1.2 The Calculus You Need

The prerequisite for differential equations is calculus. This may mean a year or more of ideas and homework problems and rules for computing derivatives and integrals. Some of those topics are essential, but others (as we all acknowledge) are not really of first importance. These pages have a positive purpose, to bring together essential facts of calculus. This section is to read and refer to—it doesn't end with a Problem Set.

I hope this outline may have value also at the end of a single-variable calculus course. Textbooks could include a summary of the crucial ideas, but usually they don't. Certainly the reader will not agree with every choice made here, and the best outcome would be a more perfect list. This one is a lot shorter than I expected.

At the end, a useful formula in differential equations is confirmed by the product rule, the derivative of e^x, and the Fundamental Theorem of Calculus.

1. Derivatives of key functions: $\boldsymbol{x^n}$ $\quad$ $\boldsymbol{\sin x}$ $\quad$ $\boldsymbol{\cos x}$ $\quad$ $\boldsymbol{e^x}$ $\quad$ $\boldsymbol{\ln x}$

The derivatives of $x, x^2, x^3, \ldots$ come from first principles, as limits of $\Delta y/\Delta x$. The derivatives of $\sin x$ and $\cos x$ focus on the limit of $(\sin \Delta x)/\Delta x$. Then comes the great function e^x. It solves the differential equation $dy/dx = y$ starting from $y(0) = 1$. **This is the single most important fact needed from calculus: the knowledge of $\boldsymbol{e^x}$.**

2. Rules for derivatives: Sum rule Product rule Quotient rule Chain rule

When we add, subtract, multiply, and divide the five original functions, these rules give the derivatives. The sum rule is the quiet one, applied all the time to *linear* differential equations. This equation is linear (*a crucial property*):

$$\frac{dy}{dt} = ay + f(t) \quad \text{and} \quad \frac{dz}{dt} = az + g(t) \quad \text{add to} \quad \frac{d}{dt}(\boldsymbol{y}+\boldsymbol{z}) = a(\boldsymbol{y}+\boldsymbol{z}) + (\boldsymbol{f}+\boldsymbol{g}).$$

With $a = 0$ that is a straightforward sum rule for the derivative of $y + z$. We can always add equations as shown, because $a(t)y$ is linear in y. This confirms *superposition* of the separate solutions y and z. Linear equations add and their solutions add.

The chain rule is the most prolific, in computing the derivatives of very remarkable functions. The chain $\boldsymbol{y = e^x}$ and $\boldsymbol{x = \sin t}$ produces $\boldsymbol{y = e^{\sin t}}$ (the composite of two functions). The chain rule gives dy/dt by multiplying the derivatives dy/dx and dx/dt:

$$\textbf{Chain rule} \qquad \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = e^x \cos t = y \cos t.$$

Then $e^{\sin t}$ solves that differential equation $\dfrac{dy}{dt} = ay$ with varying growth rate $a = \cos t$.

3. The Fundamental Theorem of Calculus

The derivative of the integral of $f(x)$ is $f(x)$. The integral from 0 to x of the derivative df/dx is $f(x) - f(0)$. One operation inverts the other, when $f(0) = 0$. This is not so easy to prove, because both the derivative and the integral involve a limit step $\Delta x \to 0$.

One way to go forward starts with numbers $y_0, y_1, \ldots, y_n$. Their differences are like derivatives. Adding up those differences is like integrating the derivative:

$$\textbf{Sum of differences } (y_1-y_0)+(y_2-y_1)+\cdots+(y_n-y_{n-1})=y_n-y_0. \tag{1}$$

Only y_n and $-y_0$ are left because all other numbers $y_1, y_2, \ldots$ come twice and cancel. To make that equation look like calculus, multiply every term by $\Delta x/\Delta x = 1$:

$$\left[\frac{y_1 - y_0}{\Delta x} + \frac{y_2 - y_1}{\Delta x} + \cdots + \frac{y_n - y_{n-1}}{\Delta x}\right] \Delta x = y_n - y_0. \tag{2}$$

Again, this is true for all numbers $y_0, y_1, \ldots, y_n$. Those can be heights of the graph of a function $y(x)$. The points $x_0, \ldots, x_n$ can be equally spaced between $x = a$ and $x = b$. Then each ratio $\Delta y/\Delta x$ is a *slope* between two points of the graph:

$$\frac{\Delta y}{\Delta x} = \frac{y_k - y_{k-1}}{x_k - x_{k-1}} = \frac{\text{distance up}}{\text{distance across}} = \text{slope.} \tag{3}$$

This slope is exactly correct if the graph is a straight line between the points x_{k-1} and x_k. If the graph is a curve, the approximate slope $\Delta y/\Delta x$ becomes exact as $\Delta x \to 0$.

The delicate part is the requirement $n\Delta x = b - a$, to space the points evenly from $x_0 = a$ to $x_n = b$. Then n will increase as Δx decreases. Equation (2) remains correct at every step, with $y_0 = y(a)$ at the first point and $y_n = y(b)$ at the last point. As $\Delta x \to 0$ and $n \to \infty$, the slopes $\Delta y/\Delta x$ approach the derivative dy/dx. At the same time the sum approaches the integral of dy/dx. Equation (2) turns into equation (4):

Fundamental Theorem of Calculus	$\displaystyle\int_a^b \frac{dy}{dx}\, dx = y(b) - y(a)$	$\displaystyle\frac{d}{dx}\int_a^x f(s)\, ds = f(x)$	(4)

The limits of $\Delta y/\Delta x$ in (3) and the sum in (2) produce dy/dx and its integral. Of course this presentation of the Fundamental Theorem needs more careful attention. But equation (1) holds a key idea: *a sum of differences*. This leads to *an integral of derivatives.*

4. The meaning of symbols and the operations of algebra

Mathematics is a language. The way to learn this language is to use it. So textbooks have thousands of exercises, to practice reading and writing symbols like $y(x)$ and $y(x + \Delta x)$. Here is a typical line of symbols:

$$\textbf{Derivative of } \boldsymbol{y} \qquad \frac{dy}{dt}(t) = \lim_{\Delta t \to 0} \frac{y(t + \Delta t) - y(t)}{\Delta t}. \tag{5}$$

I am not very sure that this is clear. One function is $\boldsymbol{y}$, the other function is its derivative y'.

Could the symbol y' be better than dy/dt? Both are standard in this book. In calculus we know $y(t)$, in differential equations we don't. The whole point of the differential equation is to connect y and y'. From that connection we have to discover what they are.

A first example is $y' = y$. That equation forces the unknown function y to grow exponentially: $y(t) = Ce^t$. At the end of this section I want to propose a more complicated equation and its solution. But I could never find a more important example than e^t.

5. Three ways to use $\quad dy/dx \approx \Delta y/\Delta x$

On the graph of a function $y(x)$, the exact slope is dy/dx and the approximate slope (between nearby points) is $\Delta y/\Delta x$. If we know *any two* of the numbers dy/dx and Δy and Δx, then we have a good approximation to the third number. All three approximations are important, because dy/dx is such a central idea in calculus.

(A) When we know Δx and dy/dx, we have $\Delta y \approx (\Delta x)(dy/dx)$.

This is linear approximation. From a starting point x_0, we move a distance Δx. That produces a change Δy. The graph of $y(x)$ can go up or down, and the best information we have is the slope dy/dx at x_0. (That number gives no way to account for *bending* of the graph, which appears in the next derivative d^2y/dx^2.)

Linear approximation is equivalent to following the tangent line —not the curve:

$$\Delta y \approx \Delta x \, \frac{dy}{dx} \qquad y(x_0 + \Delta x) \approx y(x_0) + \Delta x \frac{dy}{dx}(x_0) \tag{6}$$

(B) Δy and dy/dx lead to $\Delta x \approx (\Delta y)/(dy/dx)$. This is Newton's Method.

Newton's Method is a way to solve $y(x) = 0$, starting at a point x_0. We want $y(x)$ to drop from $y(x_0)$ to zero at the new point x_1. *The desired change in y is $\Delta y = 0 - y(x_0)$.* What we don't know is Δx, which locates x_1. The exact slope dy/dx will be close to $\Delta y/\Delta x$, and that tells us a good Δx:

$$\textbf{Newton's Method} \qquad \Delta x \approx \frac{\Delta y}{dy/dx} \qquad x_1 - x_0 = \frac{-y(x_0)}{dy/dx(x_0)} \tag{7}$$

Guess x_0, improve to x_1. This is an excellent way to solve nonlinear equations $y(x) = 0$.

(C) Dividing Δy by Δx gives the approximation $dy/dx \approx \Delta y/\Delta x$.

That is the point of equation (5), but something important often escapes our attention. *Are x and $x + \Delta x$ the best two places to compute y?* Writing $\Delta y = y(x + \Delta x) - y(x)$ doesn't seem to offer other choices. If we notice that Δx can be negative, this allows $x + \Delta x$ to be on the left side of x (leading to a backward difference). The best choice is not forward or backward but *centered around x: a half step each way.*

$$\textbf{Centered difference} \qquad \frac{dy}{dx} \approx \frac{\Delta y}{\Delta x} = \frac{y(x + \frac{1}{2}\Delta x) - y(x - \frac{1}{2}\Delta x)}{\Delta x} \tag{8}$$

Why is centering better? When $y = Cx + D$ has a straight line graph, all ratios $\Delta y/\Delta x$ give the correct slope C. But the parabola $y = x^2$ has the simplest possible bending, and **only this centered difference gives the correct slope $2x$** (varying with x).

Exact slope for parabolas by centering $$\frac{\Delta y}{\Delta x} = \frac{(x+\frac{1}{2}\Delta x)^2 - (x-\frac{1}{2}\Delta x)^2}{\Delta x} = \frac{x\,\Delta x - (-x\,\Delta x)}{\Delta x} = \mathbf{2x}$$

The key step in scientific computing is improving first order accuracy (forward differences) to second order accuracy (centered differences). For integrals, rectangle rules improve to trapezoidal rules. This is a big step to good algorithms.

6. Taylor series: Predicting $y(x)$ from all the derivatives at $x = x_0$

From the height y_0 and the slope y_0' at x_0, we can predict the height $y(x)$ at nearby points. But the tangent line in equation (6) assumes that $y(x)$ has constant slope. That first order prediction becomes a second order prediction (*much more accurate*) when we use the second derivative y_0'' at x_0.

Tangent parabola using y_0'' $$y(x_0 + \Delta x) \approx y_0 + (\Delta x)y_0' + \tfrac{1}{2}(\Delta x)^2 y_0''. \qquad (9)$$

Adding this $(\Delta x)^2$ term moves us from constant slope to constant bending. For the parabola $y = x^2$, equation (9) is exact: $(x_0 + \Delta x)^2 = (x_0^2) + (\Delta x)(2x_0) + \frac{1}{2}(\Delta x)^2(2)$.

Taylor added more terms—infinitely many. His formula gets *all derivatives correct* at x_0. The pattern is set by $\frac{1}{2}(\Delta x)^2 y_0''$. The n^{th} derivative $y^{(n)}(x)$ contributes a new term $\frac{1}{n!}(\Delta x)^n y_0^{(n)}$. The complete Taylor series includes all derivatives at the point $x = x_0$:

Taylor series $\quad y(x_0 + \Delta x) = y_0 + (\Delta x)\,y_0' + \cdots + \dfrac{1}{n!}(\Delta x)^n\,y_0^{(n)} + \cdots$

Stop at y' for tangent line

Stop at y'' for parabola $$= \sum_{n=0}^{\infty} \frac{(\Delta x)^n}{n!}\, y^{(n)}(x_0) \qquad (10)$$

Those equal signs are not always right. There is no way we can stop $y(x)$ from making a sudden change after x moves away from x_0. Taylor's prediction of $y(x_0 + \Delta x)$ is exactly correct for e^x and $\sin x$ and $\cos x$—good functions like those are "analytic" at all x.

Let me include here the two most important examples in all of mathematics. They are solutions to $dy/dx = y$ and $dy/dx = y^2$ — the most basic linear and nonlinear equations.

Exponential series with $\boldsymbol{y^{(n)}(0) = 1}$ $$y = e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots \qquad (11)$$

Geometric series with $\boldsymbol{y^{(n)}(0) = n!}$ $$y = \frac{\mathbf{1}}{\mathbf{1-x}} = 1 + x + x^2 + x^3 + \cdots \qquad (12)$$

The center point is $x_0 = 0$. The series (11) gives e^x for every x. The series (12) gives $1/(1-x)$ when x is between -1 and 1. Its derivative $1 + 2x + 3x^2 + \cdots$ is $1/(1-x)^2$.

For $x = 2$ that geometric series will certainly not produce $1/(1-2) = -1$. Notice that $1 + x + x^2 + \cdots$ becomes infinite at $x = 1$, exactly where $1/(1-x)$ becomes $1/0$.

The key point for e^x is that its n^{th} derivative is 1 at $x = 0$. The n^{th} derivative of $1/(1-x)$ is $n!$ at $x = 0$. This pattern starts with y, y', y'', y''' equal to $1, 1, 2, 6$ at $x = 0$:

$$y = (1-x)^{-1} \qquad y' = (1-x)^{-2} \qquad y'' = 2(1-x)^{-3} \qquad y''' = 6(1-x)^{-4}.$$

Taylor's formula combines the contributions of all derivatives at $x = 0$, to produce $y(x)$.

7. Application: An important differential equation

The linear differential equation $y' = ay + q(t)$ is a perfect multipurpose model. It includes the growth rate a and the external source term $q(t)$. We want the particular solution that starts from $y(0) = 0$. Creating that solution uses the most essential idea behind integration. Verifying that the solution is correct uses the basic rules for derivatives. Many students in my graduate class had forgotten the derivative of the integral.

Here is the solution $y(t)$ followed by its interpretation, with $a = 1$ for simplicity:

$$\frac{dy}{dt} = y + q(t) \qquad \textbf{is solved by} \qquad y(t) = \int_0^t e^{t-s} q(s)\, ds. \tag{13}$$

Key idea: At each time s between 0 and t, the input is a source of strength $q(s)$. That input grows or decays over the remaining time $t - s$. **The input $q(s)$ is multiplied by e^{t-s} to give an output at time t.** Then the total output $y(t)$ is the *integral* of $e^{t-s}q(s)$.

We will reach $y(t)$ in other ways. Section 1.4 uses an "integrating factor." Section 1.6 explains "variation of parameters." The key is to see where the formula comes from. *Inputs lead to outputs, the equation is linear, and the principle of superposition applies.* The total output is the sum (in this case, the integral) of all those outputs.

We will confirm formula (13) by computing dy/dt. First, e^{t-s} equals e^t times e^{-s}. Then e^t comes outside the integral of $e^{-s}q(s)$. Use the product rule on those two factors:

$$\textbf{Producing } y + q \qquad \frac{dy}{dt} = \left(\frac{de^t}{dt}\right) \int_0^t e^{-s} q(s)\, ds + \left(e^t\right) \frac{d}{dt} \int_0^t e^{-s} q(s)\, ds. \tag{14}$$

The first term on the right side is exactly $y(t)$. How to recognize that last term as $q(t)$?

We don't need to know the function $q(t)$. What we do know (and need) is the *Fundamental Theorem of Calculus*. **The derivative of the integral of $e^{-t}q(t)$ is $e^{-t}q(t)$.** Then multiplying by e^t gives the hoped-for result $q(t)$, because $e^t e^{-t} = 1$. The linear differential equation $y' = y + q$ with $y(0) = 0$ is solved by the integral of $e^{t-s}q(s)$.