[11pt]article graphicx amssymb epstopdf

Theorem [theorem]Corollary Definition

1. Implicit Functions

1.1. **Examples of Implicit Functions.** A function $f : D \to \mathbb{R}^m$ is usually defined by giving some explicit formula to calculate $f(x) \in \mathbb{R}^m$ for each $x \in D \subset \mathbb{R}^n$. Functions can also be defined *implicitly* by a system of equations

$$F(x, y) = c$$

where $F : D_1 \times D_2 \to \mathbb{R}^m$ is defined on some domain $D_1 \times D_2 \subset \mathbb{R}^n \times \mathbb{R}^m$. Given $x \in D_1 \subset \mathbb{R}^n$ we "solve" the system of equations for $y = y(x) \in D_2 \subset \mathbb{R}^m$ and in this way obtain a function $y : D_1 \to D_2$. In this section we shall examine conditions under which such an implicit function exists and is unique. Let us start by looking at some examples.

1.1.1. *Inverse Functions.* The *inverse* of a function f is the function defined implicitly as the solution of the equation $F(x, y) = x - f(y) = 0$. Solving for y gives the inverse function $y = f^{-1}(x)$. We know that we can only expect a well-defined inverse function to exist on an interval where f is one-to-one. If f is differentiable, such intervals can be found by checking where $f'(x) > 0$ (or $f'(x) < 0$). For example, let $f(x) = x^{1/x}$ for $x \geq 0$ ($f(0) = \lim_{x\to 0} x^{1/x} = \lim_{x\to 0} e^{\log(x)/x} = 0$). Then $f'(x) = x^{1/x}(1 - \log(x))/x^2 > 0$ for $0 < x < e$ and $f'(x) < 0$ for $x > e$, so f has an inverse on either of the intervals $[0, e]$ or $[e, \infty)$. Finding a "formula" for the inverse by solving $x - y^{1/y} = 0$ for y is difficult, even though we know $y = y(x)$ exists as an abstract function of x. We could argue that $y = x^y$ so substituting this equation into itself yields $y = x^{x^y}$. Repeating this indefinitely, we might conclude that

$$y = x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}$$

The difficulty of solving equations explicitly underscores the importance of having criteria that guarantee the existence and uniqueness of a solution.

1.1.2. *Solutions to Exact Differential Equations.* Recall that an *exact* differential equation is one of the form

$$\frac{dy}{dx} = -\frac{M(x,y)}{N(x,y)} \tag{1}$$

where the functions M and N satisfy

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \tag{2}$$

Equation (1) is often written

$$M(x,y)\,dx + N(x,y)\,dy = 0 \tag{3}$$

To solve the differential equation we find a function $F(x,y)$ such that

$$\frac{\partial F}{\partial x} = M \text{ and } \frac{\partial F}{\partial y} = N \tag{4}$$

The conditions (2) guarantee that such an F exists, at least in some neighborhood of a given point in the xy-plane. The conditions are certainly necessary, since if (4) holds then

$$\frac{\partial M}{\partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial N}{\partial x}$$

Given the function $F(x,y)$, (3) can be written

$$dF(x,y) = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy = M(x,y)\,dx + N(x,y)\,dy = 0$$

and the solution of the differential equation is therefore the implicit solution $y = y(x)$ of the equation $F(x,y) = c$ for some constant c.

Let us work out an example. The differential equation

$$(e^x \cos(y) - 2x)\,dx + (1 - e^x \sin(y))\,dy = 0$$

is exact since

$$\frac{\partial}{\partial y}(e^x \cos(y) - 2x) = -e^x \sin(y) = \frac{\partial}{\partial x}(1 - e^x \sin(y))$$

Integrating $M(x,y) = e^x \cos(y) - 2x$ with respect to x gives the first information about $F(x,y)$,

$$F(x,y) = e^x \cos(y) - x^2 + g(y)$$

where $g(y)$ is an unknown function of y (so $\partial g(y)/\partial x = 0$). To determine $g(y)$, compute $F_y(x,y)$ and set it equal to $N(x,y) = 1 - e^x \sin(y)$,

$$F_y(x,y) = -e^x \sin(y) + g'(y) = 1 - e^x \sin(y)$$

which implies that $g'(y) = 1$ and hence $g(y) = y$. The solution of the differential equation is therefore the function $y = y(x)$ defined implicitly by the equation

$$F(x,y) = e^x \cos(y) - x^2 + y = c$$

Below are some graphs of the implicit solutions (determined numerically) for various values of the constant c.

1.1.3. *Equations of Curves and Surfaces.* We often describe a curve in the plane or a surface in space by an equation, $F(x, y) = c$ or $F(x, y, z) = c$, respectively. For example, the unit sphere is defined as the set of points satisfying $x^2 + y^2 + z^2 = 1$. The sphere is "two-dimensional" because there are two "degrees of freedom" on the surface in the sense that any one variables can be thought of as a function of the other two. For example, $z = \pm\sqrt{1 - x^2 - y^2}$. This functional representation is less elegant than the single equation and it also has exceptions and cases (we must choose the positive or negative square root, and the representation does not work well at the points $x^2 + y^2 = 1$). It is often difficult if not impossible to solve explicitly for one variable as a function of the other two in the equation for a general surface. For example,

$$F(x, y, z) = 8(x^2 + y^2 + z^2) - 8(x^4 + y^4 + z^4) = c$$

Here are the surfaces corresponding to $c = 2$, 3, and 4:

The hand-drawn pictures were done as a homework assignment by a freshman, Cassidy Curtis, in 1988 at Brown University without the aid of a computer. For more on this story check out the link "The Best Homework Ever?" at `http://www.math.brown.edu/~banchoff/`. It is difficult to see why $c = 2$ gives a rounded cube with "dimpled" faces, $c = 3$ gives a surface with six "holes" in it, and $c = 4$ gives a surface with 12 "singular points." It would be useful to have some way of understanding the surface analytically through its equation.

1.1.4. *Systems of Equations.* Implicit functions can also be vector-valued and defined by systems of equations. For example, consider $F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F(x, y) = (x_1 - y_1^2 + y_2^2, x_2 - 2y_1y_2), \quad x = (x_1, x_2),\ y = (y_1, y_2)$$

The system of equations $F(x, y) = (c_1, c_2)$ is equivalent to

$$\begin{aligned} x_1 &= c_1 + y_1^2 - y_2^2 \\ x_2 &= c_2 + 2y_1y_2 \end{aligned}$$

We can solve this system algebraically (substitute $y_2 = (x_2 - c_2)/(2y_1)$ into the first equation) to realize y as a function of x.

$$y_1 = \pm\sqrt{\frac{\sqrt{(x_1-c_1)^2+(x_2-c_2)^2}+(x_1-c_1)}{2}}$$

$$y_2 = \pm\sqrt{\frac{\sqrt{(x_1-c_1)^2+(x_2-c_2)^2}-(x_1-c_1)}{2}}$$

The Implicit Function Theorem is a tool for understanding implicitly defined functions. It provides answers to questions raised by the above examples such as, When does an inverse function exist? When are the solutions of an exact differential equation smooth curves? When do systems of equations define smooth curves and surfaces?

1.2. **Implicit Function Theorem.** One way to find an approximate solution to a system of equations $F(x, y) = c$ is to "linearize" the system. The linear transformation, let's call it $L(x, y)$ that best approximates $F(x, y) = (F_1(x, y), \ldots, F_m(x, y))$ near (a, b) is

$$L(x,y) = \left(F_i(a,b) + \sum_{j=1}^{n} \frac{\partial F_i}{\partial x_j}(a,b)(x_j - a_j) + \sum_{j=1}^{m} \frac{\partial F_i}{\partial y_j}(a,b)(y_j - b_j)\right)_{1\le i\le m}$$

We can simplify this expression by using matrix operations,

$$L(x, y) = F(a, b) + F_x(a, b)(x - a) + F_y(a, b)(y - b)$$

where $F_x(a, b)$ is the rectangular $m \times n$ matrix

$$F_x(a,b) = \left[\frac{\partial F_i}{\partial x_j}(a,b)\right]_{\substack{1\le i\le m\\ 1\le j\le n}}$$

and $F_y(a, b)$ is the square $m \times m$ matrix

$$F_y(a,b) = \left[\frac{\partial F_i}{\partial y_j}(a,b)\right]_{\substack{1\le i\le m\\ 1\le j\le m}}$$

The linearized system of equations becomes $L(x, y) = c$ or

$$F(a, b) + F_x(a, b)(x - a) + F_y(a, b)(y - b) = c$$

If we fix $x = a$, the corresponding solution for y is

$$y = b + F_y(a, b)^{-1}(c - F(a, b)) \tag{5}$$

Note that we must assume $F_y(a, b)$ is invertible to solve for y. The solution to the linearized system suggests a method of obtaining the general solution using the Contractive Mapping Principle.

Before we proceed we need to have the notion of a norm of a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$. We may regard L as a rectangular

$n \times m$ matrix and the transformation as given by matrix multiplication, $L \cdot v \in \mathbb{R}^n$ for $v \in \mathbb{R}^m$. We define

$$\|L\| = \sup_{v \neq 0} \frac{|L \cdot v|}{|v|} = \sup_{v \neq 0} \left| L \cdot \frac{v}{|v|} \right| = \sup_{|u|=1} |L \cdot u|$$

The supremum is taken over the compact unit sphere, $|u| = 1$, $u \in \mathbb{R}^n$, so it is a finite number. The norm allows us to estimate

$$|L \cdot v| \leq \|L\| \cdot |v|, \quad \forall v \in \mathbb{R}^n \tag{6}$$

We leave it as an exercise for the reader to verify that $\|L\|$ is indeed a norm on the nm-dimensional vector space of $n \times m$ matrices and that this norm fits between the sup-norm and the usual Euclidean norm,

$$\|L\|_\infty \leq \|L\| \leq |L|$$

(The nm entries of a matrix $L = [L_{ij}]$ are the "components" of L as a vector so that $\|L\|_\infty = \max |L_{ij}|$ and $|L| = (\sum L_{ij}^2)^{1/2}$.)

Theorem 1.1 (Implicit Function Theorem)**.** *Let $F : D_1 \times D_2 \to \mathbb{R}^m$ be a C^1 function defined on a neighborhood of $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m$, and let $c = F(x_0, y_0)$. If $F_y(x_0, y_0)$ is invertible, then there is a neighborhood U of x_0 and a C^1 function $y(x) : U \to D_2$ such that*

$$F(x, y(x)) = c, \quad \forall\, x \in U$$

Furthermore, the function $y(x)$ is unique in that there is a neighborhood V of b such that the only solution of $F(x, z) = c$ for $z \in V$ is $z = y(x)$. Finally, the differential of $y(x)$ is given by implicit differentiation as

$$dy(x) = -F_y(x, y(x))^{-1} F_x(x, y(x))$$

Proof. For $x \in D_1$ define $T : D_2 \to \mathbb{R}^m$ by

$$Ty = y + F_y(x_0, y_0)^{-1}[c - F(x, y)]$$

(compare with equation (5)). Although T depends on x, we will not complicate the notation by indicating this dependence.

The first step is to show that T is a contractive mapping when restricted to a suitable neighborhood of y_0. Before we begin we must set up some notation and a few constants. Let $L(x, y)$ denote the linear approximation to $F(x, y)$ near (x_0, y_0),

$$L(x, y) = F(x_0, y_0) + F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0)$$

The fact that $L(x, y)$ approximates $F(x, y)$ near (x_0, y_0) is expressed formally in a version of Taylor's Theorem for vector-valued functions which says that for any $\lambda > 0$, there is a neighborhood $B_0 = B(x_0, \epsilon_0) \times B(y_0, \delta_0)$ such that

$$|L(x, y) - F(x, y)| < \lambda(|x - x_0| + |y - y_0|), \quad \forall\, (x, y) \in B_0 \tag{7}$$

Similarly, our assumption that F is C^1 implies that $F_y(x,y)$ is continuous in x and y, so for any $\lambda > 0$ there is a neighborhood $B_1 = B(x_0, \epsilon_1) \times B(y_0, \delta_1)$ such that

$$\|F_y(x,y) - F_y(x_0,y_0)\| < \lambda, \quad \forall\,(x,y) \in B_1$$

Let

$$M = \|F_y(x_0,y_0)^{-1}\|, \quad N = \|F_x(x_0,y_0)\|$$

and let B_0, B_1 be the neighborhoods given above when $\lambda = 1/(4M)$. Let

$$\begin{aligned} \delta &= \min\{\delta_0, \delta_1\} \\ \epsilon &= \min\left\{\epsilon_0, \epsilon_1, \frac{\delta}{4MN+1}\right\} \end{aligned}$$

and define $B = B(x_0, \epsilon) \times B(y_0, \delta)$. Then $\forall\,(x,y) \in B$

$$|L(x,y) - F(x,y)| < \frac{1}{4M}(|x - x_0| + |y - y_0|) < \frac{1}{4M}(\epsilon + \delta) \tag{8}$$

$$\|F_y(x,y) - F_y(x_0,y_0)\| < \frac{1}{4M} \tag{9}$$

We are now ready to prove that T is a contractive mapping. For $y, z \in B(y_0, \delta)$,

$$\begin{aligned} Tz - Ty &= z - y + F_y(x_0,y_0)^{-1}[F(x,y) - F(x,z)] \\ &= F_y(x_0,y_0)^{-1}[F(x,y) - F(x,z) + F_y(x_0,y_0)(z - y)] \end{aligned}$$

We now need to write $F(x,y) - F(x,z)$ in terms of $(z - y)$. In one dimension we could use the Mean Value Theorem, but in higher dimensions we must substitute an argument using the Fundamental Theorem of Calculus on the line segment $z + t(y - z)$, $0 \le t \le 1$, joining z and y. Since

$$\frac{d}{dt}F(x, z + t(y - z)) = F_y(x, z + t(y - z))(y - z)$$

we get

$$\begin{aligned} &\int_0^1 F_y(x, z + t(y - z))(y - z)\,dt \\ &= F(x, z + t(y - z))\Big|_0^1 = F(x,y) - F(x,z) \end{aligned}$$

Inserting this into the equation above and using the fact that

$$F_y(x_0,y_0)(y - z) = \int_0^1 F_y(x_0,y_0)(y - z)\,dt$$

we obtain

$$Tz - Ty = F_y(x_0,y_0)^{-1}\int_0^1 [F_y(x, z + t(y - z)) - F_y(x_0,y_0)](y - z)\,dt$$

Therefore, applying Minkowski's Inequality (Lemma **??**) and inequality (6),

$$\begin{aligned}
&|Tz - Ty| \\
&\leq \|F_y(x_0, y_0)^{-1}\| \cdot \Big| \int_0^1 [F_y(x, z + t(y - z)) - F_y(x_0, y_0)](y - z)\, dt \Big| \\
&\leq M \int_0^1 \|F_y(x, z + t(y - z)) - F_y(x_0, y_0)\| \cdot |y - z|\, dt
\end{aligned}$$

Now, $(x, y), (x, z) \in B$ implies $(x, z + t(y - z)) \in B$, so by the inequality (9),

$$\|F_y(x, z + t(y - z)) - F_y(x_0, y_0)\| < \frac{1}{4M}$$

Thus,

$$|Tz - Ty| \leq M \int_0^1 \frac{1}{4M} |y - z|\, dt = \frac{1}{4}|z - y|$$

proving that T is contractive on $B(y_0, \delta)$.

We must still show that T maps some compact neighborhood of y_0 into itself. Let $V = B(y_0, \delta/2)$ so that $\overline{V}$ is compact and $\overline{V} \subset B(y_0, \delta)$. To prove that $T : \overline{V} \to \overline{V}$ we must show that

$$|y - y_0| \leq \delta/2 \Rightarrow |Ty - y_0| \leq \delta/2$$

We first write

$$\begin{aligned}
Ty - y_0 &= y + F_y(x_0, y_0)^{-1}[F(x_0, y_0) - F(x, y)] - y_0 \\
&= F_y(x_0, y_0)^{-1}[F(x_0, y_0) - F(x, y) + F_y(x_0, y_0)(y - y_0)] \\
&= F_y(x_0, y_0)^{-1}[L(x, y) - F(x, y) - F_x(x_0, y_0)(x - x_0)]
\end{aligned}$$

where $L(x, y)$ is the linear approximation of $F(x, y)$ given above. Then, if $x \in U = B(x_0, \epsilon)$ and $y \in \overline{V}$, we apply inequality (8) to get

$$\begin{aligned}
|Ty - y_0| &\leq \|F_y(x_0, y_0)^{-1}\| \cdot |L(x, y) - F(x, y) - F_x(x_0, y_0)(x - x_0)| \\
&\leq M(|L(x, y) - F(x, y)| + \|F_x(x_0, y_0)\| \cdot |x - x_0|) \\
&< M\Big(\frac{1}{4M}(\epsilon + \delta) + N\epsilon\Big) \\
&= \frac{\delta}{4} + \Big(MN + \frac{1}{4}\Big)\epsilon \leq \frac{\delta}{4} + \frac{\delta}{4} = \frac{\delta}{2}
\end{aligned}$$

since $\epsilon \leq \delta/(4MN + 1)$.

We have now shown that for any $x \in U$, the mapping $T : \overline{V} \to \overline{V}$ is contractive. Therefore, by the Contractive Mapping Principle (Theorem **??**), there exists a *unique* fixed point $y(x) \in \overline{V}$,

$$Ty(x) = y(x)$$

and this implies that

$$y(x) = y(x) + F_y(x_0, y_0)^{-1}[c - F(x, y(x))]$$

or equivalently

$$F(x, y(x)) = c$$

Finally, we prove that $y(x)$ is C^1 and derive the formula for $dy(x)$. Since $F(x_0, y_0) = c = F(x, y(x))$ we find that

$$L(x, y(x)) - F(x, y(x)) = F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y(x) - y_0)$$

Solving for $y(x)$ gives

$$y(x) = y_0 - F_y(x_0, y_0)^{-1} F_x(x_0, y_0)(x - x_0) \tag{10}$$

$$-F_y(x_0, y_0)^{-1}[L(x, y(x)) - F(x, y(x))] \tag{11}$$

The proof will be complete once we show that the last term (11) is $o(|x - x_0|)$ for then, by definition,

$$dy(x_0) = -F_y(x_0, y_0)^{-1} F_x(x_0, y_0)$$

More precisely, we must show that for any $\lambda > 0$, there is a neighborhood $B(x_0, \mu)$ such that

$$\left| F_y(x_0, y_0)^{-1}[L(x, y(x)) - F(x, y(x))] \right| < \lambda |x - x_0|, \quad \forall\, x \in B(x_0, \mu)$$

Given $\lambda > 0$, choose $\lambda_0 > 0$ such that

$$\lambda_0 < \frac{\lambda}{M(1 + MN + \lambda)} \tag{12}$$

Using inequality (7), there is a neighborhood

$$D = B(x_0, \mu) \times B(y_0, \nu) \subset B$$

such that

$$|L(x, y) - F(x, y)| < \lambda_0(|x - x_0| + |y - y_0|), \quad \forall\, (x, y) \in D$$

By shrinking ϵ and δ in the first part of the proof, if necessary, we may assume $\delta/2 < \nu$ so that $x \in B(x_0, \mu) \Rightarrow y(x) \in B(y_0, \nu)$. Then, from (10),

$$\begin{aligned} |y(x) - y_0| &\leq \|F_y(x_0, y_0)^{-1}\| \cdot \|F_x(x_0, y_0)\| \cdot |x - x_0| \\ &\quad + \|F_y(x_0, y_0)^{-1}\| \cdot |L(x, y(x)) - F(x, y(x))| \\ &\leq MN|x - x_0| + M\lambda_0(|x - x_0| + |y(x) - y_0|) \end{aligned}$$

Solving for $|y(x) - y_0|$ gives the Lipschitz Condition

$$|y(x) - y_0| \leq \frac{M(N + \lambda_0)}{1 - \lambda_0 M} |x - x_0|$$

(Note that, by construction, $M\lambda_0 < 1$.) Therefore, $\forall\, x \in B(x_0, \mu)$,

$$\begin{aligned}
&\left|F_y(x_0, y_0)^{-1}[L(x, y(x)) - F(x, y(x))]\right| \\
&\quad\le \|F_y(x_0, y_0)^{-1}\| \cdot |L(x, y(x)) - F(x, y(x))| \\
&\quad< M\lambda_0(|x - x_0| + |y(x) - y_0|) \\
&\quad\le M\lambda_0\Big(|x - x_0| + \frac{M(N + \lambda_0)}{1 - \lambda_0 M}|x - x_0|\Big) \\
&\quad= M\lambda_0 \frac{1 + MN}{1 - M\lambda_0}|x - x_0| \\
&\quad< \lambda|x - x_0|
\end{aligned}$$

since, by (12),

$$M\lambda_0 \frac{1 + MN}{1 - M\lambda_0} < \lambda \iff \lambda_0 < \frac{\lambda}{M(1 + MN + \lambda)}$$

One last remark we need to make is that the formula for the derivative holds for all x in a neighborhood of x_0, not just at the point x_0. This follows from the assumption that $F_y(x, y)$ is continuous in x and y. Since the determinant is also a continuous function of matrix entries, and since $\det F_y(x_0, y_0) \neq 0$, we know that there is a neighborhood of (x_0, y_0) for which $\det F_y(x, y) \neq 0$ (and hence $F_y(x, y)$ is invertible) in that neighborhood. Therefore, the preceding proof can be done at any point in this neighborhood. □

Examples:

Let us return to some of the examples at the beginning of this section to see how the Implicit Functions Theorem applies to them.

1) We found the solutions to a certain exact differential equation to be given by the equations

$$F(x, y) = y + e^x \cos(y) - x^2 = c$$

for various values of c. By the Implicit Function Theorem, such an equation defines y as a function of x near any point where

$$F_y(x, y) = 1 - e^x \sin(y) \neq 0$$

The points in the xy-plane where $1 - e^x \sin(y) = 0$ are show below.

For any point not on one of these curves, y is, at least locally, a function of x. The derivative of this function is given by the Implicit Function Theorem as

$$y'(x) = -F_y(x, y)^{-1} F_x(x, y) = -\frac{e^x \cos(y) - 2x}{1 - e^x \sin(y)}$$

which is essentially the differential equation we started with.

2) We found that we could solve the system of equations

$$F(x,y) = (x_1 - y_1^2 + y_2^2, x_2 - 2y_1y_2) = (c_1, c_2)$$

algebraically for $y = (y_1, y_2)$ as a function of $x = (x_1, x_2)$. The Implicit Function Theorem guarantees that y is a C^1 function of x near any point for which $\det F_y(x,y) \neq 0$. Since

$$\det F_y(x,y) = \det \begin{bmatrix} -2y_1 & 2y_2 \\ -2y_2 & -2y_1 \end{bmatrix} = 4(y_1^2 + y_2^2)$$

we need only avoid $y = (0,0)$ and $x = (c_1, c_2)$. The differential of y is given by the Implicit Function Theorem as

$$\begin{aligned} dy(x) &= -F_y(x,y)^{-1} F_x(x,y) = -\begin{bmatrix} -2y_1 & 2y_2 \\ -2y_2 & -2y_1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \frac{1}{2(y_1^2 + y_2^2)} \begin{bmatrix} y_1 & y_2 \\ -y_2 & y_1 \end{bmatrix} \end{aligned}$$

1.3. **Inverse Function Theorem.** An immediate consequence of the Implicit Function Theorem is the following.

Theorem 1.2 (Inverse Function Theorem)**.** *Let $f : D \to \mathbb{R}^n$ be a C^1 function defined on some neighborhood $D \subset \mathbb{R}^n$ of y_0. If $df(y_0)$ is invertible, then there are neighborhoods U of $x_0 = f(y_0)$ and $V \subset D$ of y_0, and a C^1 inverse function $g : U \to V$ such that*

$$f(g(x)) = x, \quad \forall\, x \in U$$

and

$$g(f(y)) = y, \quad \forall\, y \in V$$

Moreover, if $x = f(y)$ then

$$dg(x) = df(y)^{-1}$$

Proof. Let $F(x,y) = f(y) - x$. Then $F_y(x_0, y_0) = df(y_0)$ is invertible by assumption, so the Implicit Function Theorem implies that there are neighborhoods U of x_0 and V of y_0 and a unique C^1 function $g : U \to V$ such that $F(x, g(x)) = F(x_0, y_0) = 0$, $\forall\, x \in U$. But this means $f(g(x)) = x$, $\forall\, x \in U$. Now restrict f to V and observe that if $y \in V$ and $f(y) = x \in U$, then by uniqueness $y = g(x)$. Therefore, $g(U)$ equals the open set $f^{-1}(U)$ and we may replace the neighborhood V with $g(U)$, if necessary. Thus, if $y \in V = g(U)$, then $y = g(x)$ for some $x \in U$, so $g(f(y)) = g(f(g(x))) = g(x) = y$. Finally, the Implicit Function Theorem gives the formula $dg(x) = -F_y(x,y)^{-1} F_x(x,y) = -df(y)^{-1} \cdot (-I) = df(y)^{-1}$. □

Example: Consider the function $f : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$f(y) = (y_1 + y_2^2, y_2 + y_1^2, y_3 - y_1y_2)$$

The differential of f is

$$df(y) = \begin{bmatrix} 1 & 2y_2 & 0 \\ 2y_1 & 1 & 0 \\ -y_2 & -y_1 & 1 \end{bmatrix}$$

The determinant of this matrix is $\det df(y) = 1 - 4y_1y_2$. By the Inverse Function Theorem, the function f has an inverse near any point in $\mathbb{R}^3$ not on the sheet $4y_1y_2 = 1$.

1.4. **Implicit Description of Surfaces.** We are familiar with defining curves and surfaces by equations. For example, the unit circle is $x^2+y^2 = 1$ and the unit sphere is $x^2+y^2+z^2 = 1$. Systems of equations also produce interesting geometric objects. For example, the intersection of the unit sphere with the off-center cylinder $(x-1/2)^2+y^2 = 1/4$ produces a "figure-eight" curve on the sphere.

This curve is the solution of the system of two equations in three unknowns,

$$F(x, y, z) = (x^2 + y^2 + z^2, (x - 1/2)^2 + y^2) = (1, 1/4)$$

In general, a system of equations, $F(x) = c$, given by a function $F : \mathbb{R}^n \to \mathbb{R}^{n-m}$, usually produces an m-dimensional surface, called a *level set* of the function F. The Implicit Function Theorem can be used to describe when such a level set is a smooth m-dimensional surface. One of the simplest ways to represent a smooth surface S is as the graph of a C^1 function $f : U \to \mathbb{R}^{n-m}$ defined on some subset $U \subset \mathbb{R}^m$,

$$S = \{(t, f(t)) \in \mathbb{R}^m \times \mathbb{R}^{n-m} \mid t \in U\}$$

As an example, the figure-eight curve C given above can be represented locally as the graph of the function $f : [0, 1] \to \mathbb{R}^2$

$$f(x) = (\pm\sqrt{x - x^2}, \pm\sqrt{1 - x})$$
$$C = \{(x, f(x)) \mid 0 \le x \le 1\}$$

Theorem 1.3. *Let $F : \mathbb{R}^n \to \mathbb{R}^{n-m}$ be a C^1 function and suppose $dF(x)$ has rank $n - m$ at a point x_0 in the level set*

$$S = \{x \in \mathbb{R}^n \mid F(x) = c\}$$

Then S can be represented as the graph of a differentiable function in some neighborhood of x_0. More precisely, there is a neighborhood $V \subset \mathbb{R}^n$ of x_0, an open set $U \subset \mathbb{R}^m$, and a function $f : U \to \mathbb{R}^{n-m}$ such that

$$S \cap V = \{(t, f(t)) \mid t \in U\}$$

Proof. Since $dF(x_0)$ has rank $n-m$ we can find $n-m$ variables x_j such that the columns $\partial F/\partial x_j(x_0)$ are linearly independent. Call these the s variables and the remaining variables the t variables. For simplicity, we assume the first m variables $x_1, \ldots, x_m$ are the t variables and we write the level set as $F(t, s) = c$. The independence of the columns $\partial F/\partial s_j$ is equivalent to the matrix F_s being invertible. Therefore, the Implicit Function Theorem implies there is a C^1 function $f : U \to \mathbb{R}^{n-m}$ such that $F(t, f(t)) = c$ for $t \in U$ showing that the level set can be locally represented by the graph of a differentiable function. □

Examples:

1) Consider the figure-eight curve C given above as a level set of $F(x, y, z) = (x^2 + y^2 + z^2, (x - 1/2)^2 + y^2)$. The differential

$$dF(x, y, z) = \begin{bmatrix} 2x & 2y & 2z \\ 2x - 1 & 2y & 0 \end{bmatrix}$$

has rank 2 except on the x-axis ($y = z = 0$). So, away from the point $(1, 0, 0)$ the curve C can be locally represented as the graph of a C^1 function. The picture shows that C crosses itself at the point $(1, 0, 0)$ and thus cannot be represented as the graph of a function there.

2) We have seen that the level sets of the function $F(x, y, z) = 8(x^2 + y^2 + z^2) - 8(x^4 + y^4 + z^4)$ can take on different shapes, some with singular points. By the previous theorem, we can discover which level surfaces have singularities by finding where the rank of the differential

$$dF(x, y, z) = (16x - 32x^3, 16y - 32y^3, 16z - 32z^3)$$

is less than 1. In fact the rank clearly equals 1 unless all components are zero:

$$x(2x^2 - 1) = y(2y^2 - 1) = z(2z^2 - 1) = 0$$

Therefore, we can expect singularities in the level sets that contain points all of whose coordinates are one of the values 0 or $\pm 1/\sqrt{2}$. The corresponding values of $c = F(x, y, z)$ are easily found. Let us break them down by type.

Case a) $(x, y, z) = 0$, $c = 0$. The origin is an isolated point of the level set $F(x, y, z) = 0$, since any nearby point (x, y, z), $0 < |x|, |y|, |z| < 1$, satisfies $x^4 + y^4 + z^4 < x^2 + y^2 + z^2$ and so $F(x, y, z) \neq 0$. The other points on this level set form a smooth surface. The origin is thus a singular point of the level set. The level sets are smooth surfaces for $c < 0$ and for $0 < c < 2$. Although it is not obvious, the level sets for $0 < c < 2$ have two components: an outer surface that looks like a rounded cube, and an inner surface that grows out of the origin and looks like a rounded octagon with the corners pointing towards the faces of the outer rounded cube.

Case b)

$$(x,y,z) = \big(\frac{\pm 1}{\sqrt{2}},0,0\big),\big(0,\frac{\pm 1}{\sqrt{2}},0\big),\big(0,0,\frac{\pm 1}{\sqrt{2}}\big)$$
$$c = 8\big(\frac{1}{2}\big) - 8\big(\frac{1}{4}\big) = 2$$

These 6 points lie on the level set $F(x,y,z) = 2$. The level set still has an outer surface like a rounded cube and with "dimples" on the 6 faces corresponding the 6 listed points (see the picture on p.66). The inner surface is now large enough so that its corners just touch the faces of the outer surface at the dimples. For $2 < c < 4$ the dimples break through the surface, connecting the previous outer and inner surfaces, and creating a smooth surface with 6 holes.

Case c)

$$(x,y,z) = \big(\frac{\pm 1}{\sqrt{2}},\frac{\pm 1}{\sqrt{2}},0\big),\big(\frac{\pm 1}{\sqrt{2}},0,\frac{\pm 1}{\sqrt{2}}\big),\big(0,\frac{\pm 1}{\sqrt{2}},\frac{\pm 1}{\sqrt{2}}\big)$$
$$c = 8\big(\frac{1}{2}+\frac{1}{2}\big) - 8\big(\frac{1}{4}+\frac{1}{4}\big) = 4$$

These 12 points lie on the level set $F(x,y,z) = 4$. The holes have now gotten so big that the the corners of the rounded cube are at the point of breaking off, thus creating singular points at the 12 edges of the rounded cube corresponding to the 12 listed points (see the picture on p.66). For $4 < c < 6$, the level sets break into 8 separate round surfaces that shrink down to points as c approaches 6.

Case d)

$$(x,y,z) = \big(\frac{\pm 1}{\sqrt{2}},\frac{\pm 1}{\sqrt{2}},\frac{\pm 1}{\sqrt{2}}\big)$$
$$c = 8\big(\frac{1}{2}+\frac{1}{2}+\frac{1}{2}\big) - 8\big(\frac{1}{4}+\frac{1}{4}+\frac{1}{4}\big) = 6$$

These 8 points actually comprise the level set $F(x,y) = 6$, since 6 is the maximum value of $F(x,y,z)$. (To see this just observe that $8x^2 - 8x^4$ has global maximum of 2 at $x = \pm 1/\sqrt{2}$ and $F(x,y,z)$ is the sum of three such functions). Thus, the level set is not a surface at all, but a set of 8 singular points. These points are the limit points of the level sets in the previous case. The level sets for $c > 6$ are empty.

1.5. **Exercises.**

(1) Find the implicit solutions to the differential equation

$$(y\cos(xy) - 1)dx + (x\cos(xy) + 1)dy = 0$$

Determine the points where y can be locally expressed as a C^1 function of x and where x can be locally expressed as a C^1 function of y.

(2) Find the points where the function $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y) = (\sin(x)\cosh(y), \cos(x)\sinh(y))$$

has a local inverse.

(3) Sketch the intersection of the paraboloid $z = 4 - x^2 - y^2$ and the cylinder $y^2 + (z - 2)^2 = 4$. Determine analytically where the intersection curve can be locally represented as the graph of a C^1 function.