Curves and Tangents

E. Krishnan
ekrshnan@gmail.com

1 Curves and Tangents

One of the motivations for the invention of differentiation was a problem in geometry, the computation of the slope of a tangent to a curve. But before discusssing the details, we must have a clear notion of what the two terms *curves* and *tangents* mean.

2 Curves

The intuitive visualization of a curve is easy, but it is hard to describe it precisely in words (as is the case with many geometrical concepts). Euclid (third century BCE) defined a curve as a "breadthless length" and Proclus (fourth century CE) defined it as the "flux of a point", that is the path traced by a moving point.

The invention of analytic geometry by Descartes and Fermat in the seventeenth century lead to the description of curves as algebraic equations. Under this scheme, a plane is identified with the set $\mathbb{R}^2$ of all ordered pairs of real numbers and any geometric figure in the plane with a subset of $\mathbb{R}^2$. In particular, a curve is a subset of $\mathbb{R}^2$ consisting of those ordered pairs (x, y) where x and y are related in any of the following four forms:

1. $y = f(x)$, where f is a function on a subset of $\mathbb{R}$
2. $x = f(y)$, where f is a function on a subset of $\mathbb{R}$
3. $x = f(t), y = g(t)$, where f and g are functions on a subset of $\mathbb{R}$
4. $f(x, y) = 0$, where f is a function on a subset of $\mathbb{R}^2$

On the other hand, any of these subsets of $\mathbb{R}^2$ can be geometrically represented as the set of points in the plane with coordinates (x, y) satisfying the correspoding condition.

2.1 Examples:

1. $y = f(x)$
2. $x = f(y)$
3. $x = f(t), y = g(t)$
4. $f(x, y) = 0$

3 Tangents

Intuitively we think of a tangent to a curve as a straight line which "just touches" the curve at a point, as in Fig. 1(a) below. The emphasis is on *touching* in contrast to *intersecting*. Thus the axis of a parabola it at only one point, but it doesn't conform to our intuitive idea of a tangent (see Fig. 1(b)).

Ancient mathematicians tried to distinguish between a line *touching* a curve and *intersecting* it various ways. For example, Euclid says,

Fig. 1(a): Tangent

Fig. 1(b): Not Tangent

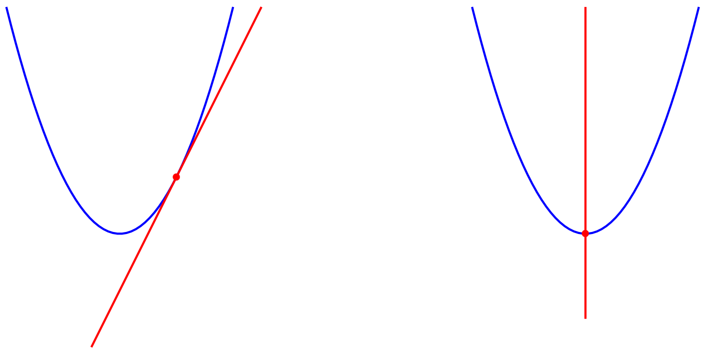

Figure 1:

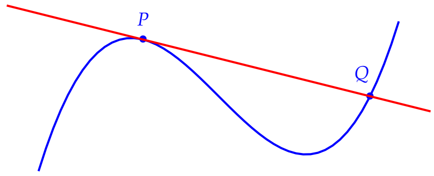

Figure 2: Tangent at P, not at Q

> *A straight line is said to touch a circle which, meeting the circle and being produced, does not cut the circle*

which defines tangents for circles pretty well, since for a circle and a line there are only there are only three alternatives: they meet in two points, one point or no points; but it does not define tangents for other curves such as the one in Figure 2.

A century late, Apollonius in his work on conic sections tires to define tangents as

> *a line such that no other straight line could fall between it and the curve*

The modern definition of a tangent to a curve is that it is the *limiting position* of lines joining two points on the curve, as one point moves closer and closer to the other. This idea emerged slowly in the seventeenth century. (See this article for the details.)

To translate this intuitive, geometric idea to precise, algebraic terms, suppose that the curve is given by an algebraic equation connecting the x and y coordinates of an arbitrary point on the curve (functional, implicit or parametric). We want to find the the equation of the tangent to the curve at a point with coordinates (a, b). For this, we must find its slope. If (x, y) is any other point on the curve, the slope the line joining (x, y) and (a, b) is $\dfrac{y-b}{x-a}$. Since the tangent line is the limiting position of such lines, its slope is the limit of these quantities, as x approaches a. This is where differentiation enters the picture. We next see how it is done for curves given by different kinds of equations.

3.1 Functional representation

Suppose the curve is given by the equation $y = f(x)$, where f is a real valued function on a subset D of real numbers. Any point on the curve is of the form $(t, f(t))$ for some t in D. Let a be an arbitrarily fixed real number in D. Then for any t in D with $t \neq a$, the slope of the line joining $(t, f(t))$ and $(a, f(a))$ is

$\dfrac{f(t) - f(a)}{t - a}$, so that the slope of the tangent at $(a, f(a))$ is

$$\lim_{t \to a} \frac{f(t) - f(a)}{t - a}$$

provided this limit exists. But this is precisely the condition that f is differentiable at a and in this case, the slope of the tangent is $f'(a)$. Thus we have the following:

3.1 Proposition. *Let $y = f(x)$ be the equation of a plane curve, where f is a real valued function on a subset D of real numbers, and let a be a number in D. If f is differentiable in an open interval containing a and contained in D, then the slope of the tangent to the curve at the point $(a, f(a))$ is the derivative $f'(a)$ of f at a and consequently, the equation of the tangent is $y = f(a) + f'(a)(x - a)$.* $\square$

But there may be points on (x, y) on the curve $y = f(x)$ where there is a tangent parallel to the y-axis and at these points f is not differentiable. For example, consider the curve $y = \sqrt[3]{x - 1}$ and the point $(0, 0)$ on it. The geometrical definition of the tangent as the limiting position of lines joining $(1, 0)$ to nearby points on the curve shows that the tangent at $(1, 0)$ is the line $x = 1$.

But f defined on $\mathbb{R}$ by $f(t) = \sqrt[3]{t - 1}$ is not differentiable at 1, since

$$\frac{f(t) - f(1)}{t - 1} = \frac{\sqrt[3]{t - 1}}{t - 1} = \frac{1}{\sqrt[3]{(t - 1)^2}}$$

and the absolute value of the expression on the right side increases without bound for values of t closer and closer to 1.

In the general case, the curve represented by the equation $y = f(x)$ has a tangent parallel to the y-axis at the point with coordinates $(a, f(a))$ if and only if the absolute values of the slopes of lines joining this point to nearby points on the curve become larger and larger, as the points are taken closer and closer to $(a, f(a))$. That is, if we define g on all points on the domain of f, except a, by $g(t) = \dfrac{f(t) - f(a)}{t - a}$, then $|g(t)|$ can be made as large as we wish for values of t sufficiently close to a. This is shortened as $\lim |g(t)| = \infty$. (It must be carefully noted that here we use the symbol ∞ as part of a shorthand notation for the behaviour of a function, and *not as a concept in itself*). Thus we have the following extension of the last result:

3.2 Proposition. *Let $y = f(x)$ be the equation of a plane curve, where f is a real valued function on a subset D of real numbers, and let a be a number in D. If f is differentiable at all points in an open interval containing a and contained in D, except at a, and $\lim_{t \to a} \left| \dfrac{f(t) - f(a)}{t - a} \right| = \infty$, then the curve has a tangent parallel to the y-axis at $(a, f(a))$ and its equation is $x = a$.* $\square$

Next we consider a curve represented by an equation of the form $x = f(y)$, where f is again a real valued function on a subset D of real numbers. In this case, any point on the curve has coordinates of the form $(f(t), t)$ for some t in D. Let $(f(a), a)$ be a fixed point on the curve. To find the slope of the tangent at this point, we consider as before the limit of the slopes of the lines joining this fixed points to other points $(t, f(t))$ near it. Note that these slopes are given by

$$\frac{t - a}{f(t) - f(a)}$$

The problem is that there may be points $(f(t), t)$ on the curve with $f(t) = f(a)$, even if $t \neq a$, and the line joining it with $(f(a), a)$ is parallel to the y-axis, so that its slope is not defined.

For example, see the curve represented by $x = y^3 - 3y$ given in Figure 3. In this figure, P, Q and R are points on the curve, which lie along a line parallel to the y axis. So, in trying to compute the slope of

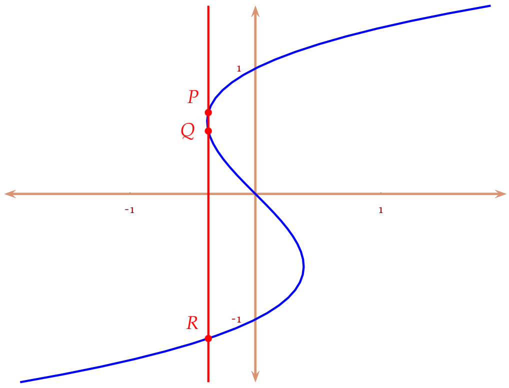

Figure 3: Vertical secant

the tangent at P, for example, as the limit of the slopes of lines joining P to nearby points on the curve, the points below P we choose must all be between P and Q.

So to find the slope at the point $(f(a), a)$ using limits, we must assume that there is an open interval containing a, in the domain of f, with $f(t) \neq f(a)$ for all $t \neq a$ in the interval. Then for any point $(f(t), t)$ with t in this interval, the slope of the line joining this point with $(f(a), a)$ is $\dfrac{t-a}{f(t)-f(a)}$, so that the slope of the tangent at $(f(a), a)$ is

$$\lim_{t \to a} \frac{t-a}{f(t)-f(a)}$$

Now it can be proved that for a real function g, if g is non-zero for all points in an open interval containing a number a and if $\lim_{x \to a} g(x) \neq 0$, then function g, then

$$\lim_{x \to a} \frac{1}{g(x)} = \frac{1}{\lim_{x \to a} g(x)}$$

So in our discussion, if $f(t) \neq f(a)$ in an open interval containing a and if

$$f'(a) = \lim_{t \to a} \frac{f(t)-f(a)}{t-a} \neq 0$$

then we have the slope of the tangent to $x = f(y)$ at $(f(a), a)$ is

$$\lim_{t \to a} \frac{t-a}{f(t)-f(a)} = \frac{1}{f'(a)}$$

Now it can be proved that if for some function h, we have $\lim_{x \to a} h(x) \neq 0$, then there exists an open interval containing a such that $f(x) \neq 0$ for any x in this interval. Thus the assumption $f'(a) \neq 0$ implies there is an open interval containing a such that $\dfrac{f(t)-f(a)}{t-a} \neq 0$ for any t in this interval, which means $f(t) \neq f(a)$ for any t in this interval. So we need not make this assumption separately in our discussion above. Thus we have the following:

3.3 Proposition. *Let $x = f(y)$ be the equation of a plane curve, where f is a real valued function on a subset D of real numbers and let a be a number in D. If D contains an open interval containing a such that f is differentiable in this interval with $f'(a) \neq 0$, then the slope of the tangent to the curve at the point $(f(a), a)$ is $\dfrac{1}{f'(a)}$ and consequently, the equation of the tangent is $x = f(a) + f'(a)(y-a)$.* □

For curves represented as $x = f(y)$, tangents parallel to the x-axis also cannot be found by differentiation. For example, consider the curve given by the equation $x = \sqrt[3]{y-1} + y$. If we define f on $\mathbb{R}$ by $f(t) = \sqrt[3]{t-1} + t$, then $x = f(y)$ and

$$f'(y) = \frac{1}{3\sqrt[3]{(y-1)^2}} + 1$$

for $y \neq 1$ and f' is not defined at 1.

Here, $f(1) = 1$, so that $(1, 1)$ is a point on this curve, and the slope of the line joining this point and another point $(f(t), t)$ on the curve is

$$\frac{t-1}{f(t) - f(1)} = \frac{t-1}{\sqrt[3]{t-1} + t - 1} = \frac{\sqrt[3]{(t-1)^2}}{1 + \sqrt[3]{(t-1)^2}}$$

and

$$\lim_{t \to 1} \frac{\sqrt[3]{(t-1)^2}}{1 + \sqrt[3]{(t-1)^2}} = 0$$

so that the slope of the tangent at $(1, 1)$ is 0. In other words, the tangent is the line $y = 1$, parallel to the x-axis.

In the general case of a curve represented by an equation of the form $x = f(y)$, and a point $(f(a), a)$ on the curve, if $\lim_{t \to a} \left| \frac{f(t)-f(a)}{t-a} \right| = \infty$ and there is an open interval containing a and contained in the domain of f such that $f(t) \neq f(a)$ for any t in this interval, then $\lim_{t \to a} \left| \frac{t-a}{f(t)-f(a)} \right| = 0$; and hence the slope of the tangent to the curve at $(f(a), a)$ is zero. In other words, the tangent is parallel to the x-axis. Thus we have another extension of Proposition 3.3 as follows:

3.4 Proposition. *Let $x = f(y)$ be the equation of a plane curve, where f is a real valued function on a subset D of real numbers and let a be a number in D. If D contains an open interval containing a with $f(t) \neq f(a)$ for any t in this interval and if f is differentiable in this interval with $\lim_{t \to a} \left| \frac{f(t) - f(a)}{t - a} \right| = \infty$, then the tangent to the curve at $(f(a), a)$ is paralel to the x-axis and its equation is $y = a$.* □

Link to the HTML version of this document.