Calculus 1

Chapter 3. Derivatives

3.11. Linearization and Differentials—Examples and Proofs

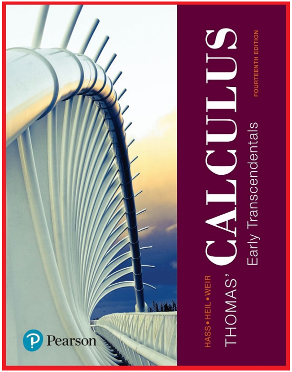

Table of contents

1. Exercise 3.11.2
2. Exercise 3.11.28
3. Exercise 3.11.38
4. Example 3.11.A
5. Exercise 3.11.44
6. Exercise 3.11.56
7. Exercise 3.11.58. Tolerance
8. Lemma 3.11.A
9. Theorem 3.2. The Chain Rule

Exercise 3.11.2

Exercise 3.11.2. Find the linearization $L(x)$ of $f(x) = \sqrt{x^2+9}$ at $x = a = -4$.

Solution. We write $f(x) = (x^2+9)^{1/2}$ so that by the Chain Rule (Theorem 3.2) we have $f'(x) = (1/2)(x^2+9)^{-1/2}[2x] = x/\sqrt{x^2+9}$.

Exercise 3.11.2

Exercise 3.11.2. Find the linearization $L(x)$ of $f(x) = \sqrt{x^2+9}$ at $x = a = -4$.

Solution. We write $f(x) = (x^2+9)^{1/2}$ so that by the Chain Rule (Theorem 3.2) we have $f'(x) = (1/2)(x^2+9)^{-1/2}\overset{\curvearrowright}{[2x]} = x/\sqrt{x^2+9}$. Now $f(a) = f(-4) = \sqrt{(-4)^2+9} = \sqrt{25} = 5$ and $f'(a) = f'(-4) = (-4)/\sqrt{(-4)^2+9} = -4/5$, so

$$L(x) = f(a) + f'(a)(x-a) = f(-4) + f'(-4)(x-(-4))$$

$$= 5 + (-4/5)(x+4) = \boxed{(-4/5)x - 9/5}.$$

□

Exercise 3.11.2

Exercise 3.11.2. Find the linearization $L(x)$ of $f(x) = \sqrt{x^2+9}$ at $x = a = -4$.

Solution. We write $f(x) = (x^2+9)^{1/2}$ so that by the Chain Rule (Theorem 3.2) we have $f'(x) = (1/2)(x^2+9)^{-1/2}\overset{\curvearrowright}{[2x]} = x/\sqrt{x^2+9}$. Now $f(a) = f(-4) = \sqrt{(-4)^2+9} = \sqrt{25} = 5$ and $f'(a) = f'(-4) = (-4)/\sqrt{(-4)^2+9} = -4/5$, so

$$L(x) = f(a) + f'(a)(x-a) = f(-4) + f'(-4)(x-(-4))$$

$$= 5 + (-4/5)(x+4) = \boxed{(-4/5)x - 9/5}.$$

□

Exercise 3.11.28

Exercise 3.11.28. Find dy when $y = \sec(x^2 - 1)$.

Solution. With $f(x) = \sec(x^2 - 1)$ we have by the Chain Rule (Theorem 3.2) that

$$dy = f'(x)\,dx = \sec(x^2 - 1)\tan(x^2 - 1)[2x]\,dx$$
$$= \boxed{2x\sec(x^2 - 1)\tan(x^2 - 1)\,dx}.$$

□

Exercise 3.11.28

Exercise 3.11.28. Find dy when $y = \sec(x^2 - 1)$.

Solution. With $f(x) = \sec(x^2 - 1)$ we have by the Chain Rule (Theorem 3.2) that

$$dy = f'(x)\,dx = \sec(x^2 - 1)\tan(x^2 - \overset{\curvearrowright}{1)}[2x]\,dx$$

$$= \boxed{2x\sec(x^2 - 1)\tan(x^2 - 1)\,dx}.$$

□

Exercise 3.11.38

Exercise 3.11.38. Find dy when $y = e^{\tan^{-1}\sqrt{x^2+1}}$.

Solution. With $f(x) = e^{\tan^{-1}\sqrt{x^2+1}}$ we have by the Chain Rule (Theorem 3.2) that

$$dy = f'(x)\,dx = e^{\tan^{-1}\sqrt{x^2+1}} \left[\frac{1}{1+(\sqrt{x^2+1})^2} [(1/2)(x^2+1)^{-1/2}[2x]] \right] dx$$

$$= \boxed{\frac{x}{(x^2+2)\sqrt{x^2+1}} e^{\tan^{-1}\sqrt{x^2+1}}\,dx}.$$

□

Exercise 3.11.38

Exercise 3.11.38. Find dy when $y = e^{\tan^{-1}\sqrt{x^2+1}}$.

Solution. With $f(x) = e^{\tan^{-1}\sqrt{x^2+1}}$ we have by the Chain Rule (Theorem 3.2) that

$$dy = f'(x)\,dx = e^{\tan^{-1}\sqrt{x^2+1}} \overset{\curvearrowright}{\left[\frac{1}{1+(\sqrt{x^2+1})^2} \overset{\curvearrowright}{[}(1/2)(x^2+1)^{-1/2} \overset{\curvearrowright}{[}2x]] \right]} dx$$

$$= \boxed{\frac{x}{(x^2+2)\sqrt{x^2+1}} e^{\tan^{-1}\sqrt{x^2+1}}\,dx}.$$

□

Example 3.11.A

Example 3.11.A. Use differentials to estimate the value of $\sin 31^\circ$.

Solution. First, we have $31^\circ = 30^\circ + 1^\circ = \pi/6 + \pi/180$ (radians). We take $f(x) = \sin x$ so that $f'(x) = \cos x$. With $a = \pi/6$ and $\Delta x = dx = \pi/180$, we have:

Example 3.11.A

Example 3.11.A. Use differentials to estimate the value of $\sin 31^\circ$.

Solution. First, we have $31^\circ = 30^\circ + 1^\circ = \pi/6 + \pi/180$ (radians). We take $f(x) = \sin x$ so that $f'(x) = \cos x$. With $a = \pi/6$ and $\Delta x = dx = \pi/180$, we have:

$$\sin 31^\circ = \sin(\pi/6 + \pi/180) = f(a + \Delta x) = f(a) + \Delta y$$

$$\approx f(a) + dy = f(a) + f'(a)\,dx = \sin(\pi/6) + \cos(\pi/6)(\pi/180)$$

$$= (1/2) + (\sqrt{3}/2)(\pi/180) = 1/2 + \sqrt{3}\pi/360 \approx \boxed{0.515115}. \quad \square$$

Example 3.11.A

Example 3.11.A. Use differentials to estimate the value of $\sin 31^\circ$.

Solution. First, we have $31^\circ = 30^\circ + 1^\circ = \pi/6 + \pi/180$ (radians). We take $f(x) = \sin x$ so that $f'(x) = \cos x$. With $a = \pi/6$ and $\Delta x = dx = \pi/180$, we have:

$$\sin 31^\circ = \sin(\pi/6 + \pi/180) = f(a + \Delta x) = f(a) + \Delta y$$

$$\approx f(a) + dy = f(a) + f'(a)\,dx = \sin(\pi/6) + \cos(\pi/6)(\pi/180)$$

$$= (1/2) + (\sqrt{3}/2)(\pi/180) = 1/2 + \sqrt{3}\pi/360 \approx \boxed{0.515115}. \qquad \square$$

Using a calculator, we have $\sin 31^\circ \approx 0.515038$. So linearization gives an approximation that is accurate to three decimal places (but not to four decimal places).

Example 3.11.A

Example 3.11.A. Use differentials to estimate the value of $\sin 31^\circ$.

Solution. First, we have $31^\circ = 30^\circ + 1^\circ = \pi/6 + \pi/180$ (radians). We take $f(x) = \sin x$ so that $f'(x) = \cos x$. With $a = \pi/6$ and $\Delta x = dx = \pi/180$, we have:

$$\sin 31^\circ = \sin(\pi/6 + \pi/180) = f(a + \Delta x) = f(a) + \Delta y$$

$$\approx f(a) + dy = f(a) + f'(a)\,dx = \sin(\pi/6) + \cos(\pi/6)(\pi/180)$$

$$= (1/2) + (\sqrt{3}/2)(\pi/180) = 1/2 + \sqrt{3}\pi/360 \approx \boxed{0.515115}. \quad \square$$

Using a calculator, we have $\sin 31^\circ \approx 0.515038$. So linearization gives an approximation that is accurate to three decimal places (but not to four decimal places).

Exercise 3.11.44

Exercise 3.11.44. For $f(x) = x^3 - 2x + 3$, $x_0 = 2$, and $dx = 0.1$, find: **(a)** the change $\Delta f = f(x_0 + dx) - f(x_0)$, **(b)** the value of the estimate $df = f'(x_0)\,dx$, and **(c)** the approximation error $|\Delta f - df|$.

Solution. First, $f'(x) = 3x^2 - 2$.

Exercise 3.11.44

Exercise 3.11.44. For $f(x) = x^3 - 2x + 3$, $x_0 = 2$, and $dx = 0.1$, find: **(a)** the change $\Delta f = f(x_0 + dx) - f(x_0)$, **(b)** the value of the estimate $df = f'(x_0)\,dx$, and **(c)** the approximation error $|\Delta f - df|$.

Solution. First, $f'(x) = 3x^2 - 2$.

(a) We have

$$\Delta f = f(x_0 + dx) - f(x_0) = f(2 + 0.1) - f(2) = f(2.1) - f(2)$$

$$= ((2.1)^3 - 2(2.1) + 3) - ((2)^3 - 2(2) + 3)) = 8.061 - 7 = \boxed{1.061}. \quad \square$$

Exercise 3.11.44

Exercise 3.11.44. For $f(x) = x^3 - 2x + 3$, $x_0 = 2$, and $dx = 0.1$, find: **(a)** the change $\Delta f = f(x_0 + dx) - f(x_0)$, **(b)** the value of the estimate $df = f'(x_0)\,dx$, and **(c)** the approximation error $|\Delta f - df|$.

Solution. First, $f'(x) = 3x^2 - 2$.

(a) We have

$$\Delta f = f(x_0 + dx) - f(x_0) = f(2 + 0.1) - f(2) = f(2.1) - f(2)$$

$$= ((2.1)^3 - 2(2.1) + 3) - ((2)^3 - 2(2) + 3)) = 8.061 - 7 = \boxed{1.061}. \qquad \square$$

(b) Next, $df = f'(x_0)\,dx = f'(2)\,dx = (3(2)^2 - 2)(0.1) = \boxed{1}$. $\square$

Exercise 3.11.44

Exercise 3.11.44. For $f(x) = x^3 - 2x + 3$, $x_0 = 2$, and $dx = 0.1$, find: **(a)** the change $\Delta f = f(x_0 + dx) - f(x_0)$, **(b)** the value of the estimate $df = f'(x_0)\,dx$, and **(c)** the approximation error $|\Delta f - df|$.

Solution. First, $f'(x) = 3x^2 - 2$.

(a) We have

$$\Delta f = f(x_0 + dx) - f(x_0) = f(2 + 0.1) - f(2) = f(2.1) - f(2)$$

$$= ((2.1)^3 - 2(2.1) + 3) - ((2)^3 - 2(2) + 3)) = 8.061 - 7 = \boxed{1.061}. \quad \square$$

(b) Next, $df = f'(x_0)\,dx = f'(2)\,dx = (3(2)^2 - 2)(0.1) = \boxed{1}$. $\square$

(c) Finally, $|\Delta f - df| = |1.061 - 1| = \boxed{0.061}$. $\square$

Exercise 3.11.44

Exercise 3.11.44. For $f(x) = x^3 - 2x + 3$, $x_0 = 2$, and $dx = 0.1$, find: **(a)** the change $\Delta f = f(x_0 + dx) - f(x_0)$, **(b)** the value of the estimate $df = f'(x_0)\,dx$, and **(c)** the approximation error $|\Delta f - df|$.

Solution. First, $f'(x) = 3x^2 - 2$.

(a) We have

$$\Delta f = f(x_0 + dx) - f(x_0) = f(2 + 0.1) - f(2) = f(2.1) - f(2)$$

$$= ((2.1)^3 - 2(2.1) + 3) - ((2)^3 - 2(2) + 3)) = 8.061 - 7 = \boxed{1.061}. \quad \square$$

(b) Next, $df = f'(x_0)\,dx = f'(2)\,dx = (3(2)^2 - 2)(0.1) = \boxed{1}$. $\square$

(c) Finally, $|\Delta f - df| = |1.061 - 1| = \boxed{0.061}$. $\square$

Exercise 3.11.56

Exercise 3.11.56. The edge x of a cube is measured with an error of at most 0.5%. What is the maximum corresponding percentage error in computing the cube's: **(a)** surface area? **(b)** volume?

Proof. The surface area of such a cube is $A = 6x^2$ and the volume of such a cube is $V = x^3$. The edge x is measured with an error of at most 0.5%, so the percentage change in the edge is $dx/x \times 100\% \leq 0.5\%$ or $dx \leq 0.005x$.

Exercise 3.11.56

Exercise 3.11.56. The edge x of a cube is measured with an error of at most 0.5%. What is the maximum corresponding percentage error in computing the cube's: **(a)** surface area? **(b)** volume?

Proof. The surface area of such a cube is $A = 6x^2$ and the volume of such a cube is $V = x^3$. The edge x is measured with an error of at most 0.5%, so the percentage change in the edge is $dx/x \times 100\% \leq 0.5\%$ or $dx \leq 0.005x$.

(a) Since $A = 6x^2$ then $dA = 12x\,dx$ and the percentage change in area is

$$\frac{dA}{A}\times 100\% = \frac{12x\,dx}{6x^2}\times 100\% \leq \frac{12x(0.005x)}{6x^2}\times 100\% = 0.010\times 100\% = \boxed{1\%}.$$

Exercise 3.11.56

Exercise 3.11.56. The edge x of a cube is measured with an error of at most 0.5%. What is the maximum corresponding percentage error in computing the cube's: **(a)** surface area? **(b)** volume?

Proof. The surface area of such a cube is $A = 6x^2$ and the volume of such a cube is $V = x^3$. The edge x is measured with an error of at most 0.5%, so the percentage change in the edge is $dx/x \times 100\% \leq 0.5\%$ or $dx \leq 0.005x$.

(a) Since $A = 6x^2$ then $dA = 12x\,dx$ and the percentage change in area is

$$\frac{dA}{A}\times 100\% = \frac{12x\,dx}{6x^2}\times 100\% \leq \frac{12x(0.005x)}{6x^2}\times 100\% = 0.010\times 100\% = \boxed{1\%}.$$

(b) Since $V = x^3$ then $dV = 3x^2\,dx$ and the percentage change in volume is

$$\frac{dV}{V}\times 100\% = \frac{3x^2\,dx}{x^3}\times 100\% \leq \frac{3x^2(0.005x)}{x^3}\times 100\% = 0.015\times 100\% = \boxed{1.5\%}.$$

□

Exercise 3.11.56

Exercise 3.11.56. The edge x of a cube is measured with an error of at most 0.5%. What is the maximum corresponding percentage error in computing the cube's: **(a)** surface area? **(b)** volume?

Proof. The surface area of such a cube is $A = 6x^2$ and the volume of such a cube is $V = x^3$. The edge x is measured with an error of at most 0.5%, so the percentage change in the edge is $dx/x \times 100\% \leq 0.5\%$ or $dx \leq 0.005x$.

(a) Since $A = 6x^2$ then $dA = 12x\,dx$ and the percentage change in area is

$$\frac{dA}{A}\times 100\% = \frac{12x\,dx}{6x^2}\times 100\% \leq \frac{12x(0.005x)}{6x^2}\times 100\% = 0.010\times 100\% = \boxed{1\%}.$$

(b) Since $V = x^3$ then $dV = 3x^2\,dx$ and the percentage change in volume is

$$\frac{dV}{V}\times 100\% = \frac{3x^2\,dx}{x^3}\times 100\% \leq \frac{3x^2(0.005x)}{x^3}\times 100\% = 0.015\times 100\% = \boxed{1.5\%}.$$

□

Exercise 3.11.58

Exercise 3.11.58. Tolerance (a) About how accurately must the interior diameter of a 10-m-high cylindrical storage tank be measured to calculate the tank's volume to within 1% of its true value? **(b)** About how accurately must the tank's exterior diameter be measured to calculate the amount of paint it will take to paint the side of the tank to within 5% of the true amount?

Solution. The volume of a cylinder of diameter $D = 2r$ and height h is $V = \pi r^2 h$. So here,

$$V = \pi(D/2)^2(10) = 5\pi D^2/2 \text{ m}^3 \text{ and } dV = 5\pi[2D]/2\, dD = 5\pi D\, dD \text{ m}^3.$$

The surface area of the side of a cylinder of diameter $D = 2r$ and height h is

$$A = 2\pi rh = 2\pi(D/2)(10) = 10\pi D \text{ m}^2 \text{ and } dA = 10\pi dD \text{ m}^2.$$

Exercise 3.11.58

Exercise 3.11.58. Tolerance (a) About how accurately must the interior diameter of a 10-m-high cylindrical storage tank be measured to calculate the tank's volume to within 1% of its true value? **(b)** About how accurately must the tank's exterior diameter be measured to calculate the amount of paint it will take to paint the side of the tank to within 5% of the true amount?

Solution. The volume of a cylinder of diameter $D = 2r$ and height h is $V = \pi r^2 h$. So here,

$$V = \pi(D/2)^2(10) = 5\pi D^2/2 \text{ m}^3 \text{ and } dV = 5\pi[2D]/2\, dD = 5\pi D\, dD \text{ m}^3.$$

The surface area of the side of a cylinder of diameter $D = 2r$ and height h is

$$A = 2\pi rh = 2\pi(D/2)(10) = 10\pi D \text{ m}^2 \text{ and } dA = 10\pi dD \text{ m}^2.$$

Exercise 3.11.58 (continued)

Solution (continued). (a) We want $dV/V \times 100\% = 1\%$, so we require the percentage change in volume to satisfy

$$\frac{5\pi D\,dD}{5\pi D^2/2} \times 100\% = \frac{2dD}{D} \times 100\% = 1\%,$$

from which we need $dD/D \times 100\% = (1/2)\%$. That is, we need D to be measured with an accuracy of $(1/2)\% = \boxed{0.5\%}$.

(b) We want $dA/A \times 100\% = 5\%$, so we require the percentage change in surface to satisfy

$$\frac{10\pi dD}{10\pi D} \times 100\% = \frac{dD}{D} \times 100\% = 5\%,$$

from which we need $dD/D \times 100\% = 5\%$. That is, we need D to be measured with an accuracy of $\boxed{5\%}$. □

Exercise 3.11.58 (continued)

Solution (continued). (a) We want $dV/V \times 100\% = 1\%$, so we require the percentage change in volume to satisfy

$$\frac{5\pi D\,dD}{5\pi D^2/2} \times 100\% = \frac{2dD}{D} \times 100\% = 1\%,$$

from which we need $dD/D \times 100\% = (1/2)\%$. That is, we need D to be measured with an accuracy of $(1/2)\% = \boxed{0.5\%}$.

(b) We want $dA/A \times 100\% = 5\%$, so we require the percentage change in surface to satisfy

$$\frac{10\pi dD}{10\pi D} \times 100\% = \frac{dD}{D} \times 100\% = 5\%,$$

from which we need $dD/D \times 100\% = 5\%$. That is, we need D to be measured with an accuracy of $\boxed{5\%}$. □

Lemma 3.11.A

Lemma 3.11.A. If $y = f(x)$ is differentiable at $x = a$ and x changes from a to $a + \Delta x$, the corresponding change Δy in f is given by $\Delta y = f'(a)\,\Delta x + \varepsilon\,\Delta x$ in which $\varepsilon \to 0$ as $\Delta x \to 0$.

Proof. The approximation error $\Delta f - df$ at $x = a$ is

$$\Delta f - df = \Delta f - f'(a)\,dx = \Delta f - f'(a)\Delta x = (f(a+\Delta x) - f(a)) - f'(a)\Delta x$$

$$= \left(\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)\Delta x = \varepsilon\,\Delta x \qquad (*)$$

where $\varepsilon = \left(\dfrac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)$.

Lemma 3.11.A

Lemma 3.11.A. If $y = f(x)$ is differentiable at $x = a$ and x changes from a to $a + \Delta x$, the corresponding change Δy in f is given by $\Delta y = f'(a)\,\Delta x + \varepsilon\,\Delta x$ in which $\varepsilon \to 0$ as $\Delta x \to 0$.

Proof. The approximation error $\Delta f - df$ at $x = a$ is

$$\Delta f - df = \Delta f - f'(a)\,dx = \Delta f - f'(a)\Delta x = (f(a+\Delta x) - f(a)) - f'(a)\Delta x$$

$$= \left(\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)\Delta x = \varepsilon\,\Delta x \qquad (*)$$

where $\varepsilon = \left(\dfrac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)$. Since $f'(a)$ exists by hypothesis, then as $\Delta x \to 0$ the difference quotient $\dfrac{f(a+\Delta x) - f(a)}{\Delta x}$ approaches $f'(a)$, so that $\dfrac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a) = \varepsilon \to 0$ as $\Delta x \to 0$.

Lemma 3.11.A

Lemma 3.11.A. If $y = f(x)$ is differentiable at $x = a$ and x changes from a to $a + \Delta x$, the corresponding change Δy in f is given by $\Delta y = f'(a)\,\Delta x + \varepsilon\,\Delta x$ in which $\varepsilon \to 0$ as $\Delta x \to 0$.

Proof. The approximation error $\Delta f - df$ at $x = a$ is

$$\Delta f - df = \Delta f - f'(a)\,dx = \Delta f - f'(a)\Delta x = (f(a+\Delta x) - f(a)) - f'(a)\Delta x$$

$$= \left(\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)\Delta x = \varepsilon\,\Delta x \qquad (*)$$

where $\varepsilon = \left(\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)$. Since $f'(a)$ exists by hypothesis, then as $\Delta x \to 0$ the difference quotient $\frac{f(a+\Delta x) - f(a)}{\Delta x}$ approaches $f'(a)$, so that $\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a) = \varepsilon \to 0$ as $\Delta x \to 0$. Also, $\Delta y = \Delta f = f'(a)\Delta x + \varepsilon\,\Delta x$ from $(*)$, as claimed. □

Lemma 3.11.A

Lemma 3.11.A. If $y = f(x)$ is differentiable at $x = a$ and x changes from a to $a + \Delta x$, the corresponding change Δy in f is given by $\Delta y = f'(a)\,\Delta x + \varepsilon\,\Delta x$ in which $\varepsilon \to 0$ as $\Delta x \to 0$.

Proof. The approximation error $\Delta f - df$ at $x = a$ is

$$\Delta f - df = \Delta f - f'(a)\,dx = \Delta f - f'(a)\Delta x = (f(a+\Delta x) - f(a)) - f'(a)\Delta x$$

$$= \left(\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)\Delta x = \varepsilon\,\Delta x \qquad (*)$$

where $\varepsilon = \left(\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a)\right)$. Since $f'(a)$ exists by hypothesis, then as $\Delta x \to 0$ the difference quotient $\frac{f(a+\Delta x) - f(a)}{\Delta x}$ approaches $f'(a)$, so that $\frac{f(a+\Delta x) - f(a)}{\Delta x} - f'(a) = \varepsilon \to 0$ as $\Delta x \to 0$. Also, $\Delta y = \Delta f = f'(a)\Delta x + \varepsilon\,\Delta x$ from $(*)$, as claimed. □

Theorem 3.2

Theorem 3.2. The Chain Rule.
If $f(u)$ is differentiable at the point $u = g(x)$ and $g(x)$ is differentiable at x, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x, and $(f \circ g)'(x) = f'(g(x))[g'(x)]$.

Proof. Let x_0 be a point at which g is differentiable and suppose f is differentiable at $g(x_0)$. We show that $\left.\dfrac{dy}{dx}\right|_{x=x_0} = f'(g(x_0))g'(x_0)$ so that the claim then follows since x_0 is an arbitrary point satisfying the hypotheses.

Theorem 3.2

Theorem 3.2. The Chain Rule.
If $f(u)$ is differentiable at the point $u = g(x)$ and $g(x)$ is differentiable at x, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x, and $(f \circ g)'(x) = f'(g(x))[g'(x)]$.

Proof. Let x_0 be a point at which g is differentiable and suppose f is differentiable at $g(x_0)$. We show that $\left.\frac{dy}{dx}\right|_{x=x_0} = f'(g(x_0))g'(x_0)$ so that the claim then follows since x_0 is an arbitrary point satisfying the hypotheses.

Let Δx be an increment in x and let $\Delta u = g(x_0) - g(x_0 + \Delta x)$ and $\Delta y = f(u_0) - f(u_0 + \Delta u)$ be the corresponding increments in u and y. By Lemma 3.11.A, we have

$$\Delta u = g'(x_0)\,\Delta x + \varepsilon_1\,\Delta x = (g'(x_0) + \varepsilon_1)\,\Delta x,$$

where $\varepsilon_1 \to 0$ as $\Delta x \to 0$.

Theorem 3.2

Theorem 3.2. The Chain Rule.
If $f(u)$ is differentiable at the point $u = g(x)$ and $g(x)$ is differentiable at x, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x, and $(f \circ g)'(x) = f'(g(x))[g'(x)]$.

Proof. Let x_0 be a point at which g is differentiable and suppose f is differentiable at $g(x_0)$. We show that $\left.\dfrac{dy}{dx}\right|_{x=x_0} = f'(g(x_0))g'(x_0)$ so that the claim then follows since x_0 is an arbitrary point satisfying the hypotheses.

Let Δx be an increment in x and let $\Delta u = g(x_0) - g(x_0 + \Delta x)$ and $\Delta y = f(u_0) - f(u_0 + \Delta u)$ be the corresponding increments in u and y. By Lemma 3.11.A, we have

$$\Delta u = g'(x_0)\,\Delta x + \varepsilon_1\,\Delta x = (g'(x_0) + \varepsilon_1)\,\Delta x,$$

where $\varepsilon_1 \to 0$ as $\Delta x \to 0$.

Theorem 3.2 (continued 1)

Proof (continued). Similarly, with $u_0 = g(x_0)$,

$$\Delta y = f'(u_0)\,\Delta u + \varepsilon_2\,\Delta u = (f'(u_0) + \varepsilon_2)\,\Delta u,$$

where $\varepsilon_2 \to 0$ as $\Delta u \to 0$. Since g is differentiable at x_0 by hypothesis, then g is continuous at x_0 by Theorem 3.1 (Differentiability Implies Continuity) so $\lim_{\Delta x \to 0} \Delta u = \lim_{\Delta x \to 0}(g(x_0) - g(x_0 + \Delta x)) = 0$ and hence $\Delta u \to 0$ as $\Delta x \to 0$. We therefore have

$$\Delta y = (f'(u_0) + \varepsilon_2)\,\Delta u = (f'(u_0) + \varepsilon_2)(g'(x_0) + \varepsilon_1)\,\Delta x,$$

so

$$\frac{\Delta y}{\Delta x} = f'(u_0)g'(x_0) + \varepsilon_2 g'(x_0) + f'(u_0)\varepsilon_1 + \varepsilon_1\varepsilon_2.$$

As $\Delta x \to 0$ we have both $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$.

Theorem 3.2 (continued 1)

Proof (continued). Similarly, with $u_0 = g(x_0)$,

$$\Delta y = f'(u_0)\,\Delta u + \varepsilon_2\,\Delta u = (f'(u_0) + \varepsilon_2)\,\Delta u,$$

where $\varepsilon_2 \to 0$ as $\Delta u \to 0$. Since g is differentiable at x_0 by hypothesis, then g is continuous at x_0 by Theorem 3.1 (Differentiability Implies Continuity) so $\lim_{\Delta x \to 0} \Delta u = \lim_{\Delta x \to 0}(g(x_0) - g(x_0 + \Delta x)) = 0$ and hence $\Delta u \to 0$ as $\Delta x \to 0$. We therefore have

$$\Delta y = (f'(u_0) + \varepsilon_2)\,\Delta u = (f'(u_0) + \varepsilon_2)(g'(x_0) + \varepsilon_1)\,\Delta x,$$

so

$$\frac{\Delta y}{\Delta x} = f'(u_0)g'(x_0) + \varepsilon_2 g'(x_0) + f'(u_0)\varepsilon_1 + \varepsilon_1\varepsilon_2.$$

As $\Delta x \to 0$ we have both $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$.

Theorem 3.2 (continued 2)

Theorem 3.2. The Chain Rule.
If $f(u)$ is differentiable at the point $u = g(x)$ and $g(x)$ is differentiable at x, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x, and $(f \circ g)'(x) = f'(g(x))[g'(x)]$.

Proof (continued). As $\Delta x \to 0$ we have both $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$. Since $g'(x_0)$ and $f'(u_0)$ are some fixed numbers, then

$$\left.\frac{dy}{dx}\right|_{x=x_0} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \left(f'(u_0)g'(x_0) + \varepsilon_2 g'(x_0) + f'(u_0)\varepsilon_1 + \varepsilon_1\varepsilon_2\right)$$

$$= f'(u_0)g'(x_0) + (0)g'(x_0) + f'(u_0)(0) + (0)(0)$$
$$= f'(u_0)g'(x_0) = f'(g(x_0))g'(x_0).$$

Since x_0 is an arbitrary point satisfying the hypotheses, then we have $dy/dx = f'(g(x))g'(x)$, as claimed. □

Theorem 3.2 (continued 2)

Theorem 3.2. The Chain Rule.
If $f(u)$ is differentiable at the point $u = g(x)$ and $g(x)$ is differentiable at x, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x, and $(f \circ g)'(x) = f'(g(x))[g'(x)]$.

Proof (continued). As $\Delta x \to 0$ we have both $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$. Since $g'(x_0)$ and $f'(u_0)$ are some fixed numbers, then

$$\left.\frac{dy}{dx}\right|_{x=x_0} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \left(f'(u_0)g'(x_0) + \varepsilon_2 g'(x_0) + f'(u_0)\varepsilon_1 + \varepsilon_1\varepsilon_2\right)$$

$$= f'(u_0)g'(x_0) + (0)g'(x_0) + f'(u_0)(0) + (0)(0)$$

$$= f'(u_0)g'(x_0) = f'(g(x_0))g'(x_0).$$

Since x_0 is an arbitrary point satisfying the hypotheses, then we have $dy/dx = f'(g(x))g'(x)$, as claimed. □