

Operator methods for integrals and differential equations

Dan Piponi

October 23, 2017

In memory of Tom Mower who
started me on this journey more
than three decades ago

Abstract

Some unpolished notes on abusing the differential operator $D = \frac{d}{dx}$ for both symbolic and numeric integration and differential equation solving. Starts with some methods we were taught at high school(!) for solving differential equations which appear not to be well known among students today. Apologies for some inconsistencies in notation.

1 Warm up

1.1 A differential equation

Suppose we are give the differential equation:

$$\frac{df}{dx} + f = x^2$$

Let's use the shorthand $D = \frac{d}{dx}$ and rewrite this as

$$(D + 1)f = x^2$$

The standard approach is first to solve the homogeneous form of the equation

$$(D + 1)f = 0$$

giving $f(x) = A \exp(-x)$ for some constant A . We then have to find a *particular integral*, i.e. any solution to the original equation. The full solution to the original equation is the particular integral plus $A \exp(-x)$.

So how do we find a particular integral? In my experience students are encouraged to use educated guesswork. Anything goes as long as you prove that what you have found is indeed a solution. So in this case a popular approach might be to assume $f(x) = Ax^2 + Bx + C$, substitute into the equation, and solve for A , B and C .

Here's a more direct way:

$$\begin{aligned}(D + 1)f &= x^2 \\ f &= \frac{1}{1 + D} x^2 \\ &= (1 - D + D^2 - D^3 - \dots)x^2 \\ &= x^2 - 2x + 2\end{aligned}$$

We have performed two unusual operations:

1. we've "divided" by $1 + D$ and
2. we've expanded $(1 + D)^{-1}$ as a power series.

1.2 An integral

Another example. Suppose we wish to find the integral:

$$\int x^3 \exp(2x) dx$$

Integration is the “inverse” of differentiation so let’s rewrite this as

$$\frac{1}{D} x^3 \exp(2x)$$

Now we use a trick known as the shift rule. It states

$$f(D) \exp(ax) = \exp(ax) f(D + a)$$

So we can rewrite our integral as

$$\exp(2x) \frac{1}{D+2} x^3 \tag{1}$$

$$= \exp(2x) \frac{1}{2} \frac{1}{1+D/2} x^3 \tag{2}$$

$$= \exp(2x) \frac{1}{2} \left(1 - \frac{D}{2} + \left(\frac{D}{2}\right)^2 - \left(\frac{D}{2}\right)^3 + \dots\right) x^3 \tag{3}$$

$$= \frac{1}{8} \exp(2x) (4x^3 - 6x^2 + 6x - 3) \tag{4}$$

$$\tag{5}$$

We have used the same division and power series operations as above. And of course for the full answer we need to add a constant C .

My goal here is to sketch (without full rigour) why you might expect such methods to work and give many more examples. I also want to show how you can stretch these methods to give some hand-wavey arguments for some well-known theorems.

2 Some justification

2.1 Truncated power series

Let $P(x)$ be a polynomial in x . Suppose we can expand $P(x)^{-1}$ as a power series with some non-empty circle of convergence:

$$\frac{1}{P(x)} = \sum_{i=0}^{\infty} a_i x^i$$

We have that

$$P(x) \sum_{i=0}^{\infty} a_i x^i = 1$$

so

$$P(x) \left(\sum_{i=0}^{n-1} a_i x^i + \sum_{i=n}^{\infty} a_i x^i \right) = 1$$

and

$$P(x) \sum_{i=0}^{n-1} a_i x^i = 1 - P(x) \sum_{i=n}^{\infty} a_i x^i$$

The left hand side

$$P(x) \sum_{i=0}^{n-1} a_i x^i$$

is a polynomial. It must be 1 followed by terms of degree n or higher.

For example

$$\begin{aligned} (1+x)(1-x+x^2-x^3+x^4) &= 1-x+x^2-x^3+x^4-x^2+x^3-x^4+x^5 \\ &= 1+x^5 \end{aligned}$$

This tells us that we can use truncated power series to compute reciprocals of polynomials as long as we don't mind some higher order terms. The nice thing about this is that we know that for any polynomial $Q(x)$, $D^n Q(x) = 0$ for n large enough. So we can use truncated power series in D to obtain exact results when applied to polynomials in x .

Let me rewrite the first argument above in a more rigorous way:

$$\begin{aligned}(D+1)f &= x^2 \\ (1-D+D^2)(D+1)f &= (1-D+D^2)x^2 \\ 1f &= x^2 - 2x + 2 \text{ (assuming that } f \text{ has no terms beyond } x^2\text{)}\end{aligned}$$

We can see this as justifying the use of the operator $(D+1)^{-1}$ as long as we remember that the original argument is just a kind of shorthand for the rigorous one.

2.2 The shift rule

By the Leibniz rule we have

$$\frac{d}{dx}(\exp(ax)f(x)) = \exp(ax)\frac{df(x)}{dx} + a\exp(ax)f(x)$$

We can rewrite this as

$$\begin{aligned}D(\exp(ax)f(x)) &= \exp(ax)(Df(x) + af(x)) \\ &= \exp(ax)(D+a)f(x)\end{aligned}$$

We can write this even more compactly as

$$D\exp(ax) = \exp(ax)(D+a)$$

as long as we remember that in this version there is an implication that the $\exp(ax)$ on the left hand side is intended to be multiplied by some further expression to its right.

We can prove more.

$$\begin{aligned}D^n \exp(ax) &= D^{n-1} \exp(ax)(D+a) \\ &= D^{n-2} \exp(ax)(D+a)^2 \\ &\vdots \\ &= \exp(ax)(D+a)^n\end{aligned}$$

More generally we have The Shift Rule

$$\boxed{f(D)\exp(ax) = \exp(ax)f(D+a)}$$

for any polynomial f . We also expect this to hold in situations where f is a power series but we know it's being applied to a polynomial so terms beyond a certain point contribute zero.

2.3 Integration again

Now consider our integral above. We can consider this as a solution to the differential equation

$$Df = \exp(2x)x^3$$

Using the shift rule we have

$$\begin{aligned}Df &= \exp(2x)x^3 \\ \exp(2x)(D+2)\exp(-2x)f(x) &= \exp(2x)x^3 \\ (D+2)\exp(-2x)f(x) &= x^3\end{aligned}$$

Writing $g(x) = \exp(-2x)f(x)$ we can now use the differential equation solving methods above to solve $(D+2)g = x^3$ and our final integral is given by $\exp(2x)g(x)$. Again, we can see the original argument as being shorthand for this more rigorous argument.

3 Diagonalisation

The functions \sin , \cos and \exp “diagonalise” various differential operators. This means that differential operators act on these functions just like multiplication by some number. (That’s what *diagonalisation* means - finding elements that operators act on like real numbers.) For example $D \exp(ax) = a \exp(ax)$. But we also have $D^n \exp(ax) = a^n \exp(ax)$ and so

$$f(D) \exp(ax) = f(a) \exp(ax)$$

You can see this is a special case of the shift rule applied to $f(D)(\exp(ax) \times 1)$.

We also have $D^2 \sin(ax) = -a^2 \sin(ax)$ and $D^2 \cos(ax) = -a^2 \cos(ax)$. So for any polynomial f

$$\begin{aligned} f(D^2) \sin(ax) &= f(-a^2) \sin(ax) \text{ and} \\ f(D^2) \cos(ax) &= f(-a^2) \cos(ax). \end{aligned}$$

Example 3.1. Find a solution to

$$\frac{d^2 f}{dx^2} - 3 \frac{df}{dx} + 2f = \exp(3x)$$

Rewrite as

$$(D^2 - 3D + 2)f = \exp(3x)$$

so

$$\begin{aligned} f(x) &= (D^2 - 3D + 2)^{-1} \exp(3x) \\ &= (3^2 - 3 \times 3 + 2)^{-1} \exp(3x) \\ &= \frac{1}{2} \exp(3x) \end{aligned}$$

Example 3.2. Find a solution to

$$(D^2 + 1)f = \cos(2x)$$

We have

$$\begin{aligned} f(x) &= (D^2 + 1)^{-1} \cos(2x) \\ &= (-2^2 + 1)^{-1} \cos(2x) \\ &= -\frac{1}{3} \cos(2x) \end{aligned}$$

4 Lots of examples

Example 4.1. Find a solution to

$$\frac{df}{dx} + f = \sin x$$

Unfortunately we have an odd power of D applied to the \sin function so we can’t directly use the diagonalisation technique. Instead we write $\sin x$ using Euler’s formula:

$$\begin{aligned} \frac{1}{1+D} \sin x &= \frac{1}{1+D} \Im(\exp ix) \\ &= \Im\left(\frac{1}{1+D} \exp ix\right) \text{ (using fact that } \Im(D(f)) = D(\Im(f))) \\ &= \Im\left(\frac{1}{1+i} \exp ix\right) \text{ (using diagonalisation for } \exp) \\ &= \Im\left(\frac{1-i}{2} \exp ix\right) \\ &= \frac{1}{2}(\sin x - \cos x) \end{aligned}$$

Example 4.2. Find a solution to

$$\frac{d^3 f}{dx^3} - f = \sin x$$

We want

$$\frac{-1}{1 - D^3} \sin x$$

We can use the fact that $D^2 \sin x = -\sin x$ to reduce this to:

$$\frac{-1}{1 + D} \sin x$$

The solution is just minus the previous example

$$f(x) = \frac{1}{2}(\cos x - \sin x)$$

Example 4.3. Find a solution to

$$\frac{d^2 f}{dx^2} - 2\frac{df}{dx} + f = e^x$$

The only slight subtlety here is noticing that when we use the shift rule, we slide $f(D)$ past $\exp x$ leaving behind a 1 that needs to be integrated twice.

$$\begin{aligned} (D^2 - 2D + 1)f &= e^x \\ f &= \frac{1}{(D-1)^2} e^x \\ &= e^x \frac{1}{D^2} 1 \\ &= \frac{x^2}{2} e^x \end{aligned}$$

Example 4.4. Find a solution to

$$4\frac{d^2 f}{dx^2} + f = x \exp(-x)$$

Solution:

$$\begin{aligned} \frac{1}{1 + 4D^2} \exp(-x)x &= \exp(-x) \frac{1}{1 + 4(D-1)^2} x \\ &= \exp(-x) \frac{1}{4D^2 - 8D + 5} x \\ &= \frac{1}{5} \exp(-x) \frac{1}{1 - 8D/5} x \quad (\text{using } D^2 x = 0) \\ &= \frac{1}{5} \exp(-x) \left(1 + \frac{8}{5}D\right) x \\ &= \frac{1}{25} \exp(-x) (5x + 8) \end{aligned}$$

Example 4.5. Find a solution to

$$\frac{d^2 f}{dx^2} + f = x e^{-x} \sin(2x)$$

No avoiding some messy complex number arithmetic here. We want

$$\begin{aligned}
\frac{1}{D^2+1}e^{-x}\sin(2x)x &= \Im\left[\frac{1}{D^2+1}e^{(2i-1)x}x\right] \\
&= e^{-x}\Im\left[e^{2ix}\frac{1}{(D+2i-1)^2+1}x\right] \\
&= e^{-x}\Im\left[e^{2ix}\frac{-1}{2+4i}\frac{1}{1-\frac{2(2i-1)}{2+4i}D}x\right] \\
&= e^{-x}\Im\left[e^{2ix}\frac{-1}{2+4i}\left(x+\frac{2(2i-1)}{2+4i}\right)\right] \\
&= \frac{1}{50}e^{-x}\Im\left[(\cos(2x)+i\sin(2x))((10i-5)x-(11-2i))\right] \\
&= \frac{1}{50}e^{-x}((2+10x)\cos(2x)-(11+5x)\sin(2x))
\end{aligned}$$

Example 4.6. Find

$$\int_0^\infty x^n \exp(-x) dx$$

First the indefinite integral

$$\begin{aligned}
&\frac{1}{D}x^n \exp(-x) \\
&= \exp(-x)\frac{1}{D-1}x^n \\
&= -\exp(-x)(1+D+D^2+\dots+D^n)x^n
\end{aligned}$$

We're going to be evaluating this at zero and in the limit as x goes to infinity. All of these terms vanish at infinity. All of the non-constant derivatives of x^n vanish at zero. So we're left with

$$\left[-\exp(-x)D^n x^n\right]_0^\infty$$

This is $n!$.

Example 4.7. Find a solution to

$$\frac{df}{dx} + af(x) = g(x)$$

This becomes

$$\begin{aligned}
f(x) &= \frac{1}{a+D}g(x) \\
&= e^{-ax}\frac{1}{D}e^{ax}g(x) \\
&= e^{-ax}\int^x e^{ay}g(y)dy \\
&= \int_{-\infty}^x e^{a(y-x)}g(y)dy \text{ (choosing a particular integral)}
\end{aligned}$$

The value at x is essentially a weighted sum of the history of g before x . You can think of f as a “leaky” integral of g - it would be the integral but f “leaks” at a rate proportional to a and f .

5 Exponentials of D

Taylor's theorem tells us that

$$f(x+a) = f(x) + af'(x) + \frac{a^2}{2!}f''(x) + \frac{a^3}{3!}f^{(3)}(x) + \dots + \frac{a^n}{n!}f^{(n)}(x) + \text{a remainder term}$$

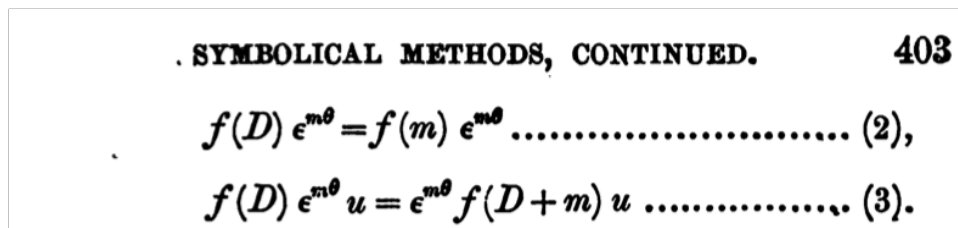


Figure 1: These techniques are old. This snippet is from Boole's *A Treatise on Differential Equations* from 1859.

The form of the remainder depends on the class of function f . For polynomials the remainder is precisely zero for n large enough. For analytic functions the remainder goes to zero as n goes to infinity so we can write

$$f(x + a) = \sum_{n=0}^{\infty} a^n \frac{D^n}{n!} (f)(x)$$

We can now write this as

$$f(x + a) = \left(\sum_{n=0}^{\infty} a^n \frac{D^n}{n!} \right) (f)(x)$$

and therefore as

$$f(x + a) = \exp(aD)(f)(x)$$

In other words, $\exp(aD)$ is the operator that shifts a function by a .

6 Sums and recurrence relations

Armed with exponentials of D we can extend our methods for integrals and differential equations to sums and recurrence relations.

Example 6.1. Find

$$\sum_{x=0}^{n-1} x^3$$

Write $f(x) = \sum_{y=0}^{x-1} y^3$. Then we want to solve

$$f(x + 1) - f(x) = x^3$$

We can now use the methods above:

$$\begin{aligned} \exp(D)f - f &= x^3 \\ (\exp(D) - 1)f &= x^3 \\ f &= \frac{1}{\exp(D) - 1} x^3 \\ f &= \frac{1}{D + \frac{1}{2}D^2 + \frac{1}{3!}D^3 + \frac{1}{4!}D^4 + \frac{1}{5!}D^5} x^3 \\ f &= \frac{1}{1 + \frac{1}{2}D + \frac{1}{3!}D^2 + \frac{1}{4!}D^3 + \frac{1}{5!}D^4} \int x^3 dx \\ f &= \frac{1}{1 + \frac{1}{2}D + \frac{1}{3!}D^2 + \frac{1}{4!}D^3 + \frac{1}{5!}D^4} \frac{1}{4} x^4 \end{aligned}$$

There are tricks we can use to minimise the work here although I'm going to be a bit more explicit than needed so everything is clear. Analogously to solving differential equations, we are going to end up with a "particular sum". The full solution is going to require determining a constant term. But it's clear that the sum needs to be zero for $x = 0$ so the constant term should be zero.

Applying D^4 to x^4 is going to give us a constant term. So we only need to keep terms up to D^3 . We get

$$\begin{aligned} f &= \frac{1}{4}(1 - (\frac{1}{2}D + \frac{1}{3!}D^2 + \frac{1}{4!}D^3) + (\frac{1}{2}D + \frac{1}{3!}D^2)^2 - (\frac{1}{2}D)^3)x^4 \\ &= \frac{1}{4}(1 - \frac{1}{2}D - \frac{1}{6}D^2 - \frac{1}{24}D^3 + \frac{1}{4}D^2 + \frac{1}{6}D^3 - \frac{1}{8}D^3)x^4 \\ &= \frac{1}{4}(1 - \frac{1}{2}D + \frac{1}{12}D^2)x^4 \text{ (nice for us, the cubic term vanishes)} \\ &= \frac{x^4}{4} - \frac{x^3}{2} + \frac{x^2}{4} \end{aligned}$$

This may seem borderline magical. One way to think about it is that if we know f is a degree 4 polynomial, then $f(x+1) = f(x) + f'(x) + \frac{1}{2}f''(x) + \frac{1}{3!}f^{(3)}(x) + \frac{1}{4!}f^{(4)}(x)$ exactly. So solving the summation is equivalent to solving the differential equation

$$\frac{1}{4!} \frac{d^4 f}{dx^4} + \frac{1}{3!} \frac{d^3 f}{dx^3} + \frac{1}{2} \frac{d^2 f}{dx^2} + \frac{df}{dx} = x^3$$

It is entirely reasonable to solve this using differential equation methods. If you extend this approach beyond polynomials you rediscover the Euler-Maclaurin summation formula.

Example 6.2. Solve

$$a_{n+2} = a_{n+1} + a_n + n^2 \text{ with } a_0 = 0, a_1 = 0$$

In this case the problem requires finding a spolution with specific initial conditions. So we're going to need both a "particular sum" and a solution to the homogeneous equation. It's well known that the solutions to the homogeneous equation are the Fibonacci numbers F_n and the Lucas numbers L_n . Any other solution is a linear combination of these. The Fibonacci numbers start with $F_0 = 0$ and $F_1 = 1$ and the Lucas numbers start with $L_0 = 2$ and $L_1 = 1$.

Let's write $a_x = f(x) + AF_x + BL_x$ so our notation matches what used earlier. We have

$$\begin{aligned} (e^{2D} - e^D - 1)f &= x^2 \\ f &= \frac{1}{e^{2D} - e^D - 1}x^2 \\ &= (-1 - D - \frac{5}{2}D^2)x^2 \\ &= -x^2 - 2x - 5 \end{aligned}$$

Because $F_0 = 0$, the initial condition at $x = 0$ immediately implies that B is $\frac{5}{L_0} = \frac{5}{2}$. The initial condition at $x = 1$ now gives $A = \frac{11}{2}$. The complete solution is

$$a_n = \frac{5}{2}L_n + \frac{11}{2}F_n - n^2 - 2n - 5$$

By the way, Mathematica makes a real mess of this. I've seen similar methods in a work at least a century old. Maybe Boole.

7 Numerical methods

7.1 Derivatives

The relation $(e^{hD} - 1)f(x) = f(x+h) - f(x)$ expresses a finite difference in terms of the differential operator. We can do this in reverse and derive formulae for derivatives in terms of finite differences. A good application of this is to derive approximations of derivatives that we can use in numerical methods on a grid. Define $E_h = e^{hD}$. We'd like to write D in terms of E_h . That seems straightforward. We expect

$$D = \frac{1}{h} \log E_h$$

The problem is, we can't simply apply a Taylor expansion to $\log E_h$ about 0. Suppose we'd like our numerical methods to work well on polynomials - typically giving exact results on low order

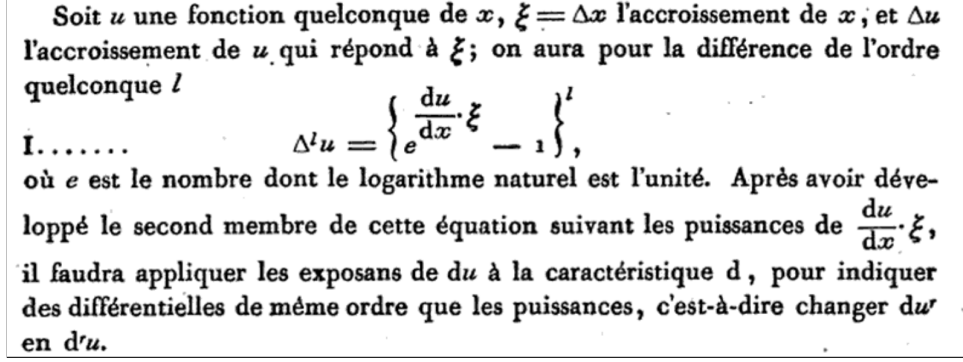


Figure 2: A snippet from Arbogast's *Du calcul des dérivations* published in 1800

polynomials. The operator $E_h - 1$ corresponds to finite differencing and we know that repeated finite differences applied to polynomials eventually give zero. So if we seek a power series in $E_h - 1$ we can guarantee convergence on polynomials. Let's define the finite difference operator $\Delta_h = e^{hD} - 1 = E_h - 1$. So we should expand $\log E_h$ around 1 and use

$$D = \log(1 + (E_h - 1))$$

Let's try a first order expansion. We get

$$\log(1 + (E_h - 1)) = E_h - 1 + \text{higher order terms}$$

In other words we can approximate the derivative of f with

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

As we might expect, this is exact if f is a linear function. What if we try second order.

$$\begin{aligned} \log(1 + (E_h - 1)) &\approx E_h - 1 - \frac{1}{2}(E_h - 1)^2 \\ &= -\frac{1}{2}E_h^2 + 2E_h - \frac{3}{2} \end{aligned}$$

So we get

$$f'(x) \approx \frac{-f(x+2h) + 4f(x+h) - 3f(x)}{2h}$$

This is known as a second-order upwind scheme. It's "upwind" because it's using values of f on one side of x to estimate the derivative. This is good in numerical methods where you only want to look at one side - for example in simulations of convection where we're trying to estimate what happens in the future and so shouldn't be using information corresponding to fluid that has already come and gone. But often we want a balanced estimate of the derivative. But a power series in e^{hD} can't ever refer to points left of x . On the other hand we could try to estimate $f'(x+h)$. In other words, we should compute $E_h D = e^{hD} D$ in terms of E_h . Clearly

$$E_h D = E_h \log E_h$$

Expanding around 1 we get

$$\begin{aligned} E_h D &= (1 + \Delta_h) \log(1 + \Delta_h) \\ &\approx (1 + \Delta_h) \left(\Delta_h - \frac{1}{2} \Delta_h^2 \right) \\ &\approx \Delta_h + \frac{1}{2} \Delta_h^2 \\ &= E_h - 1 + \frac{1}{2} (E_h - 1)^2 \\ &= -\frac{1}{2} + \frac{1}{2} E_h^2 \end{aligned}$$

So $f'(x+h) \approx \frac{1}{h}(f(x+2h) - f(x))$ or

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}.$$

which is the usual central difference formula.

More generally, we can use the power series of $\Delta_h^m \log \Delta_h$ up to terms in Δ_h^k to generate an k th order estimate for D . Even more generally, we can use the power series of $\Delta_h^m (\log \Delta_h)^n$ to estimate D^n .

Example 7.1. Derive a 3rd order upwind estimate for D using $f(x-h), \dots, f(x+3h)$. Solution:

$$\begin{aligned} (1 + \Delta_h) \log(1 + \Delta_h) &\approx (1 + \Delta_h) \left(\Delta_h - \frac{1}{2} \Delta_h^2 + \frac{1}{3} \Delta_h^3 \right) \\ &= \Delta_h - \frac{1}{2} \Delta_h^2 + \frac{1}{3} \Delta_h^3 + \Delta_h^2 - \frac{1}{2} \Delta_h^3 \\ &= \Delta_h + \frac{1}{2} \Delta_h^2 - \frac{1}{6} \Delta_h^3 \\ &= E_h - 1 + \frac{1}{2} (E_h - 1)^2 - \frac{1}{6} (E_h - 1)^3 \\ &= E_h - 1 + \frac{1}{2} (E_h^2 - 2E_h + 1) - \frac{1}{6} (E_h^3 - 3E_h^2 + 3E_h - 1) \\ &= -\frac{1}{3} - \frac{1}{2} E_h + E_h^2 - \frac{1}{6} E_h^3 \end{aligned}$$

So

$$f'(x) = \frac{-2f(x-h) - 3f(x) + 6f(x+h) - f(x+2h)}{6}$$

Example 7.2. Derive a symmetric 4th order estimate for the third derivative. Solution:

$$\begin{aligned} (1 + \Delta_h)^2 (\log(1 + \Delta_h))^3 &\approx (1 + \Delta_h)^2 \left(\Delta_h - \frac{1}{2} \Delta_h^2 + \frac{1}{3} \Delta_h^3 - \frac{1}{4} \Delta_h^4 \right)^2 \\ &\approx \Delta_h^3 + \frac{1}{2} \Delta_h^4 \\ &= \frac{1}{2} (-1 + 2\Delta_h^2 - 2\Delta_h^2 + \Delta_h^3) \end{aligned}$$

So

$$f^{(3)}(x) \approx \frac{-f(x-2h) + 2f(x-h) - 2f(x+h) + f(x+2h)}{2}$$

7.2 A generating function for all one-sided derivative schemes

Suppose we have a power series $f(x) = a_0 + a_1x + a_2x^2 + \dots$. We can form a generating function for the n th partial sum $a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ as

$$a_0 + k(a_0 + a_1x) + k^2(a_0 + a_1x + a_2x^2) + \dots$$

Assuming absolute convergence can rearrange this as

$$\begin{aligned} &(1 + k + k^2 + \dots)a_0 + k(1 + k + k^2 + \dots)a_1x + k^2(1 + k + k^2 + \dots)a_2x^2 + \dots \\ &= \frac{1}{1-k}(a_0 + a_1(kx) + a_2(kx)^2 + \dots) \\ &= f(kx)/(1-k) \end{aligned}$$

The n th order one-sided derivative is derived using n terms from $\log E_h = \log(1 + \Delta_h)$. So the coefficients from the n th order scheme are the coefficients of the polynomial in x that forms the coefficient of k^n in $\frac{\log(1+(x-1)k)}{1-k}$. We can tabulate this as

Order	Coefficients					
0	0					
1	-1	1				
2	$\frac{-3}{2}$	2	$\frac{-1}{2}$			
3	$\frac{-11}{6}$	3	$\frac{-3}{2}$	$\frac{1}{3}$		
4	$\frac{-25}{12}$	4	-3	$\frac{4}{3}$	$\frac{-1}{4}$	
5	$\frac{-137}{60}$	5	-5	$\frac{10}{3}$	$\frac{-5}{4}$	$\frac{1}{5}$

This reproduces the table in [1]. All of the results in that paper can be reproduced by forming power series from expressions of the form $(1+x)^m(\log(1+x))^n$, with m a half-integer in some cases.

Incidentally, I used the second order scheme from this table in ILM's GPU based fluid simulator basing my work on [2]. I derived it using the techniques described above but it disagreed with the paper. It turned out that the paper had an error. The moral is: it's worth knowing how to derive these schemes even if they appear to be available already in publications. (That paper is otherwise excellent and contains many powerful methods.)

7.3 Caveat

Before using any of these methods in any kind of differential equation solver please consider doing a von Neumann stability analysis. It's sometimes hard to guess which methods are and aren't stable.

7.4 Numerical Integration

We can write

$$\int_a^b f(x)dx = \left(\frac{e^{bD} - e^{aD}}{D} f \right)(0)$$

There are two ways to see this. One is to view $\frac{1}{D}f$ as the indefinite integral with e^{bD} and e^{aD} picking out the value of this indefinite integral at $x = b$ and $x = a$. Another is the following derivation:

$$\begin{aligned} \int_a^b f(x)dx &= \left(\int_a^b \exp(yD) f \right)(0) dy \\ &= \left(\int_a^b \exp(yD) dy f \right)(0) \\ &= \left(\frac{e^{bD} - e^{aD}}{D} f \right)(0) \end{aligned}$$

if you can make yourself to believe that D can behave like an ordinary number in an integration. Again we use the technique of writing $D = \frac{1}{h} \log E_h$.

7.5 Simpson's rule

Let $E = E_1$, $\Delta = \Delta_1$. We start with

$$\int_0^2 f(x)dx = \frac{E^{2D} - 1}{D} = \frac{E^2 - 1}{\log E}$$

We'll expand up to terms in Δ^2 :

$$\begin{aligned}
\frac{E^2 - 1}{\log E} &= \frac{(\Delta + 1)^2 - 1}{\Delta - \frac{1}{2}\Delta^2 + \frac{1}{3}\Delta^3} \\
&= \frac{\Delta^2 + 2\Delta}{\Delta - \frac{1}{2}\Delta^2 + \frac{1}{3}\Delta^3} \\
&= \frac{\Delta + 2}{1 - \frac{1}{2}\Delta + \frac{1}{3}\Delta^2} \\
&= (\Delta + 2)\left(1 + \frac{1}{2}\Delta - \frac{1}{3}\Delta^2 + \left(\frac{1}{2}\Delta\right)^2\right) \\
&= (\Delta + 2)\left(1 + \frac{1}{2}\Delta - \frac{1}{12}\Delta^2\right) \\
&= \Delta + \frac{1}{2}\Delta^2 + 2 + \Delta - \frac{1}{6}\Delta^2 \\
&= 2 + 2\Delta + \frac{1}{3}\Delta^2
\end{aligned}$$

Now substitute $\Delta = E - 1$ to get

$$\begin{aligned}
2 + 2(E - 1) + \frac{1}{3}(E^2 - 2E + 1) &= 2 + 2E - 2 + \frac{1}{3}E^2 - \frac{2}{3}E + \frac{1}{3} \\
&= \frac{1}{3} + \frac{4}{3}E + \frac{1}{3}E^2
\end{aligned}$$

So

$$\int_0^2 f(x)dx \approx \frac{1}{3}(f(0) + 4f(1) + f(2))$$

Rescaling the x -axis gives:

$$\int_a^b f(x)dx \approx \frac{b-a}{6}(f(x_0) + 4f(x_1) + f(x_2))$$

where $x_0 = a$, $x_1 = \frac{a+b}{2}$, $x_2 = b$. This is the usual Simpson rule. Note that we could have kept terms up to Δ^3 in the derivation above and we would have had the same result. So Simpson's rule is good for cubics as well as quadratics.

7.6 Newton-Cotes rules

The higher order integration rules known as the Newton-Cotes rules can be derived from Taylor expansions of

$$\frac{(1 + \Delta)^n - 1}{n \log(1 + \Delta)}$$

taken to terms in Δ^n and then substituting $\Delta = E - 1$. For example expanding the $n = 4$ case gives

$$\frac{1}{90}(7 + 32E + 12E^2 + 32E^3 + 7E^4)$$

These are the coefficients from Boole's integration rule

$$\int_a^b f(x)dx \approx \frac{b-a}{90}(7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4))$$

with the x_i equally spaced from a to b . This is probably the method used by Boole to derive these coefficients.

Since

$$\begin{aligned}
 1 + \Delta &= \epsilon^p, \therefore \Delta \equiv D \frac{\Delta}{D} = D \left\{ \frac{\Delta}{\log(1 + \Delta)} \right\} \\
 \Delta w_s &= \frac{d}{dx} \left\{ \frac{\Delta}{\log(1 + \Delta)} \right\} w_s \\
 &= \frac{d}{dx} \left\{ 1 + \frac{\Delta}{2} - \frac{1}{12} \Delta^2 + \frac{1}{24} \Delta^3 - \dots \right\} w_s.
 \end{aligned}$$

Integrate between limits 1 and 0, remembering that

$$[w_s]_{s=r}^{s=r+1} = \Delta w_r,$$

Figure 3: A snippet from Boole's *Calculus of Finite Differences*

7.7 Numerical integration over arbitrary regions

Suppose we wish to find a numerical method for the integral

$$I(f) = \int_D f(x, y) dx dy$$

where the integral is over some domain D in two dimensions, say. We'd like to compute this in terms of samples of f at points on some grid. How can we derive a suitable method? As we're working in two dimensions, let's define $X = \frac{\partial}{\partial x}$ and $Y = \frac{\partial}{\partial y}$. Write

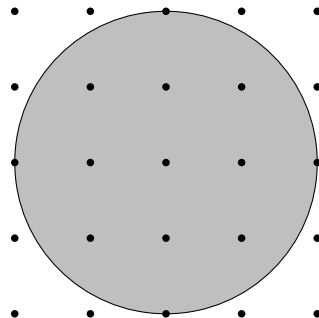
$$\begin{aligned}
 I(f) &= \int_D f(x, y) dx dy \\
 &= \int_D f(x', y') dx' dy' \\
 &= \left(\int_D e^{x'X} e^{y'Y} dx' dy' f \right) (0)
 \end{aligned}$$

So we need to expand the operator

$$\int_D e^{x'X} e^{y'Y} dx' dy'$$

in terms of $\Delta_x = E_x - 1$ and $\Delta_y = E_y - 1$ with $E_x = e^X$ and $E_y = e^Y$. (Note change of notation from Δ_h I used earlier.)

As a practical example considering integrating a function over the unit disk using a grid of 25 points. It's convenient to start with the disk of radius 2 with the 5×5 grid from $(-2, -2)$ to $(2, 2)$ illustrated below:



First we compute the integral

$$\begin{aligned} U(a, b) &= \frac{1}{\pi} \int_D \exp(ax + by) dx dy \\ &= \frac{1}{\pi} \int_0^{2\pi} \int_0^2 \exp(a \cos \theta + b \cos \theta) r dr d\theta \end{aligned}$$

This is in fact a slight variation on the classic integration that gives us the Airy disk in optics. The result is

$$U(a, b) = \frac{4I_1(2\sqrt{a^2 + b^2})}{\sqrt{a^2 + b^2}}$$

where I_1 is the first order modified Bessel function of the first kind.

Scaling by $\frac{1}{4}$ so as to get coefficients appropriate for the unit circle, we compute the Taylor series up to 4th order in Δ_x and Δ_y in

$$\frac{1}{4}(1 + \Delta_x)^2(1 + \Delta_y)^2 U(\log(1 + \Delta_x), \log(1 + \Delta_y))$$

When this expansion is written in terms of E_x and E_y we get the grid of coefficients:

$$\frac{1}{4320} \begin{array}{|c|c|c|c|c|} \hline -1 & 34 & 114 & 34 & -1 \\ \hline 34 & 464 & 444 & 464 & 34 \\ \hline 114 & 444 & -36 & 444 & 114 \\ \hline 34 & 464 & 444 & 464 & 34 \\ \hline -1 & 34 & 114 & 34 & -1 \\ \hline \end{array}$$

We can try using these coefficients to integrate an example function. First an approximation computed by another

$$\int_D \cos(2x) \log(2 + y^2) dx dy \approx 0.477976.$$

Using the coefficients above we get 0.477383.

Note that this example is sub-optimal. For example, it uses values outside of the integration region. But for smooth enough functions it gives good results and the method can be adapted to many other problems.

8 Quadratic exponentials $\exp(aD^2/2)$

Consider

$$e^{aD^2/2} f$$

To get some intuition, pick a large N and write this as

$$(e^{aD^2/2N})^N f$$

So this is multiple applications of $\exp(aD^2/2N) \approx 1 + \frac{aD^2}{2N}$. The operator D^2 measures convexity. In areas of convexity, $D^2 f$ is positive and so applying $1 + \frac{aD^2}{2N}$ will “push up” the convexity, smoothing things out. Conversely, concave areas get pushed down. So we have a sequence of N steps, each of which reduces concavity or convexity. In other words, this is a smoothing operation.

Here is an informal argument. We start by convolving f with $\exp(-\frac{y^2}{2a})$:

$$\begin{aligned} \int_{-\infty}^{\infty} \exp(-\frac{y^2}{2a}) f(x - y) dy &= \int_{-\infty}^{\infty} \exp(-\frac{y^2}{2a}) \exp(-yD) dy f(x) \\ &= \int_{-\infty}^{\infty} \exp(-\frac{(y + aD)^2}{2a}) \exp(\frac{aD^2}{2}) dy f(x) \\ &= \sqrt{2\pi} \exp(\frac{aD^2}{2}) f(x) \end{aligned}$$

We assumed that $\int_{-\infty}^{\infty} \exp(-(x-a)^2)dx = \sqrt{2\pi}$ even when a is a differential operator.

So

$$\exp\left(\frac{aD^2}{2}\right)f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2a}\right)f(x-y)dy$$

This operation is sometimes called the Weierstrass transform, but better known in the graphics and image processing world as Gaussian blur.

9 Dirac deltas of D and the Fourier transform

We now throw all caution to the wind. A standard representation of the Dirac delta function is

$$2\pi\delta(x) = \int_{-\infty}^{\infty} \exp(i\omega x)d\omega$$

So we have

$$\begin{aligned} 2\pi\delta(iD - \omega)f(x) &= \int_{-\infty}^{\infty} \exp(iy(iD - \omega))dyf(x) \\ &= \int_{-\infty}^{\infty} \exp(iy(iD - \omega))dyf(x) \\ &= \int_{-\infty}^{\infty} \exp(-iy\omega)f(x-y)dy \\ &= \int_{-\infty}^{\infty} \exp(i(x-y)\omega)f(y)dy \\ &= \exp(ix\omega) \int_{-\infty}^{\infty} \exp(-iy\omega)f(y)dy \\ &= \sqrt{2\pi} \exp(ix\omega)\tilde{F}(\omega) \end{aligned}$$

where I'm using the definition

$$\tilde{F}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-ix\omega)f(x)dx$$

Therefore

$$\delta(iD - \omega)f(x) = \frac{1}{\sqrt{2\pi}} \exp(ix\omega)\tilde{F}(\omega)$$

Notice what it's doing. It's projecting the original f to a plane wave scaled by the Fourier transform at ω . It's the projection of f onto the Fourier component corresponding to ω . We expect to be able to reassemble the original function f by summing these projections back up again. First on the right hand side:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(ix\omega)\tilde{F}(\omega)d\omega = f(x)$$

This is the usual statement of how to invert the Fourier transform. Now on the left hand side:

$$\begin{aligned}
\int_{-\infty}^{\infty} \delta(iD - \omega) d\omega f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(iD - \omega) \int_{-\infty}^{\infty} \tilde{F}(\nu) \exp(i\nu x) d\nu d\omega \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{F}(\nu) \delta(iD - \omega) \exp(i\nu x) d\omega d\nu \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{F}(\nu) \delta(i(i\nu) - \omega) \exp(i\nu x) d\omega d\nu \text{ (diagonalisation)} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{F}(\nu) \delta(-\nu - \omega) \exp(i\nu x) d\omega d\nu \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{F}(-\omega) \exp(-i\omega x) d\omega \text{ (standard property of } \delta) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{F}(\omega) \exp(i\omega x) d\omega \\
&= f(x)
\end{aligned}$$

This illustrates how we're able to work with a Dirac delta of a differential operator much like how you can work with a conventional Dirac delta.

On its own the Dirac delta of D isn't so useful. It gets more interesting when you consider $\delta(iD + ax - \omega)$ but that's for another time.

10 A short note on integration and differentiation schemes

The technique for both integration and differentiation schemes is to write your differential operator in terms of finite difference operators, expand as a power series, truncate, and only then rewrite using shift operators. It's important to first expand in terms of finite differences because high powers of finite differences eventually go to zero when applied to polynomials. So we have convergence on vector spaces polynomials. Generic linear combinations of powers of shift operators don't have this property.

11 Summary

1	identity	$f(x)$
D^n	n th derivative	$f^{(n)}(x)$
$\exp(aD)$	shift by a	$f(x + a)$
$\frac{1}{D}$	integration	$\int f(x) dx$
$\frac{1}{D+a}$	leaky integration	$\int_{-\infty}^x e^{a(y-x)} g(y) dy$
$\sinh(aD)$	finite difference	$\frac{1}{2}(f(x+a) - f(x-a))$
$\exp(aD^2/2)$	Weierstrass transform	$\int_{-\infty}^{\infty} \exp(-y^2/2a) f(x+y) dy$
$\frac{1}{\exp D - 1}$	Euler-Maclaurin sum	$\sum_x f(x)$
$\delta(iD - \omega)$	Fourier transform	$\frac{1}{\sqrt{2\pi}} \exp(ix\omega) \tilde{F}(\omega)$

12 More

So far I've looked at functions of D but it's also possible look at functions of both x and D . For example $\exp(xD)$ is an operator that appears in many places. However, I have to draw the line somewhere. So in these notes I've chosen to stick with functions of just D and leave the larger class of operators to a sequel.

I've left out any mention of fractional powers of D . Whenever I try to read about this subject I mostly find fractional powers of D being used to solve problems about fractional powers of D . But there is a large literature on the subject out there.

13 Final thoughts

There are a couple of other uniform approaches to much of what I've said above.

One is to use the shift rule to turn your problem into some kind of constraint on a polynomial. Write the polynomial in general form as $a_0 + a_1x + \dots + a_nx^n$ and write the constraint as a linear system in the vector (a_0, \dots, a_n) . This is the traditional method taught to students when solving differential equations. Guess something general enough and solve for the coefficients.

Another approach, relevant to the numerical methods, is to again assume your problem is about polynomials, and use Lagrange interpolation to fit a polynomial to your data. You now integrate or differentiate the interpolating polynomial and relate this back as a linear operation on your original data.

Both of these miss out on the hidden structure given by the rational and transcendental functions of D I've written about here.

References

- [1] Bengt Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation*, 51(184):699–699, 1988.
- [2] Jeroen Molemaker, Jonathan M. Cohen, Sanjit Patel, and Jonyong Noh. Low Viscosity Flow Simulations for Animation. In Markus Gross and Doug James, editors, *Eurographics/SIGGRAPH Symposium on Computer Animation*. The Eurographics Association, 2008.