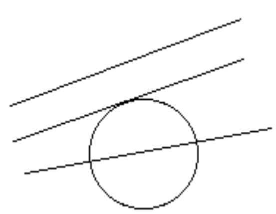
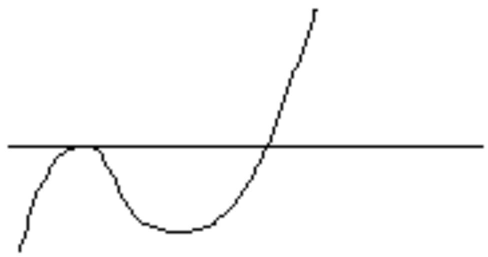
MATH 253 LECTURE NOTES for FRIDAY SEPT. 23,1988: TANGENTS
edited March 26, 2013.

Suppose that Apple Computers notices that every time they raise (or lower) the price of a $5,000 Mac II by $100, the number of Mac II sales for that month drops (or rises) by 1%. How much should they raise or lower the price to maximize their monthly gross sales? This translates mathematically, as we will see later, into a problem requiring us to find a tangent line to the graph of income versus price.
The problem of finding the tangent line to various curves interested and challenged the best mathematicians for centuries. It was solved by the ancients for circles and parabolas and some other special curves. Newton gave the "last word" on the problem with his definition of a tangent as a "limit" of secant lines (generalizing Euclid's characterization of tangents to circles as limits of secant lines in his Prop. III.16), but earlier Descartes and Fermat, among others had completely understood the problem, at least for graphs of polynomials, simply using algebra and no infinite processes such as limits. In today's notes we will see how Descartes' method works on polynomials. Later we will observe how it fails for trigonometric functions like sine and cosine, at which point we will introduce the full power of Newton's method of limits. Newton's approach will then clarify not only the question of how to find a tangent line when it exists, but also how to identify those situations in which a tangent line does not in fact exist. I hope that this presentation, which separates the two essentially different ideas of tangent and limit, will help the student to understand both ideas better.

The one curve whose tangents almost everyone is familiar with is the circle:

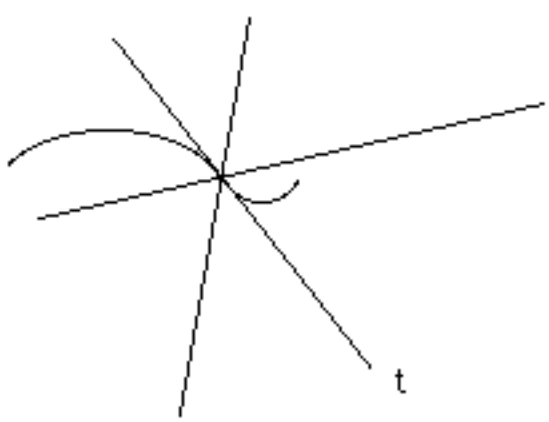


The middle line in this picture, the one meeting the circle only once, is the tangent line to the circle. The top line misses the circle entirely and the lower line cuts across the circle, meeting it again. But look at the following picture:



Here the horizontal line meets the curve twice, but if you look closely you see that it meets it differently at the two points of intersection.  At the left hand intersection point, the line just touches the curve without crossing it, the way the tangent line to a circle meets the circle.  At the second intersection point, the one on the right, the line cuts right across the curve.  This line is a tangent line to this curve, but it is only tangent to the curve at one of the two points where it meets the curve, namely the left one.  (Compare the definition of tangent line to a circle in Euclid, as a line that meets the circle "but does not cut (across) it".)  So if a line meets a curve at several points, it may be tangent to the curve at some of the points and not at others depending on the way in which it meets the curve at each point.  In particular a line is tangent to a curve at a point if it meets the curve at that point in a special way.  There will be many lines that meet a curve at a particular point, but in general only one of them will be tangent to the curve at that point.

Here is still another example of a different kind, the midpoint of an "S" curve:




This time all three lines meet the curve at the same point, and they all cross the curve in the sense that they go from one side to the other, but the line labeled "t" is different from the other lines passing through this curve at this point since it meets the curve at a smaller angle than the other lines do. (Euclid also characterizes a tangent to a circle in terms of the angle between the tangent and circle, in Prop. III.16.) It "touches" the curve similar to the way the tangent to a circle touches a circle, and it only crosses the curve because of the way the curve happens to change direction at the point. To see this it may help to look at only half the "S" curve:

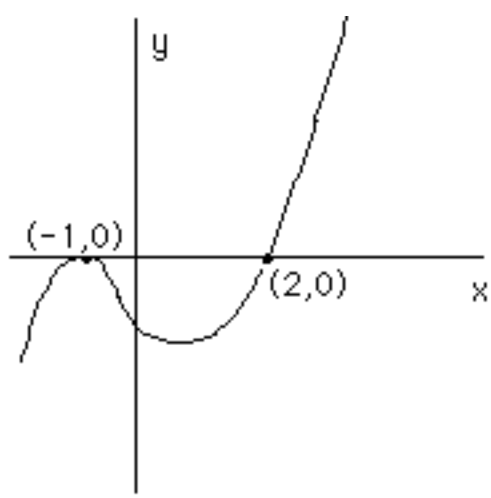

<u>DESCARTES' METHOD FOR FINDING TANGENTS</u>

Now let us recreate Descartes' approach to understanding just what it is about the way a tangent line meets a curve which differs from the way all other lines do so. **The idea is to use algebra to solve for the points where a line meets a curve, and then notice that at a point of tangency a line actually meets a curve "doubly" in the sense that the equation we solve to get the intersection points has a "repeated root" at the point of tangency.** Then all we have to do is look, among all lines that meet a curve at a given point, for the one that meets it doubly and we will have the tangent line at that point.

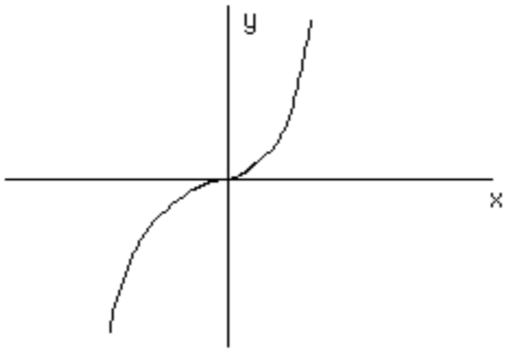
First let's examine again a tangent to a circle, say the circle with equation $x^2+y^2=4$. This is the circle of radius 2 centered at the origin and we know that at the north pole, i.e. the point (0,2), that the tangent line has equation $y = 2$. Stop and draw the graphs now on scratch paper to convince yourself of this. Now to find the points where the line meets the circle we just substitute the equation for the line into the equation for the circle and solve. So we substitute $y = 2$ into the circle's equation, getting the equation $x^2 + 4 = 4$, or $x^2 = 0$, which indeed has $x = 0$ as a double root! [Recall that we find the roots of an equation by factoring it, and repeated roots correspond to repeated factors. Here we have $(x)(x) = 0$, so x is a double factor and so 0 is a double root. In the example $(x-1)(x-2)(x-1) = 0$, we would have $(x-1)$ as a double factor and so

1 is a double root; and so on. Review your algebra text, if necessary, on the multiplicity of roots.]

Next consider the curve with equation $x^3 - 3x - 2 = y$. If you plot some points you find that this curve looks like the picture at the top of page 2, with the horizontal line as the x-axis. Thus we guess that the x-axis meets this curve at two points, and is tangent to it at exactly one of those points. To verify this we must solve algebraically for the points where the x-axis meets this curve, so again we substitute the equation of the x-axis, which is $y = 0$, into the equation of the curve, which gives us $x^3 - 3x - 2 = 0$. To factor this recall that the first guess at the roots of this polynomial should be the factors of the constant term, which are $\{+2,-2,+1,-1\}$. [For this, review an algebra lesson on rational roots.] Trying these we find that +2 and -1 are indeed roots of this polynomial and therefore that $x-2$ and $x+1$ are factors. Dividing them out leaves us with $x+1$, so we get the full factorization as $x^3 - 3x - 2 = (x-2)(x+1)(x+1)$. Hence the x-axis meets the curve at the points $x=2$ and $x=-1$, and it meets the curve doubly at $x= -1$ but only simply at $x=2$. If we look again at the graph below, we see that this reflects the fact that the line $y=0$ meets the curve $y= x^3-3x-2$ tangentially at the point (-1,0) corresponding to the double root $x= -1$, and meets it again <u>non</u>-tangentially at the point $x=2$, corresponding to the simple root $x=2$.




Finally here is another example, similar to the case of the midpoint of the "S-curve" above: the equation of the curve is $y=x^3$, and if we graph it, it looks like the following picture in which we see that the x-axis is apparently the tangent line to the curve at the point (0,0).




Now let's do the algebra to compute the multiplicity of the point (0,0) as an intersection point of this curve with the x-axis. This time we want to solve $y=0$ and $y=x^3$, so we substitute and get $0=x^3=(x)(x)(x)$, so that this time $x=0$ is actually a triple root. What this means is just that the x-axis is <u>very</u> tangent to the curve $y=x^3$ at the point (0,0) corresponding to the triple root $x=0$. So we have the following principle:

**A line is tangent to a curve at a given point if, when we substitute the equation of the line into that of the curve and solve, the point in question corresponds to a root which is at least a double root, i.e. either a double or triple or higher multiplicity root.**

[Note also that the tangent apparently cuts across the curve only in case the multiplicity is odd. In the case of a circle which has a quadratic equation, the only possible multiplicities were one and two, hence the tangent line could be characterized by Euclid as the one with even multiplicity.]

Now let's use this principle to find some tangent lines to some curves. Let the curve be the parabola $y=x^2$, and we ask for the tangent line at the point (2,4). We know from pre-calculus that any line with equation $y=m(x-2)+4$, where m is any number, will pass through the point (2,4), and hence will meet the parabola there. (Just set $x=2$ in this linear equation and see that you get $y=4$.) We want to see what the slope m should be so that the line will meet the parabola <u>doubly</u> at the point (2,4). So let's substitute the equation of the line into that of the curve and solve for x. (Of course m is also an unknown quantity here but think of it as a fixed number for now.) Substituting gives $m(x-2)+4=x^2$, and if we bring all the unknowns to one side we have $x^2-m(x-2)-4=0$. This is a quadratic equation whose two roots should give us the two points where the line $y=m(x-2)+4$ meets the curve $y=x^2$. Now we know one of those points is at (2,4) so

$(x-2)$ is guaranteed to be a factor of this quadratic. If we divide $x^2-m(x-2)-4$ by $(x-2)$, using long division (or synthetic division) of polynomials as in pre-calculus, we get $(x+2-m)$, so that $x^2-m(x-2)-4= (x-2)(x+2-m)$, and so the points where the line meets the curve correspond to the roots $x=2$ and $x=m-2$. The only way we can have a double root at $x=2$ then would be if the root $m-2$ were also equal to 2; but $m-2 = 2$ means $m=4$, so the tangent line to $y=x^2$ at the point $(2,4)$ has slope $m=4$, and thus has equation $y=4(x-2)+4$, or in point -slope form, $y-4=4(x-2)$.

Exercise: Use this same method to check that at $(3,9)$ the tangent line to $y=x^2$ has equation $y-9=6(x-3)$. Graph this parabola and these lines and see that the lines really look tangent to the curve at the given points.

Next we will derive some general formulas for the slopes of tangent lines. First consider again the parabola $y=x^2$, and the point $(a,a^2)$ on it. A line with equation $y=m(x-a)+b$ passes through this point if when we set $x=a$ we get $y=a^2$; which means b must equal $a^2$. So the equation of any line which meets the parabola at $(a,a^2)$ must have form $y=m(x-a)+a^2$, and we want to solve for the value of the slope m which will make the line tangent to the parabola at that point. As before we substitute the equation of the line into that of the parabola to get $m(x-a)+a^2=x^2$, and rearrange to $x^2-m(x-a)-a^2=0$. Since putting $x=a$ solves this equation, we know that $(x-a)$ is a factor, [if necessary, review the "factor theorem" of precalculus]. Dividing by $(x-a)$ gives $(x+a-m)$ so that $x^2-m(x-a)-a^2=(x-a)(x+a-m)$. Thus the two points where the line meets the parabola correspond to the roots $x=a$ and $x=m-a$, so there is a double root at $x=a$ only if $m-a$ is also equal to a, or only if m equals 2a. **Thus the tangent line to the parabola $y=x^2$ at the point $(a,a^2)$ must have slope $m=2a$, and equation $y=2a(x-a)+a^2$ or equivalently $y-a^2=2a(x-a)$.**

Exercise: Use this method to see that the tangent line to the curve $y=x^3$ at the point $(a,a^3)$ has slope $3a^2$, and equation $y-a^3=3a^2(x-a)$.

We do one more example:

**Theorem: The tangent line to the curve $y = x^n$ at the point $(a,a^n)$ has slope $m = na^{n-1}$, and thus has equation $y-a^n = na^{n-1}(x-a)$.**

Proof: A line through that point has equation $y=m(x-a)+a^n$, so substituting gives $m(x-a)+a^n=x^n$, or $x^n-m(x-a)-a^n=0$. We know $x=a$ is a root of this, so $(x-a)$ is a factor, and dividing by it gives this factorization: $x^n-m(x-a)-a^n =(x-a)(x^{n-1}+ax^{n-2}+a^2x^{n-3}+.....+a^{n-2}x+a^{n-1}-m)$. (You need of course to get out your pencil and paper and do this division problem yourself; again you can review these division skills in a pre-calculus book.) Thus there is at least one root given by $x=a$, and $x=a$ is a <u>double</u> root only if $x=a$ is also a root of the long factor on the right. That is $x=a$ is a double root if and only If, when we substitute $x=a$ into the long factor, we get zero. If we actually do substitute it in, we see we get n terms all equal to $a^{n-1}$, and -m at the end, which equals $na^{n-1}-m$. The only way this can be zero is when $m = na^{n-1}$, as we said. Q.E.D.

Now we want to carry this method as far as we can, and derive by the same technique, a method for computing the tangent line to the graph of any polynomial.

So let f(x) be shorthand for any polynomial at all, (such as $x^6+59x^3-x^2+4$, perhaps.) We know the point (a,b) lies on the graph of this polynomial only if b=f(a), where f(a) is shorthand for whatever you get when you set x=a in the polynomial f(x). Thus our problem is to find a way to compute the slope of the tangent line to the graph of y=f(x) at the point (a,f(a)). A line through this point has equation y=m(x-a)+f(a), so after substituting we have to solve m(x-a)+f(a)=f(x), or f(x)-f(a) -m(x-a) =0. Now you can see that if you set x=a in this you get zero, so x=a is one root and thus (x-a) is one factor. Then we would like to divide f(x)-f(a) -m(x-a) by (x-a) to get a second factor as we have done before, and then as before x=a will be a <u>double</u> root of f(x)-f(a) -m(x-a) precisely when x-a is also a root of the second factor. Since we don't know what f(x) is, we do not know exactly what we get as the second factor. We do know however that we do get something, so we just write {f(x)-f(a) -m(x-a)}/(x-a) to represent the other factor. Now as we said, x=a is one root of f(x)-f(a) -m(x-a) =0, and it is a double root only if x=a is also a root of the "other factor" {f(x)-f(a) -m(x-a)}/(x-a), i.e. only if after simplifying the quotient {f(x)-f(a) -m(x-a)}/(x-a), and then setting x=a, we again get zero. Now if we just rearrange this a bit, this quotient becomes [{f(x)-f(a)}/(x-a)] - m(x-a)/(x-a), and after cancelling, this is equal to [{f(x)-f(a)}/(x-a)] - m. Now this will be zero after simplifying and setting x=a, only if the quotient [{f(x)-f(a)}/(x-a)] becomes equal to m after simplifying and setting x=a. Thus we have found a procedure for computing m, the slope of the tangent line. We rephrase it as follows:

**To find the slope m of the tangent line to the graph of $y=f(x)$ at the point $(a,f(a))$, where f is any polynomial, just simplify the expression $[\{f(x)-f(a)\}/(x-a)]$, and then set $x=a$.**

Example: Let $f(x)=x^2+2x$, and $a=1$. Then $f(a)=1^2+2(1)=3$, so we have to simplify $[x^2+2x-3]/(x-1)$. Dividing gives us $(x+3)$, and then setting $x=1$, gives $1+3=4$. So the slope of the tangent line is 4, and the equation is $y-3=4(x-1)$.

Exercise: Use this method to find the equation of the tangent line to the graph of $y=x^3-4x^2+1$ at the point $(-1,-4)$.

What happens when you use this procedure to try to find the tangent line to the graph of $y=1/x$ at the point $(2,1/2)$? [Note: $1/x$ is <u>not</u> a polynomial.] What about $y=x^{1/2}$ ? What about $y = \sin(x)$?
In all cases except the last, the method will work if you persist. The only differences are in how to do the simplification of $\{f(x)-f(a)\}/(x-a)$. When $y=\sin(x)$ however, no simplification is possible and a genuinely new idea is needed, that of "limits". This brilliant idea of Newton's (generalizing and perhaps inspired by Euclid's Prop. III.16) confirms why Descartes' method above works, and also improves on it by showing how to evaluate $\{f(x)-f(a)\}/(x-a)$ at a, at least in some cases, even when you cannot simplify.