Math 600 Day 1: Review of advanced Calculus

Ryan Blair

University of Pennsylvania

Thursday September 8, 2010
Outline

1 Differentiation
   • Chain Rule
   • Partial Derivatives
   • Critical Points
   • Inverse Function Theorem
   • The Implicit Function Theorem
Outline

1. Differentiation
   - Chain Rule
   - Partial Derivatives
   - Critical Points
   - Inverse Function Theorem
   - The Implicit Function Theorem
Definition

A function \( f : \mathbb{R}^m \to \mathbb{R}^n \) is said to be *differentiable at the point* \( x_0 \in \mathbb{R}^m \) if there is a linear map \( A : \mathbb{R}^m \to \mathbb{R}^n \) such that

\[
\lim_{h \to 0} \frac{|f(x_0 + h) - f(x_0) - A(h)|}{|h|} = 0
\]

The linear map \( A \) is called the derivative of \( f \) at \( x_0 \) and written as either \( f'(x_0) \) or as \( df_{x_0} \).
**Theorem**

*(Chain Rule)* Let

\[ \mathbb{R}^m \rightarrow f \rightarrow \mathbb{R}^n \rightarrow g \rightarrow \mathbb{R}^p \]

with \( x_0 \rightarrow f \rightarrow y_0 \rightarrow g \rightarrow z_0 \).

Suppose \( f \) is differentiable at \( x_0 \) with derivative \( f'(x_0) \) and that \( g \) is differentiable at \( y_0 \) with derivative \( g'(y_0) \).

Then the composition \( g \circ f \) is differentiable at \( x_0 \) with derivative

\[ (g \circ f)'(x_0) = g'(y_0)f'(x_0) \].
Proof of the Chain Rule.

In an intuitively taught calculus course, the truth of the chain rule is sometimes suggested by multiplying "fractions":

\[
\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.
\]

This argument comes to grief when nonzero changes in $x$ produce zero changes in $y$. The simple finesse is to avoid fractions, as follows.
Without loss of generality, and for ease of notation, we will assume that the points $x_0 \in \mathbb{R}^m$, $y_0 \in \mathbb{R}^n$ and $z_0 \in \mathbb{R}^p$ are all located at their respective origins.

We let $L = f'(x_0)$ and $M = g'(y_0)$.

Then differentiability of $f$ and $g$ at these points means that

$$\frac{(f(x) - L(x))}{|x|} \to 0 \text{ as } x \to 0, \quad \text{and}$$

$$\frac{(g(y) - M(y))}{|y|} \to 0 \text{ as } y \to 0.$$ 

We must show that

$$\frac{(g \circ f(x) - M \circ L(x))}{|x|} \to 0 \text{ as } x \to 0.$$
Using the differentiability of $f$ and $g$ at their origins, we have that

$$\left| gf(x) - ML(x) \right|$$

$$= \left| gf(x) - Mf(x) + Mf(x) - ML(x) \right|$$

$$\leq \left| gf(x) - Mf(x) \right| + |M|\left| f(x) - L(x) \right|$$

$$< \varepsilon |f(x)| + |M|\varepsilon |x|$$

for $|x|$ sufficiently small.

Then dividing by $|x|$, we get

$$\frac{|gf(x) - ML(x)|}{|x|} < \varepsilon \frac{|f(x)|}{|x|} + |M|\varepsilon |x|$$

We must show that this is small when $|x|$ is small, and the issue is clearly to show that $\frac{|f(x)|}{|x|}$ remains bounded.
But,

\[
\frac{|f(x)|}{|x|} \leq \frac{|L(x)|}{|x|} + \frac{|f(x) - L(x)|}{|x|},
\]

and the first term on the right is bounded by $|L|$ while the second term goes to $\to 0$ as $|x| \to 0$.

It follows that $\frac{|f(x)|}{|x|}$ remains bounded as $|x| \to 0$, and this completes the proof of the chain rule. $\square$
Partial Derivatives

Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$. Then we can write

$$f(x) = (f_1(x_1, x_2, \ldots, x_m), f_2(x_1, x_2, \ldots, x_m), \ldots, f_n(x_1, x_2, \ldots, x_m)),$$

and consider the usual partial derivatives $\frac{\partial f_i}{\partial x_j}$.

If $f$ is differentiable at $x_0$, then all of the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist at $x_0$, and the derivative $f'(x_0)$ is the linear map corresponding to the $n \times m$ matrix of partial derivatives.

The converse is false, that is, the existence of partial derivatives at a point does not imply that the function is differentiable there.
Definition
Let $L(\mathbb{R}^m, \mathbb{R}^n)$ denote the set of all linear maps of $\mathbb{R}^m$ into $\mathbb{R}^n$. This set is a vector space of dimension $mn$ whose elements can be represented by $n \times m$ matrices.

Definition
Let $U$ be an open set in $\mathbb{R}^m$ and $f : U \rightarrow \mathbb{R}^n$ a differentiable map. Since the derivative $f'(x)$ at each point $x$ of $U$ is a linear map of $\mathbb{R}^m \rightarrow \mathbb{R}^n$, we can think of $f'$ as a map $f' : U \rightarrow L(\mathbb{R}^m, \mathbb{R}^n)$. We call $f'$ the derivative of $f$.

Definition
Let $U$ be an open subset of $\mathbb{R}^m$. If $f : U \rightarrow \mathbb{R}^n$ is differentiable and $f' : U \rightarrow L(\mathbb{R}^m, \mathbb{R}^n)$ is continuous, then we say that $f$ is continuously differentiable, and write $f \in C^1$. 
Theorem

Let $U$ be an open set in $\mathbb{R}^m$ and let $f : U \rightarrow \mathbb{R}^n$. Then $f$ is continuously differentiable if and only if all of the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist and are continuous on $U$. 
**Simple Fact:** Let $f$ be a differentiable real-valued function defined on an open set $U$ in $\mathbb{R}^m$. Suppose that $f$ has a local maximum or local minimum at a point $x_0$ in $U$. Then $f'(x_0) = 0$.

**Simple Fact:** Let $U$ be a connected open set in $\mathbb{R}^m$ and $f : U \to \mathbb{R}^n$ a differentiable map such that $f'(x) = 0$ for every $x \in U$. Then $f$ is constant on $U$. 
**Theorem**

Let $U$ be an open set in $\mathbb{R}^m$ and let $f : U \rightarrow \mathbb{R}$ be a function such that all partial derivatives of orders one and two exist and are continuous on $U$. Then

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

for all $1 \leq i, j \leq m$. In other words, the order of differentiation in mixed partials is irrelevant.
Remark

If all partial derivatives of orders \( \leq n \) are continuous, then the order of differentiation in them is irrelevant.
Functions with Preassigned Partial Derivatives

Let $U$ be an open set in $\mathbb{R}^m$ and $f : U \to \mathbb{R}$ a function of class $C^2$ (remember this means that all partial derivatives of orders one and two exist and are continuous). We know that

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

for all $1 \leq i, j \leq m$.

Now we run this story in reverse, and imagine that we are seeking a function $f : U \to \mathbb{R}$ of class $C^2$, where $U$ is, for simplicity, an open set in the plane $\mathbb{R}^2$. 
We are given two functions $r$ and $s : U \to \mathbb{R}$ of class $C^1$ such that

\[
\frac{\partial f}{\partial x} = r \quad \text{and} \quad \frac{\partial f}{\partial y} = s.
\]

If $f$ exists, then

\[
\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial r}{\partial y}
\]

and

\[
\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \frac{\partial f}{\partial y} = \frac{\partial s}{\partial x}
\]

hence by equality of mixed partials, we’ll have

\[
\frac{\partial r}{\partial y} = \frac{\partial s}{\partial x}.
\]

So if we want to find $f$, we’d better make sure that

\[
\frac{\partial r}{\partial y} = \frac{\partial s}{\partial x}.
\]
But is this enough to guarantee that $f$ exists?

Surprisingly, the answer is,

"Sometimes yes and sometimes no."

We will see that it depends on the topology of the domain $U$ on which these functions are defined.

This influence of the topology of a domain on the behavior of functions defined there is a theme that will be repeated throughout the course.
Theorem

Let \( r \) and \( s : \mathbb{R}^2 \to \mathbb{R} \) be \( C^1 \) functions such that \( \frac{\partial r}{\partial y} = \frac{\partial s}{\partial x} \). Then there exists a \( C^2 \) function \( f : \mathbb{R}^2 \to \mathbb{R} \) such that \( \frac{\partial f}{\partial x} = r \) and \( \frac{\partial f}{\partial y} = s \).

Remark

If two such functions \( f_1 \) and \( f_2 \) exist, then their difference \( f_1 - f_2 \) is a constant, as an immediate consequence of the mean value theorem.

Example

Let \( U = \mathbb{R}^2 - (0,0) \). Let \( r(x, y) = \frac{-y}{x^2 + y^2} \) and \( s(x, y) = \frac{x}{x^2 + y^2} \). Then \( \frac{\partial r}{\partial y} = \frac{\partial s}{\partial x} \), yet there is no function \( f : U \to \mathbb{R} \) such that \( \frac{\partial f}{\partial x} = r \) and \( \frac{\partial f}{\partial y} = s \).
Differentiation under the integral sign

The following lemma will be used in proving the theorem.

**Lemma**

Suppose $f(x, t)$ is $C^1$ for $x \in \mathbb{R}^1$ and $t \in [0, 1]$. Define $F(x) = \int_{t=0}^{1} f(x, t) dt$.

Then $F$ is of class $C^1$ and $F'(x) = \int_{t=0}^{1} \frac{\partial f(x, t)}{\partial x} dt$.

The proof is an application of the mean value theorem.
There are various generalizations of this lemma, all proven similarly. For example, we can replace $x \in \mathbb{R}^1$ by $(x, y) \in \mathbb{R}^2$, define $F(x, y) = f(x, y, t)dt$ and conclude that

$$\frac{\partial F(x, y)}{\partial x} = \int_{t=0}^{1} \frac{\partial f(x, y, t)}{\partial x} dt.$$ 

We are ready to prove our theorem, and restate it for convenience.

**Theorem**

Let $r, s : \mathbb{R}^2 \to \mathbb{R}$ be $C^1$ functions such that $\frac{\partial r}{\partial y} = \frac{\partial s}{\partial x}$. Then there exists a $C^2$ function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\frac{\partial f}{\partial x} = r$ and $\frac{\partial f}{\partial y} = s$. 
**Proof:** First suppose we are given $f(x, y)$ with $f(0, 0) = 0$. Define $g(t) = f(tx, ty)$, and note that, by the chain rule,

$$g'(t) = \frac{\partial f}{\partial x}(tx, ty)x + \frac{\partial f}{\partial y}(tx, ty)y.$$ 

Then

$$f(x, y) = g(1) = \int_{t=0}^{1} g'(t) dt$$

$$= \int_{t=0}^{1} \frac{\partial f}{\partial x}(tx, ty)x + \frac{\partial f}{\partial y}(tx, ty)y dt.$$ 

Therefore, to find a function $f(x, y)$ such that $\frac{\partial f}{\partial x} = r$ and $\frac{\partial f}{\partial y} = s$, we should define $f$ by

$$f(x, y) = \int_{t=0}^{1} r(tx, ty)x + s(tx, ty)y dt,$$

and aim to show that $\frac{\partial f}{\partial x} = r$ and $\frac{\partial f}{\partial y} = s.$
Given

\[ f(x, y) = \int_{t=0}^{1} r(tx, ty)x + s(tx, ty)y \, dt, \]

we differentiate under the integral sign, using our lemma:

\[ \frac{\partial f(x, y)}{\partial x} = \int_{t=0}^{1} r(tx, ty) + \frac{\partial r}{\partial x}(tx, ty)tx + \frac{\partial s}{\partial x}(tx, ty)ty \, dt, \]

\[ \frac{\partial f(x, y)}{\partial x} = \int_{t=0}^{1} r(tx, ty) + \frac{\partial r}{\partial x}(tx, ty)tx + \frac{\partial r}{\partial y}(tx, ty)ty \, dt. \]

Now define \( h(t) = r(tx, ty) \) and note that

\[ h'(t) = \frac{\partial r}{\partial x}(tx, ty)x + \frac{\partial r}{\partial y}(tx, ty)y. \]
Thus,

\[
\frac{\partial f(x, y)}{\partial x} = \int_{t=0}^{1} h(t) + th'(t)\,dt
\]

\[= \int_{t=0}^{1} (th(t))'\,dt\]

\[h(1) = r(x, y).\]

Likewise, \(\frac{\partial f(x, y)}{\partial y} = s(x, y)\), and our theorem is proved. □
Critical points

Let $U$ be an open set in the plane $\mathbb{R}^2$, and let $f : U \rightarrow \mathbb{R}$ be a real valued function on $U$, all of whose first and second partial derivatives exist and are continuous on $U$. In such a case, we say that $f$ is of class $C^2$ on $U$.

We know that if $f$ has a local maximum or minimum at a point $(x_0, y_0)$ of $U$, then the first partials $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are both zero at $(x_0, y_0)$.

Searching for such points, we call $(x_0, y_0)$ a critical point of $f$ if $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are both zero at $(x_0, y_0)$, and want to learn whether $(x_0, y_0)$ is a local maximum or minimum point, a saddle point, or perhaps something more exotic.
Models:

\[ f(x, y) = -x^2 - y^2 \text{ has a local maximum at } (0, 0) \]

\[ f(x, y) = x^2 + y^2 \text{ has a local minimum at } (0, 0) \]

\[ f(x, y) = x^2 - y^2 \text{ has a saddle point at } (0, 0). \]

The issue hinges upon consideration of the Hessian matrix of second partial derivatives at the point \((x_0, y_0)\):

\[
\begin{pmatrix}
\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\
\frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2}
\end{pmatrix}
\]

We know from equality of mixed partials that this matrix is symmetric.
Differentiation

Critical Points

Theorem

Suppose that \((x_0, y_0)\) is a critical point of \(f\), and let \(H\) denote the Hessian of \(f\) at \((x_0, y_0)\).

1. If \(\det(H) > 0\) and both diagonal terms are \(> 0\), then \(f\) has a local minimum at \((x_0, y_0)\).
2. If \(\det(H) > 0\) and both diagonal terms are \(< 0\), then \(f\) has a local maximum at \((x_0, y_0)\).
3. If \(\det(H) < 0\), then \((x_0, y_0)\) is a saddle point of \(f\).
4. If \(\det(H) = 0\), the test is inconclusive.
5. If \(f\) has a local minimum or local maximum at \(f\), then \(\det(H) \leq 0\).
Definition

Let $U$ be an open set in $\mathbb{R}^2$ and $f : U \to \mathbb{R}$ a real valued function on $U$ of class $C^2$. Let $(x_0, y_0)$ be a critical point of $f$, and let $H$ be the Hessian matrix of second partials of $f$, evaluated at $(x_0, y_0)$. Then $(x_0, y_0)$ is called a nondegenerate critical point if $\det(H) \neq 0$, and a degenerate critical point if $\det(H) = 0$. 
Inverse Function Theorem

**Theorem**

(Inverse Function Theorem) Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable on an open set containing $a$, with nonsingular derivative $df_a$. Then there exists an open set $V$ containing $a$ and an open set $W$ containing $f(a)$, such that $f : V \to W$ is one-one and onto, and its inverse $f^{-1} : W \to V$ is also differentiable.

Furthermore, $d(f^{-1})f(a) = (df_a)^{-1}$.

**Example**

The mapping $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x, y) = (e^x \cos(y), e^x \sin(y))$ shows that in $\mathbb{R}^2$, unlike $\mathbb{R}^1$, the derivative of $f$ can be nonsingular at each point without $f$ being a diffeomorphism on all of $\mathbb{R}^2$. 
Proof of the Inverse Function Theorem.
Following the map $f : \mathbb{R}^n \to \mathbb{R}^n$ by the linear transformation $(df_a)^{-1}$ makes the derivative at $a$ the identity, so we assume this from the start: $df_a = I$.

Since

$$\lim_{h \to 0} \frac{|f(a + h) - f(a) - df_a(h)|}{|h|} = 0,$$

with $df_a(h) = h$, we can not have $f(a + h) = f(a)$ for nonzero $h$ arbitrarily close to 0.

Hence, there is a closed rectangle $U$ centered at $a$ with

1. $f(x) \neq f(a)$ if $x \in U$ and $x \neq a$. 

Since \( f \) is \( C^1 \) on an open set containing \( a \), we can assume

(2) \( df_x \) is nonsingular for \( x \in U \),

(3) \( \left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) \right| < \frac{1}{2n^2} \) for all \( x \in U \) and all \( i, j \).

Condition (3) will force \( f \) to be one-to-one on \( U \). To that end, we first state and prove

**Lemma**

Let \( A \) be a rectangle in \( \mathbb{R}^n \), and \( g : A \rightarrow \mathbb{R}^n \) of class \( C^1 \). Suppose that \( \left| \frac{\partial g_i}{\partial x_j} \right| \leq M \) at all points of \( A \). Then \( \left| g(x) - g(u) \right| \leq n^2 M \left| x - u \right| \) for all \( x, u \in A \).
Proof of Lemma.
Going from $u$ to $x$ by changing one coordinate at a time, and applying the MVT at each step, we get

$$|g_i(x) - g_i(u)| \leq \sum_{j=1}^{n} |x_j - u_j| M \leq nM|x - u|.$$ 

Hence,

$$|g(x) - g(u)| \leq \sum_{i=1}^{n} |g_i(x) - g_i(u)| \leq n^2 M|x - u|,$$

as claimed.
Now apply this lemma to the function $g(x) = f(x) - x$, and use

(3) $|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)| < \frac{1}{2n^2}$ for all $x \in U$ and all $i, j$,

which implies that $|\frac{\partial g_i}{\partial x_j}(x) - \frac{\partial g_i}{\partial x_j}(a)| < \frac{1}{2n^2}$.

Now $\frac{\partial g_i}{\partial x_j}(a) = 0$, and hence by the Lemma we get

$$|g(x) - g(u)| \leq n^2(\frac{1}{2n^2}|x - u| = \frac{1}{2}|x - u|.$$ 

Thus, $|(f(x) - x) - (f(u) - u)| \leq \frac{1}{2}|x - u|$.
Hence, using the triangle inequality, we get

\[ |x - u| - |f(x) - f(u)| \leq |(f(x) - x) - (f(u) - u)| \leq \frac{1}{2}|x - u|. \]

So,

(4) \[ |f(x) - f(u)| \geq \frac{1}{2}|x - u|, \]

for all \( x, u \in U \), implying that \( f \) is one-to-one on \( U \), as claimed earlier.
Now $f(\partial U)$ is a compact set which does not contain $f(a)$, since $f$ is one-to-one on $U$.

Let $d = \text{distance from } f(a) \text{ to } f(\partial U)$.

Let $W = \{y : |y - f(a)| < \frac{d}{2}\} = \text{open neighborhood of } f(a)$.

Thus, if $y \in W$ and $x \in \partial U$, we have

(5) $|y - f(a)| < |y - f(x)|$. 
CLAIM. For any \( y \in W \), there is a unique \( x \in U \) with \( f(x) = y \).

**Proof.** Fix \( y \in W \) and consider the real-valued function \( g : U \to \mathbb{R} \) defined by

\[
g(x) = |y - f(x)|^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2.
\]

Since \( g \) is continuous, it has a minimum value on \( U \).

By (5) above, this min can not occur on \( \partial U \). Say it occurs at \( x \in \text{int}(U) \).

Then \( \frac{\partial g}{\partial x_j}(x) = 0 \) for all \( j \). That is,

\[
\sum_{i=1}^{n} 2(y_i - f_i(x)) \left( \frac{\partial f_i}{\partial x_j}(x) \right) = 0
\]

for all \( j \).

But the matrix \( (\frac{\partial f_i}{\partial x_j}(x)) \) is invertible. Hence \( y_i - f_i(x) = 0 \) for all \( i \), that is, \( y = f(x) \). This \( x \) is unique, since \( f \) is one-to-one on \( U \).
Now let $V = \text{int}(U) \cap f^{-1}(W)$.

By the previous claim, $f : V \to W$ is one-to-one and onto.

Let $f^{-1} : W \to V$ be its inverse. Then we rewrite (4) as

$$(6) \ |f^{-1}(y) - f^{-1}(y')| \leq 2|y - y'|,$$

showing that $f^{-1}$ is continuous. It remains to show that $f^{-1}$ is differentiable.
Proof that $f^{-1} : W \to V$ is differentiable.

Let $x \in V$, and let $y = f(x)$. Let $L = df_x$, which we already know is nonsingular.

We will show $f^{-1}$ is differentiable at $y$ with $d(f^{-1})_y = L^{-1}$.

Write $f(x') = f(x) + L(x' - x) + \phi(x' - x)$, with $\lim_{x' \to x} \frac{|\phi(x' - x)|}{|x' - x|} = 0$.

Then $L^{-1}(f(x') - f(x)) = (x' - x) + L^{-1}\phi(x' - x)$, which we rewrite as $L^{-1}(y' - y) = f^{-1}(y') - f^{-1}(y) + L^{-1}\phi(f^{-1}(y') - f^{-1}(y))$, or $f^{-1}(y') = f^{-1}(y) + L^{-1}(y' - y) - L^{-1}\phi(f^{-1}(y') - f^{-1}(y))$. 
To show that $f^{-1}$ is differentiable at $y$ with $d(f^{-1})_y = L^{-1}$, we must show that

$$\lim_{y' \to y} \frac{|L^{-1} \phi(f^{-1}(y') - f^{-1}(y))|}{|y' - y|} = 0.$$ 

Since $L^{-1}$ is linear, it is sufficient to show that

$$\lim_{y' \to y} \frac{|\phi(f^{-1}(y') - f^{-1}(y))|}{|y' - y|} = 0.$$ 

Now write the fraction $\frac{|\phi(f^{-1}(y') - f^{-1}(y))|}{|y' - y|}$ as the product of the two fractions $\frac{|\phi(f^{-1}(y') - f^{-1}(y))|}{|(f^{-1}(y') - f^{-1}(y))|}$ and $\frac{|(f^{-1}(y') - f^{-1}(y))|}{|y' - y|}$.
We must show that the product of these two fractions goes to zero as $y' \to y$.

Since $f^{-1}$ is continuous, $y' \to y$ implies $x' = f^{-1}(y') \to x = f^{-1}(y)$. The first fraction $\frac{|\phi(f^{-1}(y') - f^{-1}(y))|}{|f^{-1}(y') - f^{-1}(y)|}$ can be rewritten as $\frac{|\phi(x' - x)|}{|x' - x|}$, and this $\to 0$ as $x' \to x$ since $f$ is differentiable at $x$.

By (6), the second fraction $\frac{|(f^{-1}(y') - f^{-1}(y))|}{|y' - y|} \leq 2$. Hence the product of the two fractions $\to 0$ as $y' \to y$, completing the proof that $f^{-1}$ is differentiable at $y$ with derivative $d(f^{-1})_y = L^{-1} = (df_x)^{-1}$, and with it the proof of the Inverse Function Theorem.
The Implicit Function Theorem

In calculus, we learn that the equation \( f(x, y) = x^2 + y^2 = 1 \) can be regarded as implicitly defining \( y \) as a function of \( x \),

\[
y = \sqrt{1 - x^2} \quad \text{or} \quad y = -\sqrt{1 - x^2}
\]

. We also learn that we can compute the derivative \( \frac{dy}{dx} \) without actually solving for \( y \). Just regard \( y \) as a function of \( x \), write \( y = y(x) \), and then the equation

\[
f(x, y(x)) = 1
\]

can be differentiated with respect to \( x \) by the chain rule.
Doing this, we get

\[ \frac{\partial f}{\partial x} + \left( \frac{\partial f}{\partial y} \right) \left( \frac{dy}{dx} \right) = 0, \]

and hence

\[ \frac{dy}{dx} = \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} = -\frac{2x}{2y}. \]

There are some subtleties: we cannot solve for \( y \) as a function of \( x \) near the points \((1,0)\) and \((-1,0)\). The implicit function theorem handles these subtleties, and we begin with the simplest case.
Theorem

(Implicit function theorem) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a $C^1$ function defined on a neighborhood of $(a, b)$, with $f(a, b) = c$. Suppose that $\frac{\partial f}{\partial y}(a, b) \neq 0$. Then there is a $C^1$ function $g : \mathbb{R} \to \mathbb{R}$ defined on a neighborhood of $a$ such that $g(a) = b$ and such that $f(x, g(x)) = c$ for all $x$ in that neighborhood.

Proof. Define a $C^1$ function $F : \mathbb{R}^2 \to \mathbb{R}^2$ on the given neighborhood of $(a, b)$ by $F(x, y) = (x, f(x, y))$. The derivative $F'(a, b)$ is nonsingular because it is represented by a $2 \times 2$ matrix with determinant $\frac{\partial f}{\partial y}(a, b)$. Hence, by the Inverse Function Theorem, $F$ is a $C^1$ function with $C^1$ inverse from a neighborhood $U$ of $(a, b)$ to a neighborhood $V$ of $F(a, b) = (a, c)$. 
Let $H : V \to U$ be the inverse $C^1$ map. Since $F(x, y) = (x, f(x, y))$, we have $H(x, z) = (x, h(x, z))$. If we define $g(x) = h(x, c)$ on a neighborhood of $x$, then

$$F(x, g(x)) = F(x, h(x, c)) = FH(x, c) = (x, c)$$

so

$$f(x, g(x)) = c,$$

as desired. □
The general case is no more difficult to prove, and we style the notation so that its statement looks almost the same as the statement of its prototype above:

\[ x = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m \]
\[ y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n \]
\[ z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n, \]

and the \((i, j)\) entry of the \(n \times n\) matrix \(\frac{\partial f}{\partial y}(a, b)\) is the partial derivative \(\frac{\partial f_i}{\partial y_j}(a, b)\).
Theorem (Implicit function theorem (general case))

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ be a $C^1$ function defined on a neighborhood of $(a, b)$, with $f(a, b) = c$. Suppose that the $n \times n$ matrix $\frac{\partial f}{\partial y}(a, b)$ is nonsingular. Then there is a $C^1$ function $g : \mathbb{R}^m \to \mathbb{R}^n$ defined on a neighborhood of $a$ such that $g(a) = b$ and such that $f(x, g(x)) = c$ for all $x$ in that neighborhood.