Finally, some calculus. The coefficients in the equation of the tangent line have been identified as values of derivatives, so finding them leads us, after a long detour into theory, to finding derivatives. The point of the theory has been to show the importance of the operation of differentiating with respect to one variable while holding other variables constant. This is easy to do: you simply follow the rules from single-variable calculus. It only remains to describe the notation. If z has been defined by some expression in the variables x and y, then the derivative of z with respect to x, treating y as a constant, is denoted

$$\frac{\partial z}{\partial x} \text{ or } D_x z.$$

As usual, the result of finding this derivative is an expression that usually involves x and y. To use this result to find a tangent line, you will need to evaluate the expression at $(x, y) = (a, b)$. There is no nice notation for this.

Alternatively, you can think of the given expression as defining z as a function of x and y, which you typically write $z = f(x, y)$. This notation means that the function f requires two input variables and the value of the function is found by using the given expression with the first variable assigned to x and the second to y. One of the things that is typically done with functions is to evaluate them at arbitrary expressions. Thus, if

$$f(x, y) = x + y^2,$$

then

$$f(y, x) = y + x^2.$$

What this example shows is that one should never assume that the variables used to describe a function have any significance whatsoever. Unfortunately, most of the other notations used in the textbook violate this rule. The only notation that doesn't is one that uses f_1 to stand for the derivative of the function f with respect to its first variable. This has the advantage that one can write $f_1(a, b)$ for the result of evaluating this function at the point (a, b), i.e., the result of first differentiating the function and then evaluation the result at the base point.

Higher derivatives. Once one has a derivative, either as an expression or as a function, one can think of differentiating *that*. If one expects to do a lot of that sort of thing, an abbreviation is needed. Thus one writes

$$\frac{\partial^2 z}{\partial x^2} \text{ for } \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right); \; \frac{\partial^2 z}{\partial y \partial x} \text{ for } \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right);$$

and f_{12} for $(f_1)_2$.

Fortunately, you only need to be careful about the order of the variables in this expression when making the definition, since $f_{12} = f_{21}$ for functions that you will meet in practice.

Exercises 14.3 Find partial derivatives.

$$3x - 2y^4 \tag{11}$$

$$\frac{x - y}{x + y} \tag{15}$$

$$\sqrt{x^2 + y^2} \tag{33}$$

The chain rule.There are other derivatives involving functions of several variables that can be found. Suppose that z is given in terms of x and y, and that x and y are each given in terms of t. You could (and *Maple* does) use this to find z in terms of t and then calculate $D_t z$ (this is the neatest notation for this discussion).

Alternatively, you could apply the rules of differentiation to the expressions that you are given. Whatever the expression for z tells you is the last step in its computation is the first differentiation formula to be applied. In this process, expressions equal to $D_t z$ are obtained that can contain x, y, t, $D_t x$ and $D_t y$. Since x and y are given in terms of t, their expressions are used when you need to expand $D_t x$ and $D_t y$ in terms of t. The idea behind the chain rule for functions of several variables is to delay the expansion of $D_t x$ and $D_t y$ as long as possible. This gives an expression for $D_t z$ in which t has yet to appear outside of a subscript.

The fundamental linearity of differentiation formulas. Let's take a close look at the formulas of elementary calculus:

$$D_t(x + y) = D_t x + D_t y \qquad (S)$$

$$D_t(x \cdot y) = D_t x \cdot y + x \cdot D_t y \qquad (P)$$

$$D_t\big(f(x)\big) = f'(x) \cdot D_t x \qquad (C)$$

These rules suffice to differentiate all the functions met so far, when supplemented by special formulas for differentiating functions given by $f(x) =$ one the following expressions: a constant, x^n, e^x, $\ln x$, $\sin x$, or $\cos x$, $\arctan x$. A few more formulas are obtained to avoid deriving them from other formulas every time they are needed, but this short list of formulas is a good summary of elementary calculus. Of course, the course really deals with the understanding of functions that allows you to use these formulas and apply the results to things in the real world that can be modeled by this mathematical abstraction.

The thing to notice about formulas (S), (P) and (C) is that each *term* contains a *factor* that is a single application of D_t. This means that, in the setting at the start of this lecture,

$$D_t z = A \cdot D_t x + B \cdot D_t y, \qquad (*)$$

where A and B are expressions involving x and y.

Connection with partial derivatives. Formula $(*)$ holds independent of the dependence of x and y on t. The special cases used to define partial derivatives (or, at least, their values at particular points) are obtained by using the parameterizations: (1) $x = t$, $y = b$; or (2) $x = a$, $y = t$. This shows that $A = D_x z$ and $B = D_y z$. The usual statements of the chain rule are obtained by translating this statement into different notations.

A word about proving the chain rule. This discussion has emphasized how the chain rule appears in calculations, suppressing all aspects of proof. However, this is not completely lacking in mathematical rigor. The rules of elementary calculus were obtained with proofs based on a definitions in terms of limits. In fact, it would have been better to use the "good linear approximation" version introduced in connection with our discussion of the tangent plane. Inside these proofs are rules for obtaining a δ from a given ϵ, although we usually don't look that closely.

For functions given by expressions that we recognize, our calculation is a proof that the derivatives exist. In fact, a close examination of what we have when we have reached the stage of formula $(*)$ shows that the surface defined by the given expression of z in terms of x and y has a tangent plane wherever the calculation is valid.

What this approach does *not* do is tell how to deal with a new function of several variables. What should be done, if such a function is ever met, is to show from the definition of the function that it has a tangent plane, and then prove a version of the chain rule we have already stated in the case where the graph of $z = f(x, y)$ has a tangent plane and the curve $x = g(t)$, $y = h(t)$ is differentiable.

Since we are not likely to meet any new functions of several variables in this course, this approach is more suitable for *Advanced Calculus*. However, in a sense, this was exactly what was done in deriving formulas (S) and (P)! Look closely at the expressions $x + y$ and $x \cdot y$. What are the partial derivatives? Does this agree with the statement given for the chain rule? (Of course, these questions are rhetorical.)

Exercises

Find derivatives and identify role of the chain rule. In these examples, each variable plays a definite role. If desired, everything could be expressed in terms of the independent variable(s) only. This illustrates the way that computations are organized in a spreadsheet.

1. $z = x^2y + xy^2$, $x = 2 + t^4$, $y = 1 - t^3$.

3. $z = \sin x \cos y$, $x = \pi t$, $y = \sqrt{t}$.

7. $z = x^2 + xy + y^2$, $x = s + t$, $y = st$.

9. $z = \arctan(2x + y)$, $x = s^2t$, $y = s \ln t$.

19. $w = x^2 + y^2 + z^2$, $x = st$, $y = s \cos t$, $z = s \sin t$ (at $s = 1$, $t = 0$).

Gradients. There is one more thing to be seen in (∗). Whenever one has a sum of terms, each of which is a product of something of one type and something of another, it should be viewed as a dot product of vectors. We have already met the vector $\langle D_t x,\, D_t y\rangle$ as the velocity vector when $\mathbf{r}(t) = \langle x,\, y\rangle$ gives the position of a point at time t. This model suggests that it won't be long before we try to write $\mathbf{r}'(t)$ in the form $(ds/dt)\mathbf{T}$, but first we collect the other factors into a vector $\big\langle D_x z,\, D_y z\big\rangle$. When differentiating a function f instead of an expression z, this has the form $\langle f_1,\ f_2\rangle$. In this form, it is easy to imagine the generalization to functions of any number of variables. This vector is called the **gradient** of f and denoted ∇f. Gradients are very much a "function thing" since it emphasizes the domain of the function rather than the range — there is no good notation for the same object constructed from an expression.

Like all other derivatives, gradients will be evaluated at points of their domain when they appear in applications.

Directional derivatives. The chain rule can now be expressed as

$$D_t f\big(\mathbf{r}(t)\big) = \nabla f\big(\mathbf{r}(t)\big) \cdot \mathbf{r}'(t).$$

While f is an ordinary real valued function on some $\mathbb{R}^d$ (with $d = 2$ or $d = 3$ for most of the examples in this course, but a common description of those cases leads immediately to a vast generalization), ∇f is something else: for each point $P \in \mathbb{R}^d$, $\nabla f(P)$ is a vector in $\mathbb{R}^d$ based at P. One term frequently used for such a function $\mathbb{R}^d \to \mathbb{R}^d$ is **vector field**. The appearance of $\nabla f(P)$ in an inner product suggests that its principal interpretation will be in terms of expressions of the form $\nabla f(P) \cdot \mathbf{v}$. In particular, if $\mathbf{v}$ is a unit vector, this expression is called the **directional derivative** of f in the direction $\mathbf{v}$. The special case in which $\mathbf{v}$ lies along one of the coordinate axes has already been met under the name "partial derivative".

Now, as promised, we write $\mathbf{r}'(t) = (ds/dt)\mathbf{T}$ to express $D_t f\big(\mathbf{r}(t)\big)$ as the product of the *speed* with which $\mathbf{r}(t)$ is drawing the curve and the directional derivative of f in the *tangential direction* of the curve.

There are two important special cases: $\mathbf{v} \perp \nabla f$ is equivalent to the directional derivative being zero; the directional derivative takes its maximal value when $\mathbf{v} \parallel \nabla f$. Since

$$\nabla f \cdot (-\mathbf{v}) = -\nabla f \cdot \mathbf{v},$$

the minimum value of the directional derivative is in the *anti-parallel* direction.

An important application is that level curves of functions on $\mathbb{R}^2$, or level surfaces in $\mathbb{R}^3$, are perpendicular to the gradient of the function.

Changing coordinates. Since the last step of computing a gradient of f is to draw that vector field at points of the domain of f, the coordinates used in its computation have been pushed into the background. This suggests that it should be possible to perform these computations in other coordinate systems and get the same geometric answer. This is indeed true, with one important requirement — since inner products play an important role in the theory, only systems in which the coordinates are the components

of a set of *mutually perpendicular unit vectors* can be used. While it is usually better to use a disjoint set of names for the elements of different coordinate systems, the restriction to orthonormal coordinates allows the same name to appear in different coordinate systems as long as it means the component with respect to the same vector.

Exercises. Find gradient ∇f in general and at point P. If a direction at P is given, find the directional derivative in that direction.

#3. $f(x, y) = x^2y^3 + 2x^4y$, $P(1, -2)$, angle $\pi/3$ from positive x-axis.

#9. $f(x, y, z) = xy^2z^3$, $P(1, -2, 1)$, in direction $(1/\sqrt{3})\langle 1, -1, 1\rangle$.

#15. $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$, $P(1, 2, -2)$, direction of $\langle -6, 6, -3\rangle$.

Overview of beginning of chapter 16. If $f(x, y, z)$ is a real-valued function on $\mathbb{R}^3$, its **gradient** ∇f was introduced in Section 14.6. When it was introduced, we noted that it was an new kind of function called a **vector field**. In Section 16.1, a definition of this object finally appears, with some examples from physics. It is quickly noted that a vector field $\mathbf{F}$ that is of the form ∇f is very special. The word **conservative** is introduced to describe such vector fields, and f is given the special designation of the **potential function** of $\mathbf{F}$. This is supposed to inspire a feeling of warm fuzzies among those who have met these terms in a physics course. It is motivated by the concept of *potential energy*, which is an invention that allows one to claim a law of *Conservation of energy*. The change in the more observable *kinetic energy* is given by the **work integral**, which is the integral of the *tangential component* with respect of arc length of the force acting on an object. Formula (7) in Section 11.9 noted that only the tangential component of acceleration contributed to changes of speed, and this is what is being

measured here. This work integral takes the form

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{T}\, ds.$$

The $\mathbf{T}$ and ds in this formula should not be taken seriously, they only serve to establish the link with physics. The $\mathcal{C}$ in this notation represents the curve over which the object moves, and if $\mathcal{C}$ is given by a vector function $\mathbf{r}(t)$, then

$$\mathbf{T}\, ds = \mathbf{r}'(t)\, dt$$

from the definitions of $\mathbf{T}$ and s. The chain rule of elementary calculus shows that a change of parameter gives the usual change of variable in this integral, and hence, does not change the value of the integral. Thus, although we **describe** the integral using $\mathbf{T}\, ds$, we **compute** it using $\mathbf{r}'(t)\, dt$.

If $\mathbf{F} = \nabla f$, then the chain rule for functions of several variables shows that this integral is equal to the difference of the values of f at the ends of the curve $\mathcal{C}$. This is the first generalization of the fundamental theorem to appear in this course.

Integrals of vector fields. The text mentions some integrals of scalar functions with respect to arc length along curves. Such integrals (except for the calculation of arc length itself) are artificial. Only integrals of vector fields like the *work integral*, seem to appear in applications, and only these integrals have interesting mathematical properties. If $\mathcal{C}$ is a curve parameterized by $\mathbf{r}(t)$ for $a \le t \le b$, then we describe this integral as

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{T}\, ds$$

or

$$\int_{\mathcal{C}} \mathbf{F} \cdot\, d\mathbf{r},$$

and evaluate it as

$$\int_a^b \mathbf{F}\big(\mathbf{r}(t)\big) \cdot \mathbf{r}'(t)\, dt,$$

which we think of as a substitution in the previous expression. The value of this expression is independent of parameterization. If $\mathbf{F} = \langle P, Q, R \rangle$, then

$$\int_{\mathcal{C}} \mathbf{F} \cdot\, d\mathbf{r},$$

can be expanded as

$$\int_{\mathcal{C}} P\, dx + Q\, dy + R\, dz. \qquad (*)$$

In these formulas, P, Q, and R are functions of x, y, and z. Expanding everything in terms of the parameter t gives the same formula as before. This shows the power of the notation of expressions when doing calculus. Indeed, some of the things that you think you see in the form $(*)$ turn out to really be there and allow simplifications. For example, if you have $\int x\, dx$, it is equal to the difference of the values of $x^2/2$ at the two ends of the path.

If you need to give a piecewise definition of the curve $\mathcal{C}$, then the resulting integral will normally be written as the sum of the integrals over the pieces. For example, to integrate $\mathbf{F}$ in the counterclockwise direction around the square with vertices $(0, 0)$, $(1, 0)$, $(1, 1)$, $(0, 1)$, you form

$$\int_0^1 \mathbf{F}(x, 0)\, dx + \int_0^1 \mathbf{F}(1, y)\, dy$$

$$+ \int_1^0 \mathbf{F}(x, 1)\, dx + \int_1^0 \mathbf{F}(0, y)\, dy.$$

The fundamental theorem. If $F = \nabla f$, then the integrand in $(*)$ is

$$f_1(x, y, z)\, dx + f_2(x, y, z)\, dy + f_3(x, y, z)\, dz,$$

which the chain rule for several variables says is just $df(\mathbf{r}(t))$. The fundamental theorem of single variable calculus says that this integrates to the difference of the values of f at the points $\mathbf{r}(t)$ obtained from the endpoints of the interval in t over which you are integrating. These values of $\mathbf{r}(t)$ are just the endpoints of the arc $\mathcal{C}$. In particular, $f(x, y, z)$ can be determined up to an additive constant by integrating $\mathbf{F}$ along any path from a fixed base point (x_0, y_0, z_0) to (x, y, z).

Independence of path. The result just mentioned shows that the integral of a conservative vector field depends only on the endpoints of the arc $\mathcal{C}$ and not on the details of how $\mathcal{C}$ gets from one of those points to the other. However, you should remember to verify

the hypothesis of this theorem before jumping to its conclusion. This is mostly a theoretical result, not a shortcut for evaluating integrals. Direct use of the definition is an important skill which should not be abandoned before it is mastered.

Finding potential functions. Given a vector field $\mathbf{F}$ that is defined everywhere, we can build a path from (x_0, y_0, z_0) to (x, y, z) by following lines parallel to the axes to (x, y_0, z_0) and (x, y, z_0). If $\mathbf{F} = \nabla f$, the integral along this path gives a possible value of f, and all other choices of f differ from this by our old friend "$+ C$".

A necessary condition for a field to be conservative. Clairaut's theorem tells that $f_{ij} = f_{ji}$. If $\mathbf{F} = \nabla f$, this says that the partial derivative of the i-th component of $\mathbf{F}$ with respect to the j-th variable is everywhere equal to the partial derivative of the j-th component of $\mathbf{F}$ with respect to the i-th variable.

If $\mathbf{F}$ does not have this property, it cannot be conservative. Conversely, if it has this property, the standard path can be used to find a tentative choice of f. One can then compare the gradient of this function to $\mathbf{F}$.

Usually, one finds that $\nabla f = \mathbf{F}$. Green's theorem can be used to formulate and proof a precise statement.

The key example. If

$$\mathbf{F}(x, y) = \left\langle \frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right\rangle,$$

then it appears that our necessary condition is satisfied. However, $\mathbf{F}$ and its derivatives fail to be defined at $(0, 0)$. This causes $\mathbf{F}$ to fail to be conservative. It is easy to compute its integral around the unit circle, $x = \cos t$, $y = \sin t$, from $t = 0$ to $t = 2\pi$.

This simple example shows that independence of path requires that it be possible to compute the integral along a family of paths that describe how one can deform one path into another. A single point that prevents integrals through it from being defined can cause independence of path to fail.

The value of the integral we have given on any path not containing $(0, 0)$ is always an integer multiple of 2π. That integer can be interpreted as the number of

times the path goes around the origin in the counterclockwise sense (so that it is -1 if you go once around in the clockwise sense).

Exercises 16.2

Find $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$.

#19. $\mathbf{F} = \langle x^2 y^3, y\sqrt{x} \rangle$,

$\mathbf{r} = \langle t^2, -t^3 \rangle$,

$0 \le t \le 1$.

#21. $\mathbf{F} = \langle \sin x, \cos y, xz \rangle$,

$\mathbf{r} = \langle t^3, -t^2, t \rangle$,

$0 \le t \le 1$.

Exercises 16.3

Is $\mathbf{F} = \nabla f$?

#3. $\mathbf{F} = \langle 6x + 5y, 5x + 4y \rangle$

#5. $\mathbf{F} = \langle xe^y, ye^x \rangle$.