Corrections to IFT and S3.5 as in email. *Contents:* A few more clarifications and corrections for Chapter 3, some remarks on advanced calculus we need before starting Chapter 5, and more supplemental problems at the end. (We will skip Chapter 4 for now: many of its techniques and results are closely related to Chapter 8, so more easily done then.)

But first, comment from the grader. "It would be nice if the students don't submit their draft [of their homework]. They should work [the draft] on scratch paper [or a board, or whatever] first, then organize it and write a good solution. It is hard to read if there are arrows everywhere and information spread all over the pace. In Math we are persuading others with arguments, so the arguments should be nicely written. Also, they should staple their work. There are staplers in the library, at the front desk or by printing." In the future, he may deduct points for poor organization and exposition, even if the solution is correct (or seems to be, as well as he can make it out).

Additional Chapter 3 Corrections.

Last sentence, p. 81 should say "contact of at least order 3" (i.e., add "at least").

Proposition 3.3.3 Proof, last line on p. 82, delete "sign(f')", so it says $\tau_\xi = \tau$.

Exercise 3.3.3, p. 83, change "osculating circle" to "osculating plane."

Theorem 3.4.2, p. 84, should require $\kappa(s) > 0$; that is, the curvature cannot be zero anywhere (because if it were, the torsion τ would be undefined).

Some advanced calculus

The authors of our text want the book to be accessible to a wide range of students, so do not expect their readers to have any background in analysis beyond calculus. With your background of Math 327/8 and 441 or 424, or 334/5, you deserve, and are expected to handle, more sophisticated analysis. In fact, this may sometimes make definitions and proofs simpler! So this supplement summarizes some definitions and results about differentiation for functions of several variables. If you have taken or are taking Math 326, Math 334/5, or Math 425, much of this material may be review.

The main reference is Folland's *Advanced Calculus*, which is on reserve in the Math Library. References below to [F] indicate that book, while [B&L] indicates our text.

Let S be a subset of $\mathbb{R}^n$, usually an open one, and suppose we have a function $f : S \to \mathbb{R}^m$ and $a \in S$ (or sometimes the closure of S). For this supplement (at least), I will dispense with putting arrows over vectors and vector-valued functions. We define

$$\lim_{x\to a} f(x) = L$$

to mean that for every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$||f(x) - L|| < \varepsilon \text{ whenever } 0 < ||x - a|| < \delta \text{ and } x \in S.$$

See [F], pp. 13-14, if you would like more discussion of this definition or what follows. (The reference uses $|x|$ instead of $||x||$ for the Euclidean norm of a vector.) When $n = 1$, you should confirm that this definition reduces to Definition 1.1.4, pp. 3-4, of [B&L]. Hopefully you will think that the new definition is an obvious generalization of 1.1.4, at least in hindsight. Continuity for a function at a point is defined as usual, that the limit at the point exists and is equal to the value of the function there.

The strict analog of Proposition 1.1.6, p. 4, in [B&L] holds: If $f(x) = (f_1(x), f_2(x), ..., f_m(x))$ and $L = (L_1, ..., L_m)$, then $\lim_{x\to a} f(x) = L$ if and only if $\lim_{x\to a} f_j(x) = L_j$ for $j = 1, 2, ..., n$. However, as mentioned the first week of class, if we try to make a similar proposition that focusses on components in the domain instead of the codomain, the result is not true in general. Here are the standard counterexamples; they are worked out in [F], pp. 14-15.

Example 1: A function that is continuous along both axes, but discontinuous along lines through the origin (so discontinuous at (0,0)):

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Example 2: A function that is continuous along every line through the origin, but nonetheless discontinuous at (0,0):

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

These examples show we must use the definition of limit carefully, and not just think about "approaching" a point in various ways. However, the basic facts about continuity - compositions of continuous functions and algebraic functions are continuous - follow as expected.

Next we consider differentiablity of $f : S \to \mathbb{R}^m$ for $S \subseteq \mathbb{R}^n$, $n > 1$. At a minimum, we would like to have the directional derivatives at a make sense. Recall that the directional derivative in a given direction is the derivative of $f(x(t))$, where $x(t)$ is a unit speed parametrization of the line with the given direction. In particular, the partial derivatives of f at a point a are directional derivatives. Existence of the partial derivatives at a is insufficient to ensure differentiability there, as we can see in example 1 above: The partials exist at the origin, but the directional derivatives in other directions do not.

The essential feature of a derivative is that it gives linear approximation of the function. For single variable calculus, this means

$$f(x) - f(a) \approx f'(x)(x-a), \text{ for } x, a \in \mathbb{R}.$$

To quantify how good the approximation is, let's write it as

$$f(x) - f(a) - f'(x)(x-a) \to 0,$$

but we need to say how quickly it goes to zero as x approaches a. In fact what we want is

$$\lim_{x\to a} \frac{f(x) - f(a) - f'(x)(x-a)}{x-a} = 0. \tag{1}$$

Exercise: Show that that the equation in (1) holds if and only if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(x).$$

If $f : \mathbb{R}^n \to \mathbb{R}^m$, to make sense of the approximation we need the number $f'(x)$ to be replaced by a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$; for the moment, let's call it L. (Remark: The terms *function, transformation, map,* and *mapping* are essentially interchangeable.) And since we can't divide by the vector $x - a$, let us rewrite the limit in (1) as

$$\lim_{x \to a} \frac{f(x) - f(a) - L(x - a)}{||x - a||} = 0, \text{ for } x, a \in \mathbb{R}^n. \tag{2}$$

Definitions. We define the function f to be *differentiable* at a if there is a linear function L so that the equation in (2) holds. This definition, with slightly different notation, appears at the top of p. 108 in [F]. Names and notation for the linear function L vary. Our text calls it the *differential* of f at a, and writes it $d\vec{f}_{\vec{a}}$ (p. 120). Folland calls it the *Fréchet derivative* and writes $Df(a)$. Let's use the [B&L] notation but allow omitting the arrows, and call df_a either the *derivative* or the *differential*. (In other books, you may see it called the *total derivative*, and denoted in many different ways, including $D_a f$, $Tf(a)$, $f_*(a)$, and even $f'(a)$. That last one is my least favorite because it tends to lead students to forget there's more than one independent variable.)

The matrix for df_a, which we'll denote $[df_a]$, has the first partials of f as its entries or columns. (Columns, you ask? If the codomain is $\mathbb{R}^m$ for $m > 1$, each partial is a vector.) To see this, let E_i be the vector with 1 in the ith slot and 0 in all other slots; in other words, the unit vector parallel to the x_i-axis, pointing in the positive x_i direction. If you set $x - a = hE_i$ in (2), you should see that you must have $L(E_i) = df_a(E_i) = \partial f/\partial x_i$ for (2) to hold. But $L(E_i) = df_a(E_i)$ is also the ith column in the matrix for df_a, giving the result claimed. The matrix $[df_a]$ is sometimes called the *Jacobian matrix* for f at a, and its determinant is the *Jacobian determinant* (or just the Jacobian).

If f is differentiable at a, then the directional derivative in the direction of a unit vector u is given by $df_a(u)$. However, existence of directional derivatives in all directions is not enough to ensure differentiability. Here is simple example to show this.

Example 3:

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Suppose $df_{(0,0)}$ exists, and let $u(u_1, u_2)$ be a unit vector. Then

$$df_{(0,0)}(u) = \frac{d}{dt} f(tu_1, tu_2) = \frac{d}{dt} \frac{t^3 u_1^2 u_2}{t^2} = u_1^2 u_2.$$

For $u = (1, 0)$ or $(0, 1)$, the directional derivative is zero. (We could also deduce this from the fact that f is identically zero along both axes, so the partial derivatives vanish at the origin.) So if the derivative $df_{(0,0)}$ exists, it must be the zero transformation. But then we would get zero for all the directional derivatives at the origin, which the computation above shows we do not.

Fortunately, if we require a bit more than just existence of the partial derivatives, we can be deduce that the function is differentiable. Specifically, if f and its first partial derivatives exist on a neighborhood of a and are continuous at a, then f is differentiable at a ([F], p. 57). For functions of several variables, we say the function is C^k (or *class* C^k) if the function and all its partials up to order k are continuous. Thus requiring a function to be C^1 is sufficient to ensure its differentiability.

Now let us turn to the Chain Rule. You may recall it as a formidable expression in terms of partial derivatives. Using the (total) derivative instead of partial derivatives, the Chain Rule for functions of several variables looks, except for the description of the functions and the notation for the derivative, exactly like the chain rule in first year calculus.

Chain Rule. Suppose $f : S \to \mathbb{R}^m$ for $S \subseteq \mathbb{R}^n$ and f is differentiable at a for some $a \in S$. Also suppose $g : T \to \mathbb{R}^p$, where $T \subseteq \mathit{image}(f) \subseteq \mathbb{R}^m$, and that g is differentiable at $f(a)$. Then $h = g \circ f$ is differentiable at a, and $dh_a = dg_{f(a)} \circ df_a$.

Here the linear function dh_a is given as the composition of two linear functions. If we look at the corresponding matrices, we get $[dh_a] = [dg_{f(a)}] \cdot [df_a]$, where $\cdot$ is included to emphasize the matrix multiplication on the right side of the equation. Exercise: For $n = 2$, $m = 3$, and $p = 1$, show that the matrix multiplication produces the formulas for the sums of products of partial derivatives you were given when you first saw the multivariable Chain Rule.

Next we review the Inverse and Implicit Function Theorems. These two theorems are closely related, and in fact essentially equivalent, in the following sense. Some books prove the former first, and use it to prove the latter. Other books do the reverse!

The Inverse Function Theorem, Theorem 3.18, p. 137 [F]. Suppose S is an open set in $\mathbb{R}^n$ and f is a C^1 function from S to $\mathbb{R}^n$. Also suppose $a \in S$ and df_a is invertible. (Requiring invertible is equivalent to requiring df_a is one-to-one, or that the determinant of $[df_a]$ is nonzero.) Then there is an open set $U \subseteq S$ with $a \in U$ and an open subset $V \subseteq \mathbb{R}^n$ with $f(a) = b \in V$ such that the restriction of f to U is bijective, and its inverse function g is also C^1. Furthermore, $[dg_{f(x)}] = [df_x]^{-1}$, so the partial derivatives of g are given by the entries in the inverse matrix for df. If f is C^k for $k > 1$, then g is also C^k.

The Implicit Function Theorem: the Hypersurface (codimension 1) Case, Theorem 3.1, pp. 114-115 in [F]. Suppose S is an open set in $\mathbb{R}^{n+1}$ and $F : S \to \mathbb{R}$ is C^1. Represent points in the domain as $(x, z) \in \mathbb{R}^n \times \mathbb{R}$, so $x \in \mathbb{R}^n$ and $z \in \mathbb{R}$, and consider the set $\{(x, z) : F(x, z) = c\}$ for a constant $c \in \mathbb{R}$. If $(a, b) \in S$, $F(a, b) = c$, and $\partial F/\partial z(a, b) \neq 0$, then there are open sets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}$ with $(a, b) \in U \times V \subseteq S$ such that the following hold.

1. There is a C^1 function $g : U \to V$ such that $F(x, g(x)) = c$ everywhere on U.
2. For every $x \in U$, the only solution $z \in V$ for $F(x, z) = c$ is $z = g(x)$.
3. The partial dervatives of g are given by $\dfrac{\partial g}{\partial x^i} = -\dfrac{\partial F/\partial x^i}{\partial F/\partial z}$.

4. If F is C^k for $k > 1$, then g is also C^k.

Other versions of the Implicit Function Theorem are similar. For a function $F : \mathbb{R}^{n+k} \to \mathbb{R}^k$, instead of requiring one partial derivative to be nonzero, we must have a $k \times k$ submatrix matrix of the derivative dF be nonsingular. The deteminant for this submatrix appears in the denominator of the formula for the partials of g, and the numerator is another determinant. See e.g., Theorem 3.9, p.118 in [F].

Supplementary problems.

Problem S3.5. The goal of this problem is to prove the uniqueness part of Theorem 3.4.1. Suppose $\vec{x}(s)$ and $\vec{y}(s)$ are two unit speed parametrized curves $I \to \mathbb{R}^3$ with the same nonzero curvature function κ and torsion function τ and the same initial conditions: for some $s_0 \in I$,

$$\vec{x}(s_0) = \vec{y}(s_0), \quad \vec{x}'(s_0) = \vec{y}'(s_0), \quad \text{and} \quad \vec{x}''(s_0) = \vec{y}''(s_0).$$

Denote the Frenet frames of $\vec{x}$ and $\vec{y}$ by $\{\vec{T}_x, \vec{P}_x, \vec{B}_x\}$ and $\{\vec{T}_y, \vec{P}_y, \vec{B}_y\}$, respectively. Let $f(s) = \vec{T}_x(s) \cdot \vec{T}_y(s) + \vec{P}_x(s) \cdot \vec{P}_y(s) + \vec{B}_x(s) \cdot \vec{B}_y(s)$.

(a) Prove that f is constant.

(b) Determine the constant value of f, and explain why this means that $\vec{T}_x(s) = \vec{T}_y(s)$ for all $s \in I$.

(c) Conclude that $\vec{x}(s) = \vec{y}(s)$ for all $s \in I$.

Remark: For part (c), we do need to solve a system of differential equations, but it is a much simpler system than the original one in (3.17), p. 85. In fact for part (c), first year calculus results are sufficient to draw the conclusion.

Problem S3.6. Suppose $G(x, y, z) = (3yz^2, -e^{xy})$ and $F(s, t) = (s \ln t, F_2(s, t), 9 - 2t)$, where you know the following about the function F_2: it is differentiable at (3,2),

$$F_2(3, 2) = -2, \quad \frac{\partial F_2}{\partial s}(3, 2) = \sqrt{2}, \quad \text{and} \quad \frac{\partial F_2}{\partial t}(3, 2) = -1.$$

Use the Chain Rule to find the derivative (Jacobian matrix) for $H = G \circ F$ at the point with coordinates $(s, t) = (3, 2)$.

Problem S3.7. Consider the equation $\quad e^w(x^3 + y^3 + z^3) = \sqrt{1 + w^2} + 3xyz \quad$ near the point $(x, y, z, w) = (1, 0, 0, 0)$.

(a) Does the Implicit Function Theorem say you can solve for w as a differentiable function $f(x, y, z)$ on a neighborhood of $(x, y, z) = (1, 0, 0)$, with $f(1, 0, 0) = 0$? If the answer is yes, find the approximate value of $f(1 + s, t, u)$ as a linear function of (s, t, u) for small (s, t, u).

(b) Does the Implicit Function Theorem say you can solve for z as a differentiable function $g(x, y, w)$ on a neighborhood of $(x, y, w) = (1, 0, 0)$, with $g(1, 0, 0) = 0$? If the answer is yes, find the approximate value of $g(1 + s, t, v)$ as a linear function of (s, t, v) for small (s, t, v).