§ 4.3 Approximation of functions: Taylor polynomials.

Newton's interpolating polynomial gives us the values of the (unknown) function in between tabulated (known) values, and tries to approximate the function over the whole interval of tabulation. This means that this type of polynomial represents a global approximation, utilizing the information contained in all tabulated values.

Brook Taylor (1715), on the other hand, was interested in constructing an interpolating polynomial capable of approximating a function locally — that is, in the neighborhood of a point. The essence of his method is as follows.

- Consider a function $f: x \mapsto f(x)$ at the points

$$\begin{aligned} x_0 &= x_0 \\ x_1 &= x_0 + \Delta x \qquad (\Delta x > 0, \text{ say}) \\ x_2 &= x_0 + 2\Delta x \\ &\ldots\ldots \end{aligned}$$

to which correspond the values

$$\begin{aligned} y_0 &= f(x_0) \\ y_1 &= f(x_1) = f(x_0 + \Delta x) \\ y_2 &= f(x_2) = f(x_0 + 2\Delta x). \\ &\ldots\ldots \end{aligned}$$

Next we apply the Newton forward difference formula (4.7′) starting with $n=1$; noting that $h = \Delta x$, we get

$$(4.9) \qquad f(x) \approx f(x_0) + (x - x_0)\frac{\Delta y_0}{\Delta x} = f(x_0) + (x - x_0)\left[\frac{f(x_0+\Delta x) - f(x_0)}{\Delta x}\right].$$

Then, for $n = 2$ we obtain

$$(4.10) \quad f(x) \approx f(x_0) + (x - x_0)\frac{\Delta y_0}{\Delta x} + \frac{1}{2}(x - x_0)(x - x_1)\frac{\Delta^2 y_0}{\Delta x^2},$$

and similarly for higher order approximations. At this point Taylor wondered what would happen if the interpolating points were chosen closer and closer to each other. This can easily be done by choosing Δx smaller and smaller. For example, for $\Delta x = 0.4$, 0.2, and 0.1 the second order approximation afforded by (4.10) might look something like this:

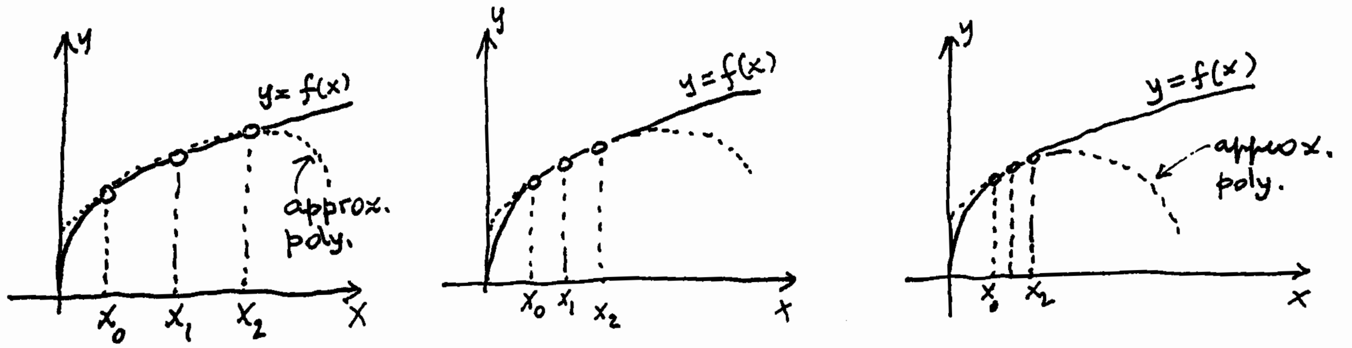

Fig. 4.4

Well, <u>what would happen if Δx became an infinitesimal dx</u>?

Remember, this is 1715; Euler is 8 yrs old, and the limit concept is way into the future!

Remembering the algebra of infinitesimals (See Divertissement #4) we would say with Taylor that

$$x_1 = x_0 + dx = x_0 \,,$$

$$x_2 = x_0 + 2dx = x_0 \,,$$

and so on. But it's easier to use the modern notation and write

$$x_1 \longrightarrow x_0 \quad \text{as } \Delta x \to 0 \,,$$

$$x_2 \longrightarrow x_0 \quad \text{as } \Delta x \to 0 \,,$$

$$\frac{\Delta y_0}{\Delta x} = \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x} \longrightarrow f'(x_0) \quad \text{as } \Delta x \to 0$$

As for $\dfrac{\Delta^2 y_0}{\Delta x^2}$, Taylor simply <u>assumed</u> that

$$\frac{\Delta^2 y_0}{\Delta x^2} \longrightarrow f''(x_0) \quad \text{as } \Delta x \to 0 \,, \tag{4.11}$$

(Taylor never proved it).

as well as

$$(4.12) \qquad \frac{\Delta^k y_0}{\Delta x^k} \longrightarrow f^{(k)}(x_0) \quad \text{as } \Delta x \to 0$$

for higher order approximations.

Summarizing, the limit process applied to the r.h.s. of (4.9) gives

$$P_{1,x_0}(x) = f(x_0) + (x-x_0)\, f'(x_0) \; ;$$

first-order Taylor polynomial

applied to the r.h.s. of (4.10), the limit process gives (using the assumption (4.11))

$$P_{2,x_0}(x) = f(x_0) + (x-x_0)\, f'(x_0) + (x-x_0)^2 \, \frac{f''(x_0)}{2!}$$

Second-order Taylor polynomial

and so on. The general n^{th} order Taylor polynomial is therefore given by (using assumption (4.12))

$$(4.13) \qquad \boxed{P_{n,x_0}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k .}$$

■ Remark 4.3. Note the big difference between (4.7) and (4.13): The former utilizes information from all data points in order to produce values in between the data points (interpolation); in contrast, the Taylor polynomial utilizes the information from only one point, x_0, in order to predict the values at other points (extrapolation). ■

■ Remark 4.4 - Maclaurin's Approach. About 27 years after Taylor published his result, Colin Maclaurin (1698–1746) rederived it by a different method. Here is how he did it.

4.10

- Given f and a point x_0 in its domain, we seek a polynomial

(4.14) $$p(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n$$

such that the coefficients are determined by the condition that all the derivatives of p at x_0 coincide with the derivatives of f at the same point; i.e.,

(4.14′) $$p^{(k)}(x_0) = f^{(k)}(x_0) \quad ; \quad k = 0, 1, 2, \ldots, n.$$

So, putting $x = x_0$ in (4.14) gives

$$p(x_0) = a_0 = f(x_0). \qquad \left(\text{By } (4.14') \text{ with } k=0\right).$$

Next we take the derivative of (4.14),

$$p'(x) = a_1 + 2a_2(x-x_0) + \cdots + n a_n (x-x_0)^{n-1},$$

and then set $x = x_0$, with the result

$$p'(x_0) = a_1 = f'(x_0). \quad \left(\text{By assumption } 4.14', \text{ with } k=1\right).$$

The process can clearly be continued for as long as necessary, and we get

$$a_2 = \frac{1}{2!} f''(x_0),$$
$$a_3 = \frac{1}{3!} f'''(x_0),$$
$$\vdots$$
$$a_n = \frac{1}{n!} f^{(n)}(x_0).$$

Putting everything together, we see that the assumed polynomial (4.14) becomes

(4.14′) $$p(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2!} f''(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n,$$

which is just $P_{n,x_0}(x)$ as given by Taylor. ■

■ Remark 4.5. Consider the first-order Taylor polynomial, which is simply found by setting $n=1$ in (4.13),

$$P_{1,x_0}(x) = f(x_0) + f'(x_0)(x-x_0).$$

When we want to use this as an approximation to f in the vicinity of x_0, we shall write

$$(4.15)\qquad \boxed{f(x) \approx f(x_0) + f'(x_0)(x-x_0)}$$

and refer to this as the linear approximation to f. The reason for the name is that — as you can easily show for yourselves — Eq.(4.15) is the equation of the tangent line to the graph of f at the point x_0.

Note the relation to the linear interpolation formula (see Fig. 4.3): They are different, but become the same in the limit as $x_1 \to x_0$ (geometrically, when $x_1 \to x_0$ the secant tends to the tangent).

Of course, the linear approximation to f will be less and less accurate as we move further away from x_0. Higher accuracy will be produced by $P_{2,x_0}(x)$, $P_{3,x_0}(x)$, and so on, as we are going to see soon. ■

• A historical remark.

In most calculus books the "Maclaurin" polynomials are those given by (4.14') for the special case $x_0 = 0$. In other words, they are $P_{n,0}(x)$. This is not true, for Maclaurin, in his book (1742), acknowledges the work of Taylor but never claims to have invented "the Maclaurin polynomials".

Misnomers of this type are very frequent. As another example concerning Maclaurin — he invented, among other things, the rule for solving systems of linear algebraic equations that we now call Cramer's rule!

§ 4.4 – Application of Taylor polynomials: Linearization.

The study of physical phenomena, or the behavior of complex systems, usually lead to nonlinear mathematical models; that is, the equations of the models (algebraic, differential, integral, and so on) contain nonlinear terms, and this fact makes it next to impossible to find exact solutions. Hence, most of the time we must study our mathematical models in an approximate manner, be that analytical or numerical. The technique of linearization is one of the most frequently used; and the following examples give the gist of the method.

▲ Example 1. Power dissipation in a resistor. (This is a trivial case, for the nonlinear term does not offer any trouble. But it serves to illustrate that linearization can do a very good job in the neighborhood of the region of interest.)

Problem: Find a linear model for the power variation in a resistor of 10 (Ω), valid for an operating current $I \approx 0.5 (A)$.

Solution: We know from physics that the power P dissipated in a resistor is given by $P(I) = RI^2$. In order to linearize this, we need to derive the first-order Taylor polynomial approximation of P(I) for $I_0 = 0.5$.
The calculation is straightforward:

$$P(I_0) = RI_0^2 = 10\,(0.5)^2 = 2.5 \ (W)$$

$$P'(I_0) = 2RI_0 = \cdots = 10$$

Recall:
Ω stands for ohm
A " " ampere
W " " watts

and so

$$P_{1,0.5}(I) \approx 2.5 + 10\,(I - 0.5) = \underbrace{10\,I - 2.5}_{\text{linear approximation}}\ , \ (W).$$

The table below shows how good the linearization is:

Current I (A)	0.5	0.499	0.501	0.49	0.51	04	0.6
True Power (W)	2.5	2.49001	2.51001	2.401	2.601	1.6	3.6
$P_{1,0.5}(I)$ (W)	2.5	2.49	2.51	2.4	2.6	1.5	3.5

• It is clear that the linear approximation is quite good when I is in the neighborhood of the operating point $I_0 = 0.5$, but bad otherwise. ▲

▲ Example 2. "The simple pendulum" problem.

A pendulum consists of a mass m suspended by a rod of length ℓ from a fixed point P, as shown in Fig. 4.5.

It is called "simple" because it's very idealized: The rod ℓ is supposed to be absolutely rigid and massless; the pivot P is assumed frictionless; the mass m is taken to be concentrated at a single point; and no air drag is allowed.

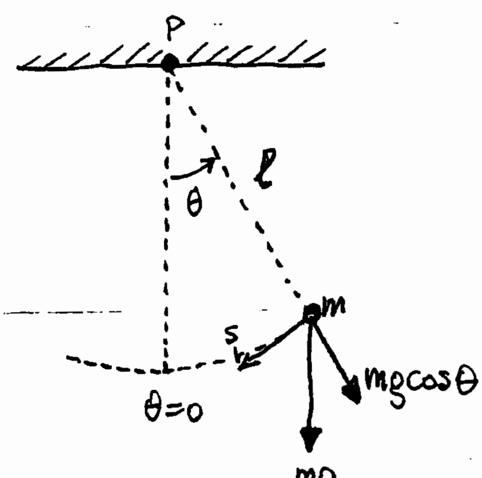

Fig. 4.5

This is typical of math. modeling — whenever we study a phenomenon for the first time, we always make the model as simple as it can be (provided the main characteristics are retained).

Under these conditions, gravity is the only force acting on the bob m, and Newton's second law gives us the differential equation (D.E.)

$$m\frac{d^2s}{dt^2} = -mg\sin\theta .$$

Here the minus sign shows that

the tangential component of the force always acts in the direction of the equilibrium position ($\theta = 0$), while the l.h.s. is just the expression of "mass x acceleration". The latter is given in terms of the arc $s(t)$, but can be easily related to $\theta(t)$ by the well known formula $s = \ell\theta$ (see MEM, p. 114). Hence, the equation of motion of the simple pendulum is given by

$$(4.16) \qquad \frac{d^2\theta}{dt^2} + \frac{g}{\ell}\sin\theta = 0,$$

Isn't it intriguing that it doesn't depend on m?

which is a nonlinear D.E. (because of $\sin\theta$).

Problem: Linearize (4.16).

Sol'n. We simply approximate the sine fct., $f(\theta) = \sin\theta$, by the first order Taylor polynomial around the equilibrium position $\theta = 0$, $P_{1,0}(\theta)$. We have,

$$f(0) = \sin(0) = 0 \quad ; \quad f'(0) = \frac{d}{d\theta}\sin\theta\bigg|_{\theta=0} = \cos(0) = 1 ;$$

and so

$$P_{1,0}(\theta) = 0 + 1\cdot(\theta - 0) = \theta,$$

which gives us

$$\sin\theta \approx \theta,$$

and the D.E. (4.16) becomes (using the overdot notation)

$$(4.16') \qquad \ddot{\theta} + \frac{g}{\ell}\theta = 0. \quad \blacktriangle$$

NOTE. You will find more examples in the Assignment. Generally speaking, it is difficult to underestimate the importance of the linearization process, for most of physics and engineering is based on linear math. models.

§ 4.5 - Application of Taylor polynomials : Root finding.

One of the most frequent problems of applied mathematics is that of computing the roots of the equation

$$f(x) = 0 .$$

This is also called the problem of finding the zeros of f.

The simplest cases occur when f is a first or second degree polynomial, for then it's easy to find the exact solution. Things get complicated even for polynomial equations of degree 3 or 4, and no exact solutions exist for polynomials of degree ≥ 5. Thus, approximation methods are a must, and the following example gives you an idea of how they work.

▲ Example (Newton, 1736). To find a real root of the equation

$$x^3 - 2x - 5 = 0 .$$

Sol'n. We have to find a real zero of the function $f(x) = x^3 - 2x - 5$. The idea is this: Approximate f by $P_{1,x_0}(x)$ and solve the equation $P_{1,x_0}(x) = 0$. In other words, we take the zeros of the first order Taylor polynomial as an approximate value of the zeros of f. Therefore, we need to find a suitable point x_0 close enough to the real zero so that the linear approximation will be good. And we are going to do this by educated guessing.

The domain of f is the entire real line; however, because the highest power is x^3, we know that the values of f will be negative for large negative x values. So we check a few values of x starting at $x=0$, and we get

$$f(0) = -5 \quad ; \quad f(1) = -6 \quad ; \quad f(2) = -1 \quad ; \quad f(3) = 16 .$$

f changes sign

Now, because of the change in sign, we know from the Intermediate Value Theorem (IVT) that f must vanish for an x between 2 and 3; and since $|f(2)| \ll |f(3)|$, the zero must be closer to 2 than to 3. So we guess

$$x_0 = 2$$

and compute $P_{1,2}(x)$:

$$f(2) = -1 \quad ; \quad f'(2) = (3x^2-2)\Big|_{x=2} = 10 \quad ;$$

$$P_{1,2}(x) = -1 + 10(x-2) = 10x - 21 .$$

We now replace the original equation by $10x - 21 = 0$ and solve the latter for x (call x_1 the sol'n)

$$x_1 = 2.1$$

which we expect to be closer to the true root than $x_0 = 2$. This algorithm can be repeated, using x_1 instead of x_0; we get

$$P_{1,2.1}(x) = f(2.1) + f'(2.1)(x - 2.1) = 11.23x - 23.52$$

and setting this equal to zero and solving for x (call x_2 the solution) gives

$$11.23\,x_2 - 23.52 = 0 \quad \Rightarrow \quad x_2 \approx 2.09,$$

and so on (for as many times as desired).

Geometrical Meaning (Fig. 4.6)

First application of the algorithm: Draw the tangent line to the graph of f at x_0; its intercept x_1 is closer to the true root.

Second application: Draw the tangent line at $f(x_1)$; its intercept x_2 will be closer to the true root.

And so on. ▲

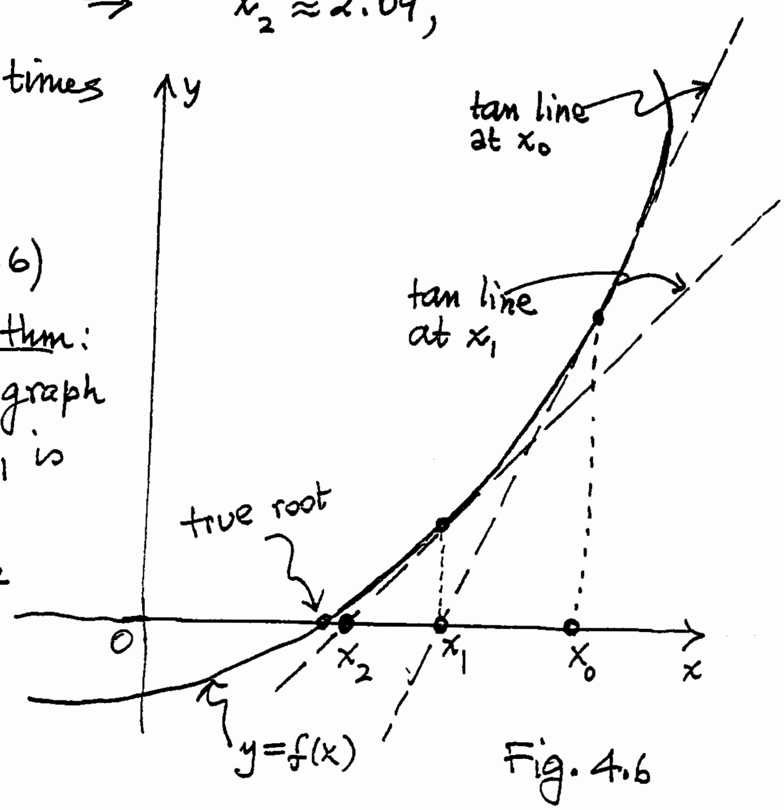

Fig. 4.6

4.17

Remark 4.6. The example above shows that the result of the algorithm is a sequence of approximate values $x_0, x_1, x_2, \dots, x_n$, which presumably converges to the value of the true root. However, at the moment we don't know that the sequence converges — clearly, there must be certain conditions to be satisfied in order for this to happen. In fact, it can be shown that everything depends on how close to the real root the starting guess is. We shall come back to this point a bit later. For now, work on the examples given in the Assignment. ■

Remark 4.7. In modern terms, Newton's root-finding algorithm is best presented in a compact form which is easily derived as follows.

The linear approximation to f at the starting guess x_0 is given by

$$f(x) \approx \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{put}=0 \text{ and solve for } x\ (x_1)}$$

from which

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Next we take the linear approximation at x_1 and repeat:

$$f(x) \approx \underbrace{f(x_1) + f'(x_1)(x - x_1)}_{=0 \text{ and solve for } x\ (x_2)}$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)},$$

and so on. Thus, the general form of the algorithm is given by

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}} \quad ; \quad n = 0, 1, 2, \dots \qquad (4.17)$$ ■