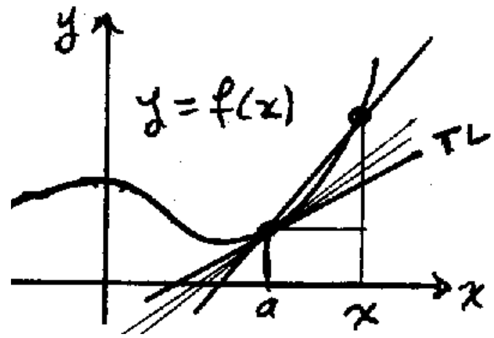
Math E-21c – Notes Differentiability, Linear Approximation, and Higher Order Approximation

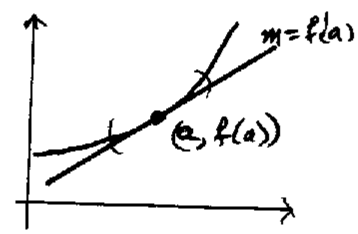
This introductory lecture will focus on basic ideas about differentiability and the derivative with applications to linear approximation, quadratic approximation, and higher order approximation.

Differentiability and the Derivative



Differential Calculus is built upon the dual idea ideas of rates of change and the derivative of a function. When we consider the graph of a function $f(x)$, the notion of rate of change is best understood in terms of slope, i.e. how fast the values of the outputs change relative to a changing input. Over an interval $[a, x]$ the values may change from $f(a)$ to $f(x)$, and the average rate of change over this interval will be correspond to the slope of the line from $(a, f(a))$ to $(x, f(x))$ given by $\dfrac{\Delta y}{\Delta x} = \dfrac{f(x)-f(a)}{x-a}$. Such a line is generally referred to as a secant line.

In order to get a more instantaneous measure of the rate of change at the single point $(a, f(a))$, we simply draw *x* closer to *a* and use the idea of a limit to determine the slope of the tangent line (TL) to the graph at this point. If this limit exists, we say that the function is **differentiable at *a***, and the value of this limit is the **derivative** $\boxed{\lim_{x \to a}\left(\dfrac{f(x)-f(a)}{x-a}\right) = f'(a)}$ at this point. Geometrically, we can think of differentiability at a point as saying that the graph at this point can be well-approximated by a tangent line at this point over some small interval around this point.



Linear Approximation

From the definition of the derivative as a limit, we can say that as long as *x* is close to *a* (how close is something that will need further clarification), then the approximation $\dfrac{f(x)-f(a)}{x-a} \cong f'(a)$ will be valid. This can also be expressed as $f(x)-f(a) \cong f'(a)(x-a)$ or as $\boxed{f(x) \cong f(a) + f'(a)(x-a)}$. Another way to say this is that in the vicinity of this point the values on the actual graph $y = f(x)$ will be approximately the same as the values on the **tangent line** $\boxed{y = f(a) + f'(a)(x-a)}$. We also refer to the function that defines this tangent line as the **linearization** of f at a, i.e. $L(x) = f(a) + f'(a)(x-a)$.

Derivatives: The calculation of derivatives from the definition is generally made simpler by using the revised definition for derivative at any point *x* given by $f'(x) = \dfrac{d}{dx}(f(x)) = \lim_{h \to 0}\left(\dfrac{f(x+h)-f(x)}{h}\right)$. In addition to finding derivatives for a variety of elementary functions, we can also prove some <u>*differentiation rules*</u> that make the calculation of derivatives much simpler and more routine. Given two differentiable functions $f(x)$ and $g(x)$ and any constant *c*, we have the following rules:

Power Rule: $\boxed{\dfrac{d}{dx}(x^p) = p\,x^{p-1}}$ We proved this for any positive integer *p*, but it's true for any (fixed) power *p*.

Sums: $\boxed{\dfrac{d}{dx}[f(x)+g(x)] = f'(x)+g'(x)}$

Differences: $\boxed{\dfrac{d}{dx}[f(x)-g(x)] = f'(x)-g'(x)}$

Constant multiples: $\boxed{\dfrac{d}{dx}[c \cdot f(x)] = c \cdot f'(x)}$

Products: $\boxed{\dfrac{d}{dx}[f(x)\cdot g(x)] = f(x)g'(x) + f'(x)g(x)}$

Quotients: $\boxed{\dfrac{d}{dx}\left(\dfrac{f(x)}{g(x)}\right)=\dfrac{g(x)f'(x)-f(x)g'(x)}{[g(x)]^2}}$

From the product rule, we can also show (for positive integer powers *p*, but it's true in general) that $\boxed{\dfrac{d}{dx}[f(x)]^p = p[f(x)]^{p-1}f'(x)}$, a fact that is a special case of the Chain Rule.

Chain Rule: If $y=f(u)$ and $u=g(x)$ are differentiable functions (where appropriate), then the composition $(f\circ g)(x)=f(g(x))$ is also differentiable and $\boxed{\frac{d}{dx}(f\circ g)(x)=f'(g(x))\cdot g'(x)}$. Using the Leibnitz notation for derivatives this can also be expressed as $\boxed{\frac{dy}{dx}=\frac{dy}{du}\frac{du}{dx}}$.

Implicit differentiation: One particularly valuable application of the Chain Rule is that it allows us to indirectly calculate derivatives from relations defined by equations without explicitly solving for one variable in terms of the other. This is often simpler even in the case where it's possible to find solve explicitly for such a function. For example, the equation $x^2+y^2=4$ implicitly defines two functions whose graphs are the upper and lower semicircles of the full circle defined by this equation. Differentiating implicitly we get:

$$\frac{d}{dx}\left(x^2+y^2\right)=\frac{d}{dx}(4) \quad\Rightarrow\quad 2x+2y\frac{dy}{dx}=0 \quad\Rightarrow\quad \frac{dy}{dx}=-\frac{x}{y}$$

In this example, since we can actually solve explicitly (for the upper semicircle) for $y=\sqrt{4-x^2}$, we could express the derivative as $\frac{dy}{dx}=-\dfrac{x}{\sqrt{4-x^2}}$. We could also have calculated this derivative directly using a combination of the Power Rule and Chain Rule.

Exponential and logarithmic functions: From the definition we can show that $\boxed{\dfrac{d}{dx}(e^x)=e^x}$, the only function whose derivative is equal to itself. A quick application of the Chain Rules gives that for any constant *k*, $\boxed{\dfrac{d}{dx}(e^{kx})=ke^{kx}}$. Using implicit differentiation we can then find the derivative of the natural logarithm function $\ln x$, the inverse of the function e^x. Specifically, if $y=\ln x$, we can exponentiate both sides to get $e^y=x$, differentiate both sides to get $e^y\frac{dy}{dx}=1$, so $\frac{dy}{dx}=\dfrac{1}{e^y}=\dfrac{1}{x}$. That is $\boxed{\dfrac{d}{dx}(\ln x)=\dfrac{1}{x}}$.

We can use these facts, plus the rules of exponents (and the fact that exponential functions and their corresponding logarithmic functions are inverses of each other) to find the derivative of <u>any</u> exponential function. Specifically, we can write $a^x=(e^{\ln a})^x=(e^{x\ln a})=(e^x)^{\ln a}$, so using previously stated facts we have $\dfrac{d}{dx}(a^x)=\dfrac{d}{dx}\left[(e^x)^{\ln a}\right]=(\ln a)(e^x)^{\ln a-1}(e^x)=(\ln a)(e^x)^{\ln a}=(\ln a)a^x$, or, more succinctly, $\boxed{\dfrac{d}{dx}(a^x)=a^x\ln a}$. We could also have done this using logarithmic differentiation, i.e. taking the logarithm of both sides and then use implicit differentiation. Specifically, $y=a^x \quad\Rightarrow\quad \ln y=x\ln a \quad\Rightarrow\quad \frac{1}{y}\frac{dy}{dx}=\ln a \quad\Rightarrow\quad \frac{dy}{dx}=a^x\ln a$, so $\boxed{\frac{d}{dx}\left(a^x\right)=a^x\ln a}$. Similarly, using the fact that $\log_a x=\frac{\ln x}{\ln a}$, we can show that $\boxed{\frac{d}{dx}\left(\log_a x\right)=\frac{1}{x\ln a}}$.

Trigonometric functions: There are three basic trigonometric functions, $\sin x$, $\cos x$, and $\tan x = \dfrac{\sin x}{\cos x}$; and their respective reciprocal functions $\csc x$, $\sec x$, and $\cot x = \dfrac{\cos x}{\sin x}$. We used the definition of the derivative and the important limits $\lim_{h\to 0}\left(\dfrac{\sin h}{h}\right) = 1$ (proved using the Squeeze Theorem) and $\lim_{h\to 0}\left(\dfrac{1-\cos h}{h}\right) = 0$ to find the derivatives of $\sin x$ and $\cos x$. The other derivatives then followed by using the Quotient Rule and the Pythagorean identities: $\boxed{\sin^2 x + \cos^2 x = 1}$, $\boxed{\tan^2 x + 1 = \sec^2 x}$, and $\boxed{1 + \cot^2 x = \csc^2 x}$. We can display all of the derivatives of the trigonometric functions in a table:

$\dfrac{d}{dx}(\sin x) = \cos x$	$\dfrac{d}{dx}(\csc x) = -\csc x \cot x$
$\dfrac{d}{dx}(\cos x) = -\sin x$	$\dfrac{d}{dx}(\sec x) = \sec x \tan x$
$\dfrac{d}{dx}(\tan x) = \sec^2 x$	$\dfrac{d}{dx}(\cot x) = -\csc^2 x$

One way to help remember these is the observation that “whatever is true for a trig function is also true for its co-function – except that you put a ‘co’ on everything and put in a minus sign.” Obviously, this needs to be correctly interpreted in order to be helpful.

We can use the basic derivatives of polynomial functions, exponential functions, and trigonometric functions in conjunction with the previous rules to calculate the derivatives of more complicated expressions involving these functions.

For example, if $g(x) = \dfrac{e^x \tan x}{x^3}$, then:

$$g'(x) = \frac{d}{dx}\left(\frac{e^x \tan x}{x^3}\right) = \frac{x^3 \frac{d}{dx}(e^x \tan x) - (e^x \tan x)\frac{d}{dx}(x^3)}{(x^3)^2} = \frac{x^3(e^x \sec^2 x + e^x \tan x) - (e^x \tan x)(3x^2)}{x^6}$$

$$= \frac{e^x(x\sec^2 x + x\tan x - 3\tan x)}{x^4}$$

Second derivatives: It should be clear by now that the derivative of a function (giving the slope of a tangent line at any point) is also a function, i.e. its values vary from point to point. Consequently, it makes sense to talk about “the derivative of the derivative”, $\dfrac{d}{dx}(f'(x)) = f''(x)$, or the rate of change of the slopes. If the slopes are increasing, then the rate of change of the slopes will be positive (so the second derivative will be positive), and this corresponds to the graph being “concave up”. If the slopes are decreasing, then the rate of change of the slopes will be negative (so the second derivative will be negative), and this corresponds to the graph being “concave down”. The alternate (Leibnitz) notation for the second derivative is $\dfrac{d}{dx}\left(\dfrac{dy}{dx}\right) = \dfrac{d^2 y}{dx^2}$.

Motion described using derivatives: If we have defined a “position function” $s(t)$ that gives the position *s* of an object at any time *t* for a “particle” moving along a line with coordinate *s*, then the interpretation of the derivative $s'(t) = \dfrac{ds}{dt}$ is as the “time rate of change of the position”, i.e. how fast the position is changing in time. This is the definition of instantaneous velocity, $v(t) = s'(t) = \dfrac{ds}{dt}$. Of course, the velocity may also be changing in time and we identify this “time rate of change of velocity” as the (instantaneous) acceleration,

$a(t) = v'(t) = \frac{d}{dt}v(t) = \frac{d}{dt}\left(\frac{ds}{dt}\right) = \frac{d^2s}{dt^2} = s''(t)$. For any given problem involving motion along a straight line, these three functions: position $s(t)$, velocity $v(t)$, and acceleration $a(t)$ can be used in tandem to answer most questions regarding the motion.

Linear Approximation calculations

The approximation $f(x) \cong f(a) + f'(a)(x-a)$ is valid for values $x \approx a$ (x near a) if the given function is differentiable at a. This can be used to give approximate values as long as we can easily calculate the values of $f(a)$ and $f'(a)$.

Example: Use the apparatus of linear approximation to provide an approximate value of $\sqrt{28}$.

Solution: Consider the differentiable function $f(x) = \sqrt{x}$. Its derivative is $f'(x) = \frac{1}{2\sqrt{x}}$ and both of these are easy to calculate at the nearby point $a = 25$. Specifically, $f(25) = 5$ and $f'(25) = \frac{1}{10}$. Linear approximation then gives that for $x \approx 25$ we have $\sqrt{x} \cong 5 + \frac{1}{10}(x-25)$. Therefore $\sqrt{28} \cong 5 + \frac{1}{10}(28-25) = 5.3$. Note that although this is a valid approximation, we don't really yet know *how good* this approximation is.

How can we improve this approximation? There are several good answers to that question. One approach might be to note that since $(5.3)^2 = 28.09$ is much closer to 28 and we can simply calculate the values of the function and its derivative at this point, we could use $a = 28.09$ as our base point, i.e. $f(x) \cong f(28.09) + f'(28.09)(x-28.09)$. We calculate that $f(28.09) = 5.3$ and $f'(28.09) = \frac{1}{10.6}$, so we now have that $\sqrt{x} \cong 5.3 + \frac{1}{10.6}(x-28.09)$ and $\sqrt{28} \cong 5.3 + \frac{1}{10.6}(28-28.09) = 5.3 - \frac{0.09}{10.6} \cong 5.291509$. At this point it is perhaps helpful to note that the calculator value for $\sqrt{28}$ is 5.29150262213, so were able to reduce the error greatly using this revised linear approximation (in this case, about 5 decimal place accuracy).

Newton's Method

The above calculation is somewhat similar to the well-established ***Newton's Method*** for finding the roots of almost any equation. The basic idea is this: If you would like to find the roots of a function $g(x)$, i.e. solutions to the equation $g(x) = 0$, start by making a good guess at an approximate root x_0, use linear approximation to find the equation of the tangent line to the graph at this point $y = g(x_0) + g'(x_0)(x - x_0)$, then see where this line crosses the x-axis to define an improved guess x_1. That is $0 = g(x_0) + g'(x_0)(x_1 - x_0)$ which we can solve to get $x_1 = x_0 - \frac{g(x_0)}{g'(x_0)}$. As long as our initial guess is not near a critical point, this should provide a greatly improved estimate of the root. This process can then be repeated to give $x_2 = x_1 - \frac{g(x_1)}{g'(x_1)}$ or more generally;

$$\boxed{x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}}.$$

In the previous example, we can seek the value of $\sqrt{28}$ by noting that if $x = \sqrt{28}$, then $x^2 = 28$ or $x^2 - 28 = 0$. So we are seeking a (positive) root of the function $g(x) = x^2 - 28$. Its derivative is $g'(x) = 2x$. If we choose $x_0 = 5$ as our first guess, then $g(5) = -3$ and $g'(5) = 10$, so $x_1 = 5 - \frac{-3}{10} = 5.3$. We then calculate that $g(5.3) = 0.09$ and $g'(5.3) = 10.6$, so $x_1 = 5.3 - \frac{0.09}{10.6} \cong 5.291509$ (which you might note is exactly the same as in our previous method). If we apply Newton's Method just one more time, we get $x_2 = 5.29150262213$ which is indistinguishable from our calculator value (11-place accuracy, in this case). It is very simple to program Newton's Method (step-by-step) into a programmable calculator, so these recursive calculations can be done very quickly.

Common Linear Approximation Expressions

Though the method for finding expressions for linear approximation can be done around any point where a given function is differentiable, it's useful to develop a library of example that are valid in the vicinity of $a=0$. Some common, simple to calculate, examples are:

$\boxed{\dfrac{1}{1-x}\cong 1+x}$ and $\boxed{(1+x)^r \cong 1+rx}$ and $\boxed{e^x \cong 1+x}$ and $\boxed{\ln(1+x)\cong x}$ and $\boxed{\sin x \cong x}$. These are valid for $x\approx 0$, but we have not yet addressed the issue of how good these approximations are or how close we must be for the approximations to remain valid.

Quadratic Approximation

If approximation by a straight (tangent) line can provide a good approximation of the graph of a function at a point $x=a$, then surely we could produce a better approximation at a point of differentiability if we instead approximate the graph by a parabola, i.e. the graph of a quadratic function. It is best to express such a quadratic function in powers of $x-a$ in order that each successive term provide small corrections to the previous terms. That is we want to approximate a function $f(x)$ with a function of the form $Q(x)=A+B(x-a)+C(x-a)^2$. If we insist that the values, slopes, and concavity (value of the second derivative) all match at $x=a$ in order to provide the best fit, and note that $Q'(x)=B+2C(x-a)$ and $Q''(x)=2C$, we get that $f(a)=Q(a)=A$, $f'(a)=Q'(a)=B$, and $f''(a)=Q''(x)=2C$, so $A=f(a)$, $B=f'(a)$, and $C=\frac{1}{2}f''(a)$. The best quadratic approximation for $x\approx a$ is therefore $\boxed{f(x)\cong f(a)+f'(a)(x-a)+\frac{1}{2}f''(a)(x-a)^2}$.

Example: Use the apparatus of linear approximation to provide an approximate value of $\sqrt{28}$.

Solution: Consider the differentiable function $f(x)=\sqrt{x}$. Its derivatives are $f'(x)=\dfrac{1}{2\sqrt{x}}$ and $f''(x)=-\dfrac{1}{4\left(\sqrt{x}\right)^3}$ and these are easy to calculate at the nearby point $a=25$. Specifically, $f(25)=5$, $f'(25)=\frac{1}{10}$, and $f''(25)=-\frac{1}{500}$. Quadratic approximation then gives that for $x\approx 25$ we have $\sqrt{x}\cong 5+\frac{1}{10}(x-25)-\frac{1}{1000}(x-25)^2$. Therefore $\sqrt{28}\cong 5+\frac{1}{10}(28-25)-\frac{1}{1000}(28-25)^2=5.3-.009=5.291$.
Though this is a valid approximation, we still can't precisely say *how good* this approximation is.

Common Quadratic Approximation Expressions

Though the method for finding expressions for quadratic approximation can be done around any point where a given function is differentiable, it's useful to develop a library of example that are valid in the vicinity of $a=0$. Some common, simple to calculate, examples are:

$\boxed{\dfrac{1}{1-x}\cong 1+x+x^2}$ and $\boxed{(1+x)^r \cong 1+rx+\frac{1}{2}r(r-1)x^2}$ and $\boxed{e^x \cong 1+x+\frac{1}{2}x^2}$ and $\boxed{\ln(1+x)\cong x-\frac{1}{2}x^2}$ and $\boxed{\sin x\cong x}$ (same as its linear approximation) and $\boxed{\cos x\cong 1-\dfrac{x^2}{2}}$.

Again, these approximations are only valid for $x\approx 0$.

Tips and Tricks

If we use the above expressions for approximation and understand that they are all valid for sufficiently small values of the given variable, then we can easily modify them via substitution and algebraic combination (taking care to neglect any terms of order higher than quadratic). Several examples follow.

Example: Given the fact that $(1+x)^r \cong 1+rx+\frac{1}{2}r(r-1)x^2$, then with $r=\frac{1}{2}$ we get that $\sqrt{1+x} \cong 1+\frac{1}{2}x-\frac{1}{8}x^2$ for small values of *x*. Replacing x by kx, we can then say that $\sqrt{1+kx} \cong 1+\frac{k}{2}x-\frac{k^2}{8}x^2$ for small values of *x*.

Example: Since $\dfrac{1}{1-x} \cong 1+x+x^2$ for small values of *x*, replacing x by $-x$ gives $\dfrac{1}{1+x} \cong 1-x+x^2$.

Example: Since $\dfrac{1}{1-x} \cong 1+x+x^2$ for small values of *x*, if we multiply both sides by x and neglect higher order terms we get that $\dfrac{x}{1-x} \cong (1+x+x^2)x = x+x^2+x^3 \cong x+x^2$ to 2nd order.

Example: Since $\dfrac{1}{1-x} \cong 1+x+x^2$ for small values of *x*, if we multiply both sides by $1+x$ and neglect higher order terms we get that $\dfrac{1+x}{1-x} \cong (1+x+x^2)(1+x) = 1+2x+2x^2+x^3 \cong 1+2x+2x^2$ to 2nd order.

Nth Order (Taylor) Approximation

If we use the same idea of matching derivatives at a given point, we can approximate a function at a point of sufficient differentiability to approximate it by an nth order polynomial. If the point about which we approximate is $x=a$, calculation shows that the coefficients must be of the form $a_n = \dfrac{f^{(n)}(a)}{n!}$. The resulting ***Taylor polynomial*** approximation is therefore

$$\boxed{f(x) \cong f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n}.$$

In the special case where $a=0$, this is known as the ***Maclaurin polynomial***:

$$\boxed{f(x) \cong f(0)+f'(0)\cdot x+\frac{f''(0)}{2!}\cdot x^2+\cdots+\frac{f^{(n)}(0)}{n!}\cdot x^n}$$

The Maclaurin polynomial is relatively simple to calculate for some simple functions. For example, if we calculate the appropriate derivatives and evaluate it's easy to derive that:

$\boxed{\dfrac{1}{1-x} \cong 1+x+x^2+\cdots+x^n}$ (though this approximation really only makes sense for small *x* (specifically $|x|<1$) – more on this later

$\boxed{e^x \cong 1+x+\frac{1}{2!}x^2+\frac{1}{3!}x^3+\cdots+\frac{1}{n!}x^n}$ (this will eventually actually work for all *x*, but it works well for $x \approx 0$)

$\boxed{\sin x \cong x-\dfrac{x^3}{3!}+\dfrac{x^5}{5!}-\cdots+(-1)^n\dfrac{x^{2n+1}}{(2n+1)!}}$ (this will eventually actually work for all *x*, but it works well for $x \approx 0$)

$\boxed{\cos x \cong 1-\dfrac{x^2}{2!}+\dfrac{x^4}{4!}-\cdots+(-1)^n\dfrac{x^{2n}}{(2n)!}}$ (this will eventually actually work for all *x*, but it works well for $x \approx 0$)

Note: Later, when we've discussed ***sequences*** and ***convergence***, we'll take up the issue of what happens if we continue ad infinitum and properly define ***infinite series*** and ***power series***. This will enable us to make sense out of formal power series expressions like

$$\boxed{f(x) \cong f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\cdots=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n} \text{ (Taylor series)}$$

and $\boxed{f(x) \cong f(0) + f'(0)\cdot x + \frac{f''(0)}{2!}\cdot x^2 + \cdots + \frac{f^{(n)}(0)}{n!}\cdot x^n}$ (Maclaurin series).

We can certainly find these formal expressions and work with them, but we really have to do some additional work before fully understanding what these formal expressions really mean. For now, just think of these expressions as a convenience with which we can work with the intention of eventually truncating at whatever power is appropriate for our needs.

In particular, we have the following short list of especially handy Maclaurin series:

$$\boxed{\frac{1}{1-x} \cong 1 + x + x^2 + \cdots + x^n + \cdots = \sum_{n=0}^{\infty} x^n}$$ (only valid for small x, specifically $|x| < 1$)

$$\boxed{e^x \cong 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}}$$

$$\boxed{\sin x \cong x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}}$$

$$\boxed{\cos x \cong 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}}$$

Tips & Tricks (again)
Once we address what it means for a series to converge, it will be possible to justify why it's valid to manipulate known series to get new series via substitution, algebraic combination, term-by-term differentiation, and term-by-term integration.

Example: Suppose we would like a Maclaurin series expressions for the function $f(x) = \tan^{-1} x$ valid for values of $x \approx 0$. We could just start calculating derivatives and evaluating at 0 to determine the necessary coefficients, but this can prove nightmarish very quickly. Instead, note that $\frac{d}{dx}\left[\tan^{-1} x\right] = \frac{1}{1+x^2}$, so if we can find a power series expression for this, we'll have a series for the derivative of what we want, and we can then just integrate term-by-term (being mindful of the arbitrary constant, of course) to produce the desired power series.

We know that $\frac{1}{1-x} \cong 1 + x + x^2 + \cdots + x^n + \cdots = \sum_{n=0}^{\infty} x^n$ for $x \approx 0$.

Replacing x by $-x$ gives $\frac{1}{1+x} \cong 1 - x + x^2 - \cdots + (-1)^n x^n + \cdots = \sum_{n=0}^{\infty} (-1)^n x^n$, also valid for $x \approx 0$.

Replacing x by x^2 gives $\frac{1}{1+x^2} \cong 1 - x^2 + x^4 - \cdots + (-1)^n x^{2n} + \cdots = \sum_{n=0}^{\infty} (-1)^n x^{2n}$, also valid for $x \approx 0$.

(Note that we could have combined these two steps by simply replacing x by $-x^2$.)

Integration term-by-term them gives that $\tan^{-1} x \cong x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots + C = C + \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$

and evaluation for $x = 0$ gives that $C = 0$. So $\boxed{\tan^{-1} x \cong x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^n \frac{x^{2n+1}}{2n+1}}$.

Notes by Robert Winters