarantees about convergence. Taylor's Theorem tells us that an infinitely differentiable function gives a formal function $f(x) = \sum_{n=0}^{\infty} f_{(n)} x^n/n!$ where $f_{(n)} = f^{(n)}(0)$ is the nth derivative of f at 0. Two particularly important formal functions are

$$\begin{aligned} \exp(x) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ \frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n \end{aligned}$$

Each of these can reasonably called the formal function corresponding to the sequence of all 1s, depending on the convention of coefficients. Preferring agnosticism, we adopt the notation that $f(x) = \sum f_{(n)} x^n/n! = \sum f_{[n]} x^n$. Later we might argue that the () convention is better than the [] convention, because it performs more simply under calculus operations. For example, we interpret $f_{(n)}$ as the nth derivative of f at 0, and the derivative of f is given by the shift operation in the () convention:

$$(f')_{(n)} = f_{(n+1)}$$

We talked last time about how to add and multiply formal functions.

Exercise 2.1 *If f and g are honest functions with $f(x) = a(x) + o(x^n)$ and $g(x) = b(x) + o(x^n)$ where a and b are polynomials, show that $f(x)\,g(x) = a(x)\,b(x) + o(x^n)$.*

Exercise 2.2 *Prove that if f and g are differentiable in a neighborhood of 0, then so is the product $f \cdot g$. Argue by induction that if f and g are n-times differentiable in a neighborhood of 0, then so is the product $f \cdot g$.*

Exercise 2.3 *By approximating f and g by Taylor polynomials of degree n, prove the generalized product rule:*

$$\frac{d^n}{dx^n}\left[f(x)\,g(x)\right]\Big|_{x=0} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(0)\,g^{(n-k)}(0) \tag{6}$$

Definition 2.1 *If $m_1, \ldots, m_k$ are non-negative integers with $m_1 + \ldots + m_k = n$, then the* multinomial coefficient *is the number of ways of sorting n labeled objects into k boxes, such that the jth box has m_j numbers in it. The formula is given by*

$$\binom{n}{m_1, \ldots, m_k} = \binom{n}{\vec{m}} = \frac{n!}{m_1! \ldots m_k!}$$

If $k = n = 0$, so that the string $\vec{m} = (m_1, \ldots, m_k)$ is the empty string, then we declare that

$$\binom{0}{} = 1$$

Exercise 2.4 *Generalize the product rule further: Let $f_1, f_2, \ldots, f_k$ be k different smooth (infinitely differentiable) functions. Show that the nth derivative of the product is given by*

$$\frac{d^n}{dx^n}\left[\prod_{j=1}^{k} f_j\right]\Bigg|_{x=0} = \sum_{\substack{\vec{m}\geq 0 \ s.t.\\ \sum m_j = n}} \binom{n}{\vec{m}} \prod_{j=1}^{k} f^{(m_j)}(0) \tag{7}$$

The sum ranges over strings of length k of nonnegative integers such that the sum of the string is n.

In addition to addition and multiplication, functions related to each other through another binary operation: composition. What's the composition of two formal functions? We manipulate the summation notation uncritically:

$$\begin{aligned} f(g(x)) &= \sum_{k=0}^{\infty} f_{[k]} \left(\sum_{m=0}^{\infty} g_{[m]} x^m\right)^k \\ &= \sum_{k=0}^{\infty} f_{[k]} \sum_{m_1, m_2, \ldots, m_k} g_{[m_1]} g_{[m_2]} \cdots g_{[m_k]} x^{m_1+m_2+\ldots+m_k} \\ &= \sum_{\vec{m}} f_{[|m|]} x^{\sum m_j} \prod_{j=1}^{|\vec{m}|} g_{[m_j]} \end{aligned}$$

In the last line, the sum ranges over all vectors $\vec{m} = (m_1, m_2, \ldots, m_k)$ of any (finite) length k, made with non-negative integers. We set $|\vec{m}| = k$ and $\sum m_j = \sum_{j=1}^{k} m_k$. Combining like terms we have

$$f(g(x)) = \sum_{n=0}^{\infty} x^n \sum_{\substack{\vec{m}\geq 0 \text{ s.t.}\\ \sum m_j = n}} f_{[|\vec{m}|]} \prod_{j=1}^{|\vec{m}|} g_{[m_j]}$$

Thus, if we believe the uncritical computation above, we see that

$$(f \circ g)_{[n]} = \sum_{\substack{\vec{m}\geq 0 \text{ s.t.}\\ \sum m_j = n}} f_{[|\vec{m}|]} \prod_{j=1}^{|\vec{m}|} g_{[m_j]} \tag{8}$$

Exercise 2.5 *Find a formula for $(f \circ g)_{(n)}$ in terms of $f_{(j)}$ and $g_{(k)}$.*

Equation (8) is very problematic. The terms m_j might be 0, and so for any given n, there are infinitely many vectors which sum to n. One potential fix is to hope that the infinite sum converges, but this goes against the philosophy of using formal functions. On the other hand, if $g_{[m]} = 0$ for some m, then any vector $\vec{m}$ with $m_j = m$ for some j has $\prod_{j=1}^{|\vec{m}|} g_{[m_j]} = 0$. Since those terms vanish, we can restrict attention to only those vectors $\vec{m}$ with $g_{[m_j]} \neq 0$ for every j. Well, if $g_{[0]} = 0$, then the problem of infinite sums disappears: for any n, there are only finitely many vectors $\vec{m}$ of *positive* integers such that $\sum m_j = n$.

Definition 2.2 *If f and g are formal functions with $g_{[0]} = 0$, then the* composition $f \circ g$ *is defined by*

$$(f \circ g)_{[n]} = \sum_{\substack{\vec{m}>0 \text{ s.t.} \\ \sum m_j = n}} f_{[|\vec{m}|]} \prod_{j=1}^{|\vec{m}|} g_{[m_j]} \tag{9}$$

The sum ranges over vectors of positive integers of any length. $\sum m_j$ is the sum of the terms, and $|\vec{m}|$ is the number of terms.

We interpret a vector $\vec{m}$ with $\sum m_j = n$ as a(n ordered) *partition* of n: it breaks up a set of n elements into a set with m_1 elements, a set with m_2 elements, a set with m_3 elements, and so on. Then the interpretation of the formula is clear. We think of f_n as the way that n xs combine via an f into a number. Then $(f \circ g)_n$ should consist of the way that n xs combine into some gs, and those combine into an f. Indeed, the formula (9) bears this out: we take the n xs, and put the first m_1 into a g, then the next m_2 into a g, and so on, and then take those $|\vec{m}|$ gs and put them into an f.

Exercise 2.6 *When $g_0 = 0$, your formula in Exercise 2.5 is well-defined. Interpret it combinatorially.*

Exercise 2.7 *Define* $\ln(1 + x)$ *to be for formal power series with* $\ln_{[0]} = 0$ *and* $\ln_{[n]} = (-1)^{n-1}/n$. *Compute* $\exp(\ln(1 + x))$. *Also,* $\exp(x) - 1$ *has* $(\exp - 1)_0 = 0$. *Compute* $\ln(1 + (\exp(x) - 1))$.

We think of $f_{(n)} = n! f_{[n]}$ as the nth derivative of f at 0 — indeed, Taylor's theorem makes this precise. Thus, taking derivatives of a composition of analytic functions:

$$\left. \frac{d^n}{dx^n} [f(g(x))] \right|_{x=0} = (f \circ g)_{(n)}$$

Thus, Equation (9) describes how to take derivatives at 0, in the case when $g_0 = 0$. By translating, we have $(f(x + a))_{(n)} = f^{(n)}(a)$. Then

$$\begin{aligned} \left. \frac{d^n}{dx^n} [f(g(x))] \right|_{x=a} &= \left(\tilde{f} \circ \tilde{g} \right)_{(n)} \\ \text{where } \tilde{f}(x) &= f(x + g(a)) \\ \text{and } \tilde{g}(x) &= g(x + a) - g(a) \end{aligned}$$

Exercise 2.8 *Check this.*

Thus, since $(g(x)+a)_n = g_{(n)}$ for any constant a and $n \geq 0$, we have The Generalized Chain Rule:

Theorem 2.1 (Generalized Chain Rule) *If $f(x)$ and $g(x)$ are n-times differentiable functions everywhere, then*

$$\frac{d^n}{dx^n}\left[f(g(x))\right]\bigg|_{x=a} = \sum_{\substack{\vec{m}>0 \ s.t. \\ \sum m_j = n}} \frac{1}{|\vec{m}|!}\binom{n}{\vec{m}} \frac{d^{|\vec{m}|}f}{dx^{|\vec{m}|}}\bigg|_{x=g(a)} \prod_{j=1}^{|\vec{m}|} \frac{d^{m_j}g}{dx^{m_j}}\bigg|_{x=a} \tag{10}$$

This formula was first discovered by Francesco Faà di Bruno. A string $\vec{m}$ with $\sum m_j = n$ we interpret as an ordered partition of a set of size n, so that the first m_1 objects go into the first box, etc. Then the multinomial coefficient counts how many isomorphic (ordered) partitions there are allowing objects to go into any box. The $|\vec{m}|!$ adjusts for the fact that many ordered partitions are isomorphic as unordered partitions.

Exercise 2.9 *Another way to describe a partition is by counting how many boxes of each size. If $\vec{b} = (b_1, b_2, \ldots, b_n)$ is a string of non-negative integers such that $b_1 + 2b_2 + \ldots + nb_n = n$ and $b_1 + b_2 + \ldots + b_n = k$, then we can construct a partition in the previous sense by setting $m_1 = m_2 = \ldots = m_{b_1} = 1$, and in general setting the next b_j boxes to size j. Prove the following alternate formula for the chain rule (this was the one originally found by Faà di Bruno):*

$$\frac{d^n}{dx^n}\left[f(g(x))\right]\bigg|_{x=a} = \sum_{\substack{\vec{b}\geq 0 \ s.t. \\ \sum jb_j = n}} \frac{n!}{1!^{b_1}\, b_1!\, 2!^{b_2}\, b_2! \, \ldots \, n!^{b_n}\, b_n!} f^{(\sum b_j)}(g(a)) \prod_{j=1}^{n} \left(g^{(j)}(a)\right)^{b_j} \tag{11}$$

Here the sum ranges over strings $(b_1, \ldots, b_n)$ of non-negative integers with $b_1 + 2b_2 + \ldots + nb_n = n$. One really ought to come up with good notation for that fraction of factorials.

Day 3: Feynman Diagrams

I've been skirting an issue of what to call the nth term in a Taylor series, and so my goal for today will be to explain what I think is the best convention. In doing so, I'll try to explain exactly what a Feynman diagram is, and how to work with it.