An easier derivation of the curvature formula from first principles

Robert Ferguson
Florida, USA
bobferg13@comcast.net

Introduction

The radius of curvature formula is usually introduced in a university calculus course. Its proof is not included in most high school calculus courses and even some first-year university calculus courses because many students find the calculus used difficult (see Larson, Hostetler & Edwards, 2007, pp. 870–872). Fortunately, there is an easier way to motivate and prove the radius of curvature formula without using formal calculus. An additional benefit is that the proof provides the coordinates of the centre of the corresponding circle of curvature. Understanding the proof requires only what advanced high school students already know: e.g., algebra, a little geometry about circles, and the intuition that both a circle and a curve have a slope and curvature. The proof is suitable for first-year university and advanced high school students.

In terms of algebra, students need to know how to solve two simultaneous linear equations with two unknowns. In terms of geometry, students need to know that:

1. The radius of a circle drawn to a point of tangency between the circle and a tangent line is perpendicular to the tangent line.
2. If two separate lines are tangent to a circle at two different points, the lines drawn perpendicular to the tangent lines at their points of tangency intersect each other at the circle's centre.
3. Each perpendicular line's segment from its point of tangency to the point of intersection is a radius.

The paper's development of the radius of curvature formula can be used as an insightful application of the mathematics advanced high school students already know.

The procedure for finding the radius of curvature

Consider a curve given by a twice differentiable function $y = f(x)$.[1] This function gives a curve $(x, f(x))$ consisting of points in the Cartesian plane. Here is the procedure for finding the centre of curvature at any point (x_0, y_0) on the curve. Definitions and formulae for the radius and circle of curvature follow naturally.

Students can gain intuition by considering a circle. Students will readily accept that a circle's curvature is constant. Considering any two points on the circle, the perpendiculars to the tangents at these two points will meet at the centre of the circle. This is true for any two points on the circle, so students will readily accept that the circle's centre is the centre of curvature for any point on the circle, the radius of curvature at any point on the circle is the radius of the circle, and the circle of curvature is the circle itself.

In general:

- Consider two close points on the curve, (x_0, y_0) and (x_1, y_1).
- Consider the line perpendicular to the tangent to the curve at (x_0, y_0) and the line perpendicular to the tangent to the curve at (x_1, y_1).
- Find the point of intersection of these two perpendicular lines and denote it by (x, y).[2]
- Keeping x_0 fixed. (x, y) may approach a limiting point as x_1 approaches x_0.[3] This point is defined to be the curve's centre of curvature at (x_0, y_0).
- The radius of curvature at (x_0, y_0) is defined to be the distance from (x, y), the centre of curvature, to (x_0, y_0).
- The circle of curvature at (x_0, y_0) is defined to be the circle whose centre is the centre of curvature at (x_0, y_0) and whose radius is the radius of curvature at (x_0, y_0).

The intuition behind the procedure

The intuition behind the procedure is that:

- The curvature of a circle usually is defined as the reciprocal of its radius (the smaller the radius, the greater the curvature).
- A circle's curvature varies from infinity to zero as its radius varies from zero to infinity.
- A circle's curvature is a monotonically decreasing function of its radius. Given a curvature, there is only one radius, hence only one circle that matches the given curvature.

1 Twice differentiable can be presented to advanced high school students as 'smooth'.

2 There will be a point of intersection for most functions of interest. An obvious exception is a straight line.

3 This will be true for most functions of interest, except at particular points, e.g., critical points or points of inflection.

- For any two close points on a curve, at least one circle minimises the absolute area between the curve and the circle between the two points. This circle's radius can be viewed as approximating the curve's radius of curvature.
- A curve's curvature between two points approaches its curvature at a point as the two points approach each other.
- The curvature of a circle that minimises the absolute area between the curve and the circle between two close points on the curve approaches the curve's curvature as the two points approach each other.
- The radius of the absolute area minimizing circle approaches the curve's radius of curvature as the two points approach each other.

The algebra

Consider a curve given by:

$$y = f(x) \tag{1}$$

Consider two points on the curve, (x_0, y_0) and (x_1, y_1). The two points are presumed not to be problematic, e.g., critical points or points of inflection. Using the point-slope form of a line and denoting its slope by m_0, the tangent line to the curve at (x_0, y_0) is:

$$(y - y_0) = m_0(x - x_0) \tag{2}$$

Using the property that the slope of a line perpendicular to another line is the negative inverse of that line's slope, the line perpendicular to the tangent line at (x_0, y_0) is:

$$\left(y - y_0\right) = -\frac{\left(x - x_0\right)}{m_0} \tag{3}$$

Similarly, the corresponding tangent and perpendicular lines at the point are:

$$(y - y_1) = m_1(x - x_1) \tag{4}$$

and

$$\left(y - y_1\right) = -\frac{\left(x - x_1\right)}{m_1} \tag{5}$$

The intersection of the two perpendicular lines approximates the curve's centre of curvature. As x_1 approaches x_0, the intersection becomes the centre of the circle of curvature that matches exactly the curve's curvature at the point (x_0, y_0).

The intersection is the solution of equations (3) and (5).

Substitute from equation (3) into equation (5).

$$\left(y-y_0\right)=\left(y-y_1\right)+\left(y_1-y_0\right)=-\frac{\left(x-x_0\right)}{m_0} \tag{6}$$

$$\left(y-y_1\right)=-\frac{\left(x-x_0\right)}{m_0}-\left(y_1-y_0\right) \tag{7}$$

$$\left(x-x_1\right)=\left(x-x_0\right)-\left(x_1-x_0\right) \tag{8}$$

$$-\frac{\left(x-x_0\right)}{m_0}-\left(y_1-y_0\right)=-\frac{\left(x-x_0\right)}{m_1}+\frac{\left(x_1-x_0\right)}{m_1} \tag{9}$$

$$-\frac{\left(x-x_0\right)}{m_0}=-\frac{\left(x-x_0\right)}{m_1}+\frac{\left(x_1-x_0\right)}{m_1}+\left(y_1-y_0\right) \tag{10}$$

Since $f(x)$ is twice differentiable, $(y_1 - y_0) = m(x_1 - x_0)$, where m denotes the curve's slope somewhere in the interval $[x_0, x_1]$.

$$\left(x-x_0\right)\left(\frac{1}{m_1}-\frac{1}{m_0}\right)=\frac{\left(x_1-x_0\right)}{m_1}+m\left(x_1-x_0\right) \tag{11}$$

$$\left(x-x_0\right)\left(\frac{m_0-m_1}{m_0 m_1}\right)=\left(\frac{1}{m_1}+m\right)\left(x_1-x_0\right) \tag{12}$$

$$\left(x-x_0\right)=\frac{\left(\frac{1}{m_1}+m\right)}{\left(\frac{m_0-m_1}{m_0 m_1}\right)}\left(x_1-x_0\right) \tag{13}$$

$$\left(x-x_0\right)=\frac{m_0\left(1+m m_1\right)}{\left(m_0-m_1\right)}\left(x_1-x_0\right) \tag{14}$$

Substitute equation (14) into equation (3) to obtain

$$\left(y-y_0\right)=-\frac{\left(x-x_0\right)}{m_0}=-\frac{\left(1+m m_1\right)}{\left(m_0-m_1\right)}\left(x_1-x_0\right) \tag{15}$$

One way of helping students gain intuition about the original function and the function that represents its slope (its derivative) is to point out that:

- a function (the curve) has a slope at each point;
- these slopes can be viewed as another function;
- just as the tangent line to the original curve at (x_0, y_0), e.g., $y - y_0 = m_0(x - x_0)$, is a good approximation to the original curve near (x_0, y_0), the tangent line to the slope function at (x_0, y_0), e.g., $(m - m_0) = n_0(x - x_0)$, where n_0 denotes the slope (rate of change) of the slope function is a good approximation to the slope function near (x_0, y_0).

The $(m_1 - m_0)$ term in equation (15) is the difference in the original curve's slopes at the two points (x_1, y_1) and (x_0, y_0), respectively. Since the function is twice differentiable,

$$\left(m_1 - m_0\right) = n\left(x_1 - x_0\right) \tag{16}$$

where n denotes the slope function's slope somewhere in the interval $[x_0, x_1]$.

Substitute equation from (16) into equations (14) and (15).

$$\left(x - x_0\right) = -\frac{m_0\left(1 + mm_1\right)}{n} \tag{17}$$

$$\left(y - y_0\right) = \frac{\left(1 + mm_1\right)}{n} \tag{18}$$

As x_1 approaches x_0, all the slopes approach their values at (x_0, y_0), hence $(x - x_0)$ and $(y - y_0)$ approach:

$$\left(x - x_0\right) = -\frac{m_0\left(1 + m_0^2\right)}{n_0} \tag{19}$$

$$\left(y - y_0\right) = \frac{\left(1 + m_0^2\right)}{n_0} \tag{20}$$

Equations (19) and (20) give the x and y coordinates of the centre of the circle corresponding to the radius of curvature at (x_0, y_0).

The radius of curvature, R, is the distance between the point (x, y) given by equations (19) and (20) and the point (x_0, y_0).

$$\begin{aligned} R^2 &= \left(x - x_0\right)^2 + \left(y - y_0\right)^2 \\ &= \frac{m_0^2\left(1 + m_0^2\right)^2}{n_0^2} + \frac{\left(1 + m_0^2\right)^2}{n_0^2} \\ &= \frac{\left(1 + m_0^2\right)^3}{n_0^2} \end{aligned} \tag{21}$$

$$R = \left|\frac{\left(1 + m_0^2\right)^{\frac{3}{2}}}{n_0}\right| \tag{22}$$

For students that have not had calculus, note that in calculus, the term m_0 is denoted by $\frac{dy}{dx}$ or $f'(x)$ or $y'(x)$ and the term n_0 is denoted by $\frac{d^2y}{dx^2}$ or $f''(x)$ or $y''(x)$, so that they can see that equation (22) is equivalent to the various formulas, below, found in calculus textbooks.

$$R = \left|\frac{\left(1 + \left(\frac{dy}{dx}\right)^2\right)^{\frac{3}{2}}}{\frac{d^2y}{dx^2}}\right| = \left|\frac{\left(1 + \left(f'(x)\right)^2\right)^{\frac{3}{2}}}{f''(x)}\right| = \left|\frac{\left(1 + \left(y'(x)\right)^2\right)^{\frac{3}{2}}}{y''(x)}\right| \tag{23}$$

Teaching the radius of curvature formula

First year university and advanced high school students can evaluate equation (22) without calculus by evaluating the slopes (derivatives) and changes in slopes (second derivatives) using an Excel spreadsheet and suitably small values for $(x_1 - x_0)$ and the changes in x used to compute the slopes and changes in slopes. Additional insight can be gained by evaluating (x, y) found from equations (14) and (15) for a decreasing sequence of values for $(x_1 - x_0)$, plotting the resultant points, (x, y), and observing how they approach a limiting point that is the centre of the circle of curvature corresponding to the curve's radius of curvature.

Example

Consider the function $y = x^3$ and the points on this curve at (0.5, 0.125) and $(0.5 + \Delta x,).125 + \Delta y)$. Figure 1 plots the function and the (x, y) points from equations (14) and (15) for several values of Δx ranging from 0.5 to 0.0001. The slopes at the two points are computed using changes in x that are one-tenth of the Δx values. Figure 1 also contains a plot of the radius of curvature for each value of Δx and a plot of the function.

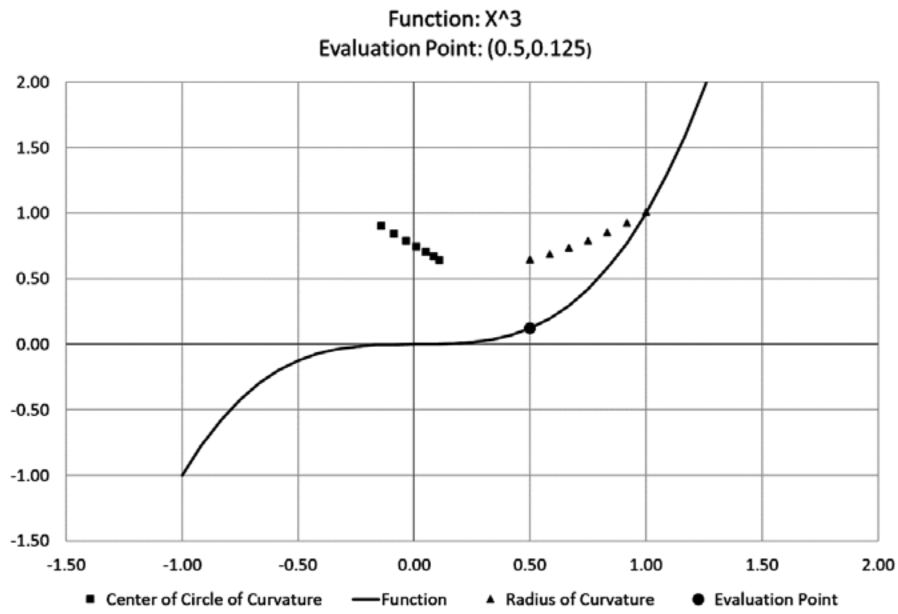

Figure 1

The iterations for the centre of the circle of curvature are shown by the square markers, which run from upper left to lower right as Δx changes from 0.5 to 0.0001. The points converge to the coordinates given by equations (19) and (20) evaluated with calculus of (0.1094, 0.6458). The iterations for the radius of curvature are shown by the triangular markers, which run from upper right to lower left as Δx changes from 0.5 to 0.0001. The points converge to the radius of curvature given by equation (23), about 0.6510.

Conclusion

This paper presented an easier way to motivate and prove the radius of curvature formula that uses mostly high school mathematics and no formal calculus. An additional benefit is that the proof provides the coordinates of the centre of the corresponding circle of curvature. Understanding the proof requires what high school students already know: algebra, a little geometry about circles, and the intuition that both a circle and a curve have a slope and curvature. The proof is suitable for first-year university and advanced high school students.

Advanced high school students can find a function's radius of curvature and the centre of the corresponding circle of curvature without use of calculus by evaluating the required slopes and changes in slopes along a curve using an Excel spreadsheet and suitably small distances between two points on the curve. Additional insight can be gained by doing so for a decreasing sequence of values for the distance between the points, plotting the resulting sequence of radii of curvature and their centres of the corresponding circles of curvature, and observing how they approach a limit.

Reference

Larson, R., Hostetler, R. P. & Edwards, B. H. (2007). *Calculus: Early transcendental functions* (4th ed.). Boston, MA: Houghton Mifflin.