Computations in Differential Calculus of Two Variables Translated From Geometric Intuition
or
Gradients, Differentials, Tangent Planes, Directional Derivatives, Chain Rules and Partial Derivatives from One Picture

Mark Feldman

This article describes an approach I use when I teach differential calculus of several variables. It has the benefit of giving the students a single image from which they can visualize any of the tools described in the title(the gradient, the differential, the tangent plane, the directional derivative, the chain rule, and partial derivatives) and see how to go from that visualization to the necessary computations. Thus, they learn not only the geometric concepts in calculus of several variables, but, they also learn how to translate geometric ideas into mathematical computations. I have had considerable success with this method with students of varying abilities. To avoid having to continually write "I then tell them...", the rest of this article is mainly a summary of my lectures.

1 One Variable Derivatives

How should we visualize the derivative of a function of one variable? That's not hard — it's just a tangent line. How do we actually use this tangent line for calculations? We use the slope of the tangent line to f at the point a to estimate the value of f at another (close) point $a + h$.(See figure 1.)

Let's carefully focus on using $f'(a)h$ to approximate the increase (or decrease) in f as x changes from a to $a + h$. It's just the increase, as we go from the origin to h, up the straight line with slope $f'(a)$. (See figure 2.) But what lousy notation! Should we use x or h for what is normally the x-axis, etc... So, just forget the notation for now. We will get back to that later. Just focus on the picture: $f'(a)h$ is simply how much you go up a straight line through the origin — a straight line parallel to the tangent line to the function at $(a, f(a))$ — when you go over h units from the origin. We use $f'(a)h$ to approximate the change in f as we go from a to $a + h$: $f(a + h) \approx f(a) + f'(a)h$.

Before we go on to discuss the generalization of this idea to two variables, let's stop for a minute to discuss notation. It's important and can be confusing. Recall that if we just want the change in y as we change x from a to $a+h$, we use the approzimaiton $f'(a)h$, where h represents the change in x. Why don't we just use standard notation for the change in x? Let's just call it Δx. Then the approximation is $f'(a)\Delta x$. Let's give this quantity a name. How about Δy? No, that won't work. That's because Δy is already used for something - the actual change in y. So, let's use something close. How about dy? Then we get $dy = f'(a)\Delta x$. Now, that looks sloppy. Ok, let's use dx for Δx. Finally, we get

$$dy = f'(a)dx,$$

a very suggestive notation.

2 Two Variables

Now, instead of a curve, we have a surface and we want to approximate $f(a+h, b+k)$, assuming we are lucky enough to have the right kind of information about f. In one variable, we needed $f(a)$ and the tangent line. Maybe that will work here, if we have $f(a,b)$ and the tangent plane. (See figure 3.) We should approximate the change in f as though we were going up the tangent plane, from (a,b) to $(a+h, b+k)$. That is,

$$f(a+h, b+k) \approx f(a,b)+$$

(distance up or down tangent plane when the xy-coordinate changes from (a,b) to $(a+h, b+k)$.

(See figure 3.)

If we had the equation of a parallel tangent plane through the origin, we could solve for the change in z in terms of h and k, and we would be done. So, we just need a normal vector to the tangent plane to our surface at $(a,b,f(a,b))$. To get the normal vector, we just need two vectors in the tangent plane. Then, we can take their cross product to get the normal vector $\mathbf{n}$. Getting these two vectors isn't hard.

2.1 Getting Two Vectors in the Tangent Plane to a Surface Given by z=f(x,y)

Look at the function of *one* variable, $g(x) = f(x,b)$. Its tanget line is in the tangent plane to the surface. (Look at figure 4.) The slope is the slope of the function $f(x,b)$ at b. In other words, its just $\frac{df}{dx}(x,b)$, evaluated at a. Let's call it $f_x(a,b)$. Then, if we go from (a,b) to $(a+1,b)$, we go up this line, $f_x(a,b)\cdot 1 = f_x(a,b)$. In the tangent plane, we are going from $(a,b,f(a,b))$ to $(a+1, b, f(a,b)+f_x(a,b)\cdot 1)$. In other words, the vector $(1,0,f_x(a,b))$ is in the tangent plane. Doing the same thing in the y direction, we conclude that

$(0, 1, f_y(a, b))$ is in the tangent plane. Thus, a normal vector is $(1, 0, f_x(a, b)) \times (0, 1, f_y(a, b)) = (-f_x(a, b), -f_y(a, b), 1)$.

Recall that we wanted the equation of the plane through the origin, parrallel to our tangent plane. Its equation is:

$$\mathbf{n} \cdot (x, y, z) = 0$$

or

$$z = f_x(a, b)x + f_y(a, b)y.$$

Thus, the approximation to $f(a + h, b + k)$ that we are looking for is:

$$f(a + h, b + k) \approx f(a, b) + f_x(a, b)h + f_y(a, b)k.$$

2.2 Important Notation and Other Matters

Let's look more closely at the estimate

$$f(a + h, b + k) - f(a, b) \approx f_x(a, b)h + f_y(a, b)k.$$

We would like to call it Δz but, again, that's wrong. That's because Δz is already used for the exact change in the value of f, as we go from (a, b) to $(a + h, b + k)$, whereas $f_x(a, b)h + f_y(a, b)k$ is the change if the surface defined by f were replaced by its tangent plane at $(a, b, f(a, b))$. Let's call this change dz. We, then, might as well (as before) use dx for h and dy for k. And, since this formula works at any (a, b), we might as well just use (x, y). Then, we get the formula for the *differential* fo f,

$$dz = f_x(x, y)dx + f_y(x, y)dy.$$

The differential is a function of *four* variables. It depends on the point where we are approximating with the tangent plane, (x, y), and the change, (dx, dy).

Also, note that $f_x(x, y)$ is just the derivative of f as a function of x, with y held fixed. Similarly, with $f_y(x, y)$.

3 Directional Derivatives

Suppose we are on the surface corresponding to $f(x, y)$. What is the rate of change of f (i.e., the change in z), if we move in some direction? Here is the idea.

It's simply how much z changes in the tangent plane if we move from (x, y) to $(x, y) + (h, k)$, where $|(h, k)| = 1$. It's the change in z per unit change in the direction given by the vector (h, k). Since (h, k) is a unit vector, it is $(cos(\theta), sin(\theta))$, where θ gives the direction. Thus, the directional derivative in the direction θ is given by

$$
\begin{aligned}
\mathrm{D}_{\theta} f(a,b) &= f_x(a,b)cos(\theta) + f_y(a,b)sin(\theta) \\
&= (\mathrm{f}_x(a,b), f_y(a,b)) \cdot (cos(\theta), sin(\theta)) \\
&= \text{component}_{(cos(\theta),sin(\theta))}(f_x(a,b), f_y(a,b)).
\end{aligned}
$$

The value, $(component_{(cos(\theta),sin(\theta))}(f_x(a,b), f_y(a,b))$ is the component of $(f_x(a,b), f_y(a,b))$ along the vector $(cos(\theta), sin(\theta))$.

Note that this directional derivative is a maximum when $(cos(\theta), sin(\theta))$ and $(f_x(a,b), f_y(a,b))$ point in the same direction. Thus, the vector in the xy-plane that points in the direction where the rate of change is greatest has the same direction as $(f_x(a,b), f_y(a,b))$. This is called the gradient of f at (a,b). Its magnitude gives the maximum rate of change along the surface.

4 The Chain Rule

Suppose you moving along a path on the surface. The projection of your path is a curve in the xy-plane. Suppose you know the parameterization in time of the projection of your path, and, the function $z = f(x,y)$ that defines your surface. Can you compute how fast you are moving up and down the surface, that is, the rate of change in z, as you move along the surface? Yes. If the curve in the xy-plane is parameterized by $\mathbf{r}(t) = (x(t), y(t))$, then, at time t_o, if you move along the tangent to the curve $\mathbf{r}(t)$ at $\mathbf{r}(t_o)$ for one unit of time, you will move from $(x(t_o), y(t_o))$ to $(x(t_o), y(t_o)) + \mathbf{r}'(t_o)$. (See figure 5.) Thus, we only need to apply the differential at $r(t_0)(= (x(t_o, y(t_0)) = (x_0, y_0))$ to the change $r'(t_0)(= (x(t_o, y(t_0)) = (x_0, y_0))$. Thus, the derivative of $(f \circ r)(t)$ is given by the chain rule:

$$(f \circ r)'(t) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t).$$

(This is usually written suggestively as:

$$\frac{dz}{dt} = \frac{dz}{dx}\frac{dx}{dt} + \frac{dz}{dy}\frac{dy}{dt}.$$

A more complicated situation arises when the variables x and y are given in terms of other variables s and t. That is, $(x,y) = (x(s,t), y(s,t))$ and we want to find the partial derivatives of the composite function $f(x(s,t), y(s,t))$ with respect to the variables s and t. We can do this just as above when we realize that by fixing t at some t_0 and letting s change, we are moving along the parameterized curve $(x(s,t_0), y(s,t_0))$ and calculating the change in z when we change s by one unit. For example,

$$(f \circ r)'(s) = f_x(x(s,t_0), y(s,t_0))x_s(s,_t 0) + f_y(x(s,t_0), y(s,t_0))y_s(s,t_0).$$

5 Differentiability

Up to now, I have avoided addressing the fact that there are functions where f_x and f_y both exist, yet f is not differentiable. To finish this note, I will just explain how I address this.

After giving an example of such a function, I tell the students that what we mean by differentiable in one variable is that there is a tangent line approximation where the relative error between the approximation and h goes to zero as h goes to zero. It is then easy for them to see that the same idea is what is needed in the case of two variables.

Now, I'm ready to talk about limits in two variables and then proceed in a rather standard way to the other differentiability topics covered in a course in multivariable calculus.

Figure 1

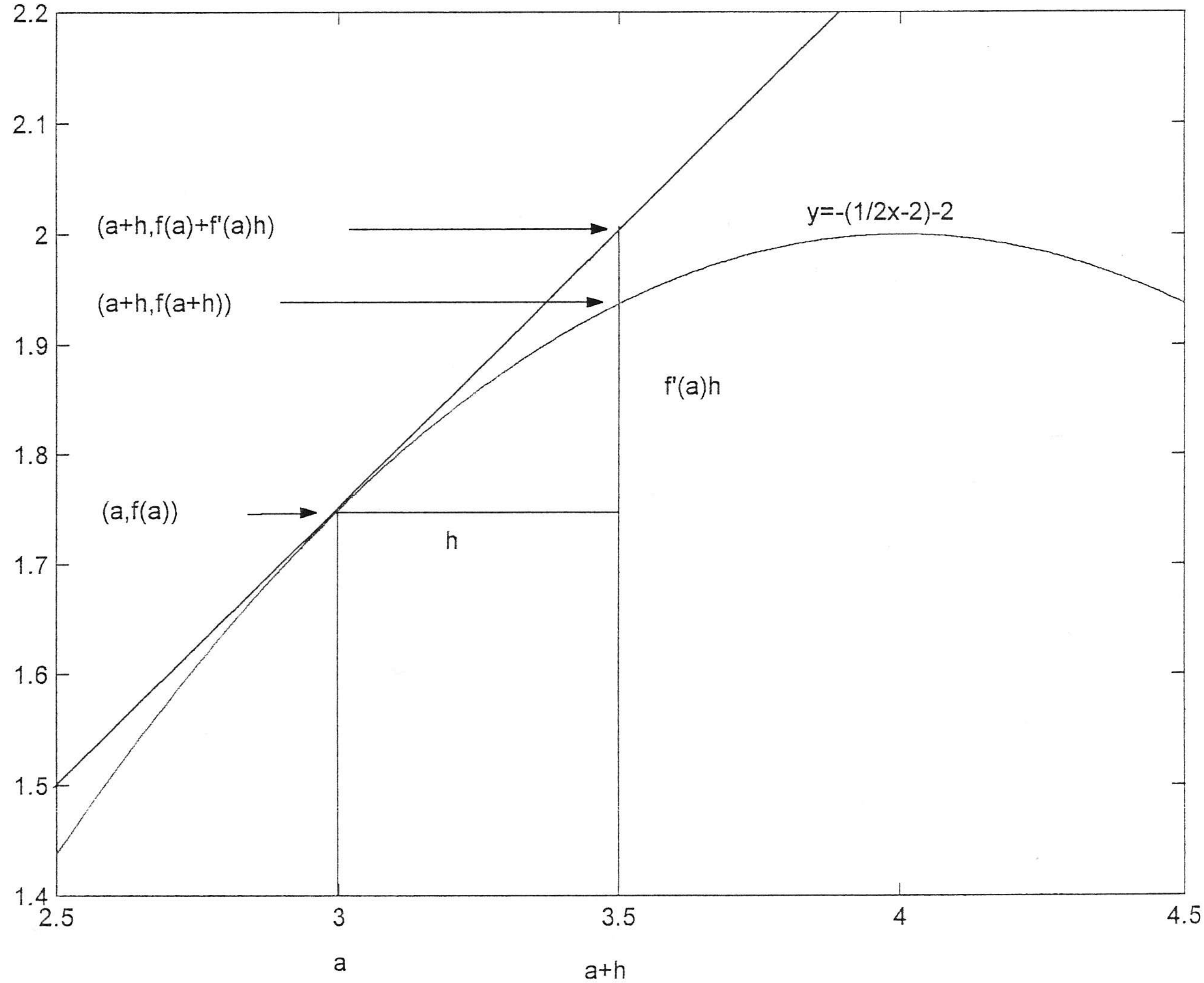

2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
2.5
3
3.5
4
4.5
(a+h,f(a)+f'(a)h)
(a+h,f(a+h))
(a,f(a))
y=-(1/2x-2)-2
f'(a)h
h
a
a+h

Figure 2(a)

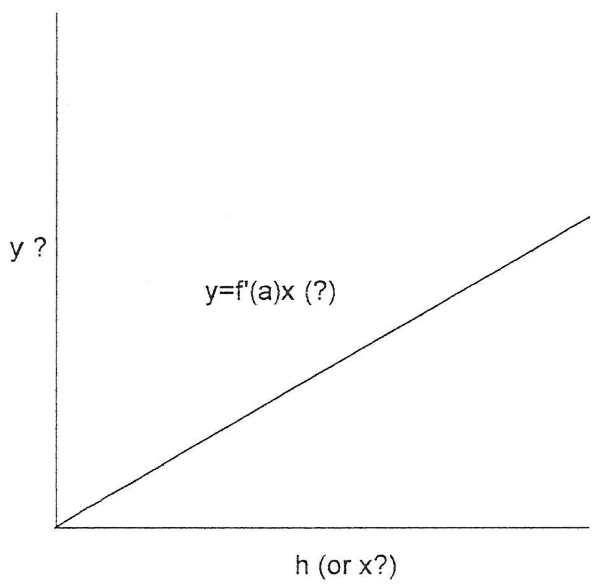
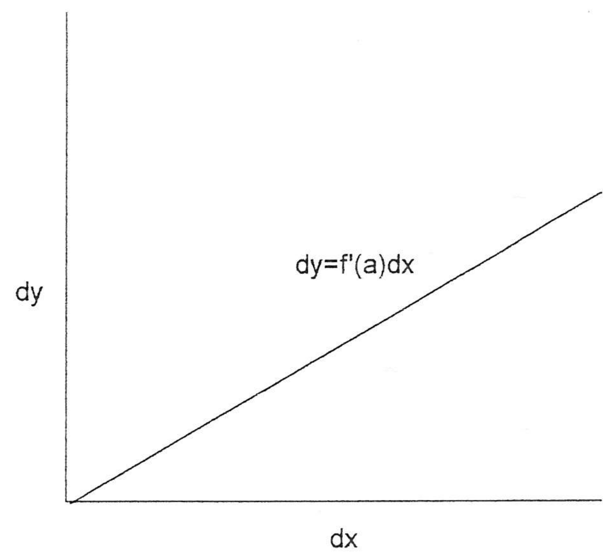


Figure 2(b)



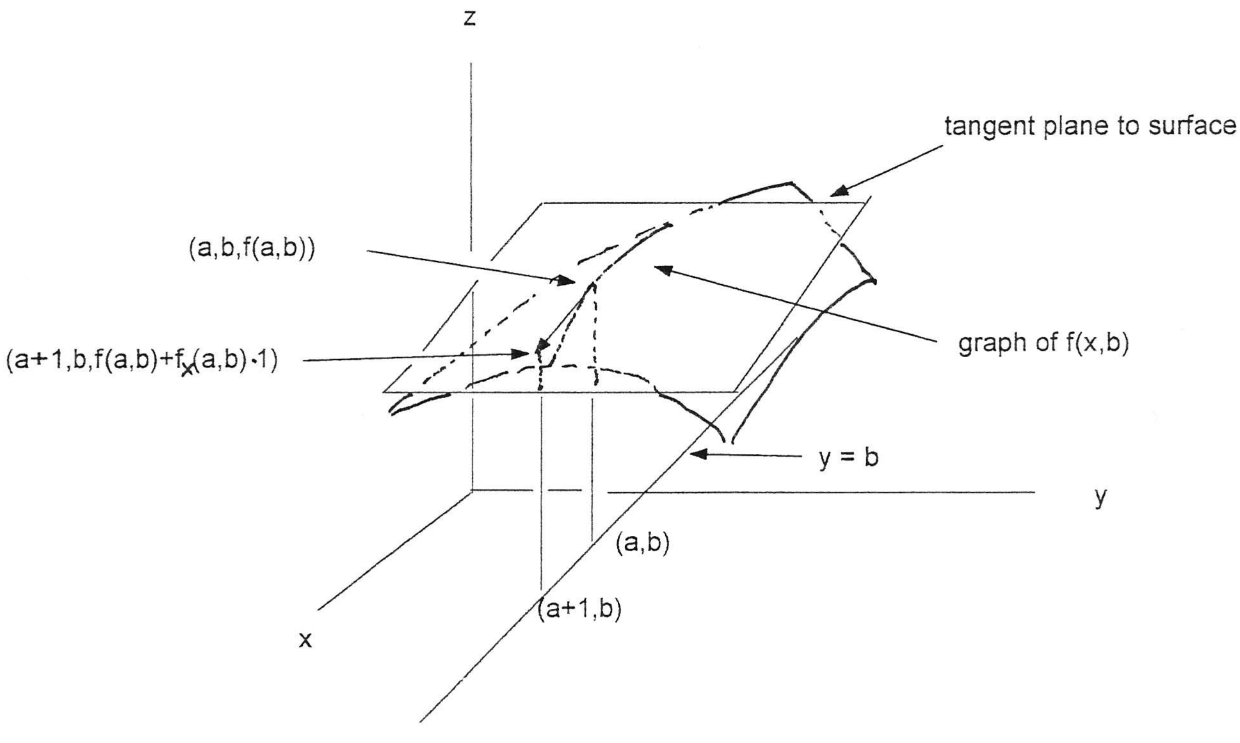
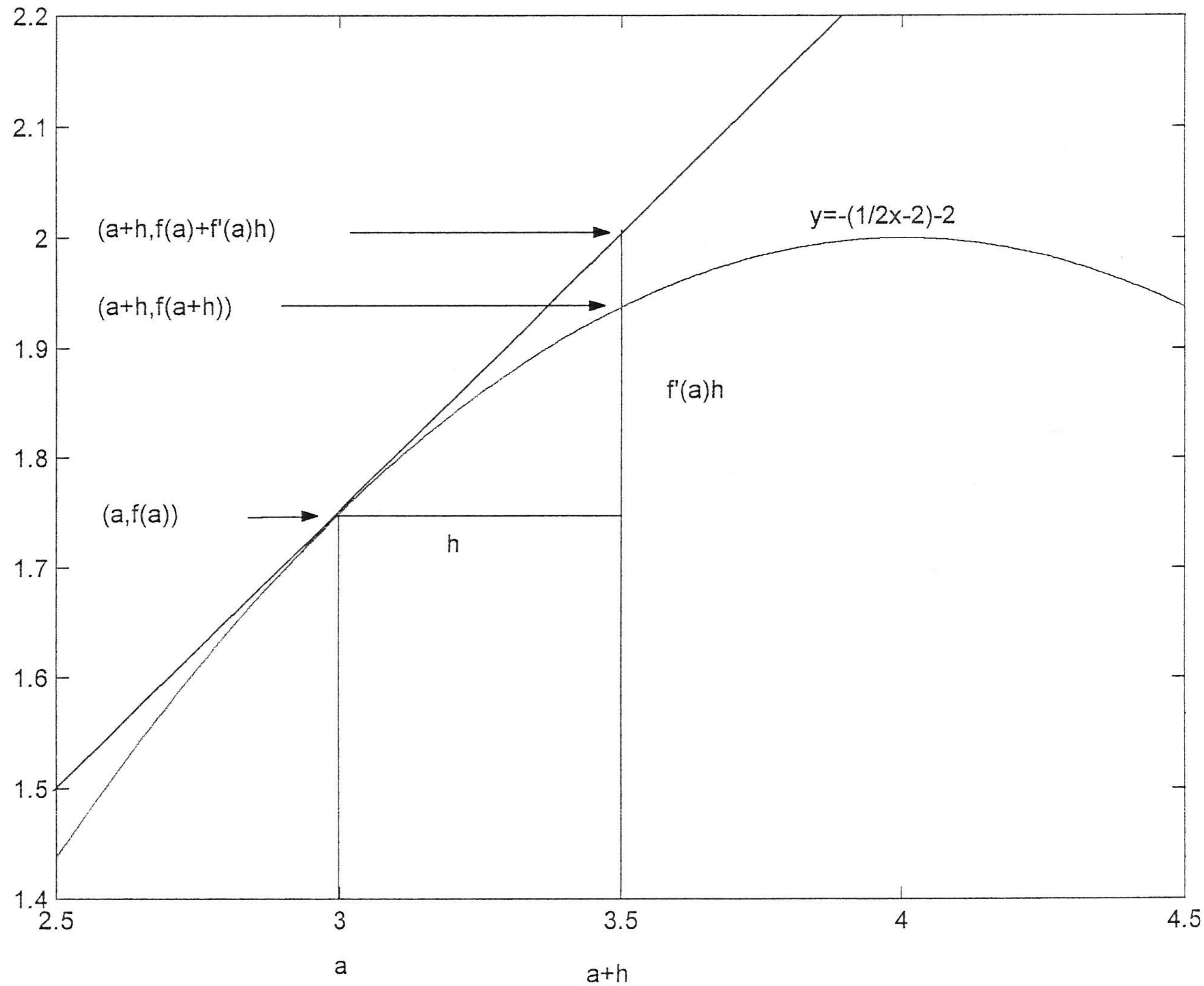

z
tangent plane to surface
(a,b,f(a,b))
graph of f(x,b)
(a+1,b,f(a,b)+f$_x$(a,b)·1)
y = b
y
(a,b)
(a+1,b)
x

Figure 4(b)

Figure 1



2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
2.5
3
3.5
4
4.5
(a+h,f(a)+f'(a)h)
(a+h,f(a+h))
(a,f(a))
y=-(1/2x-2)-2
f'(a)h
h
a
a+h

Figure 2(a)

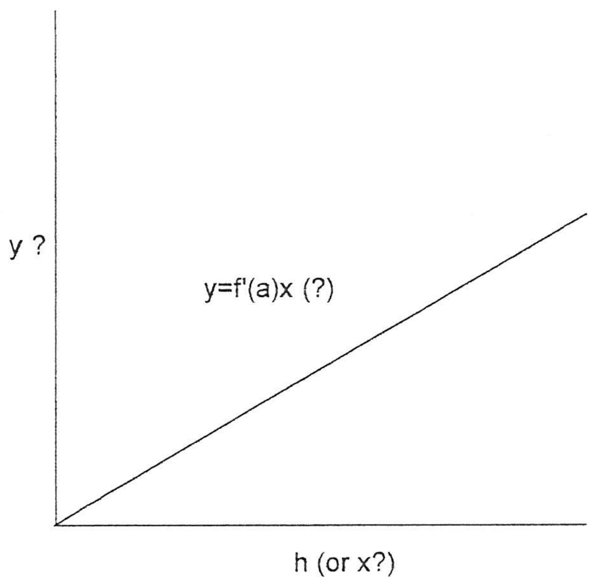
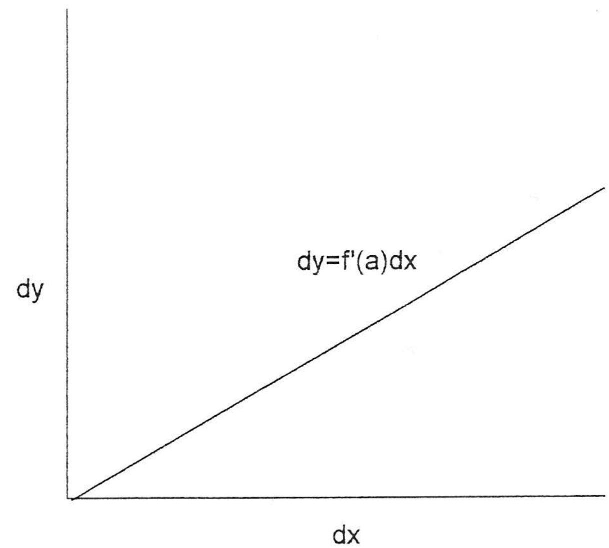


Figure 2(b)



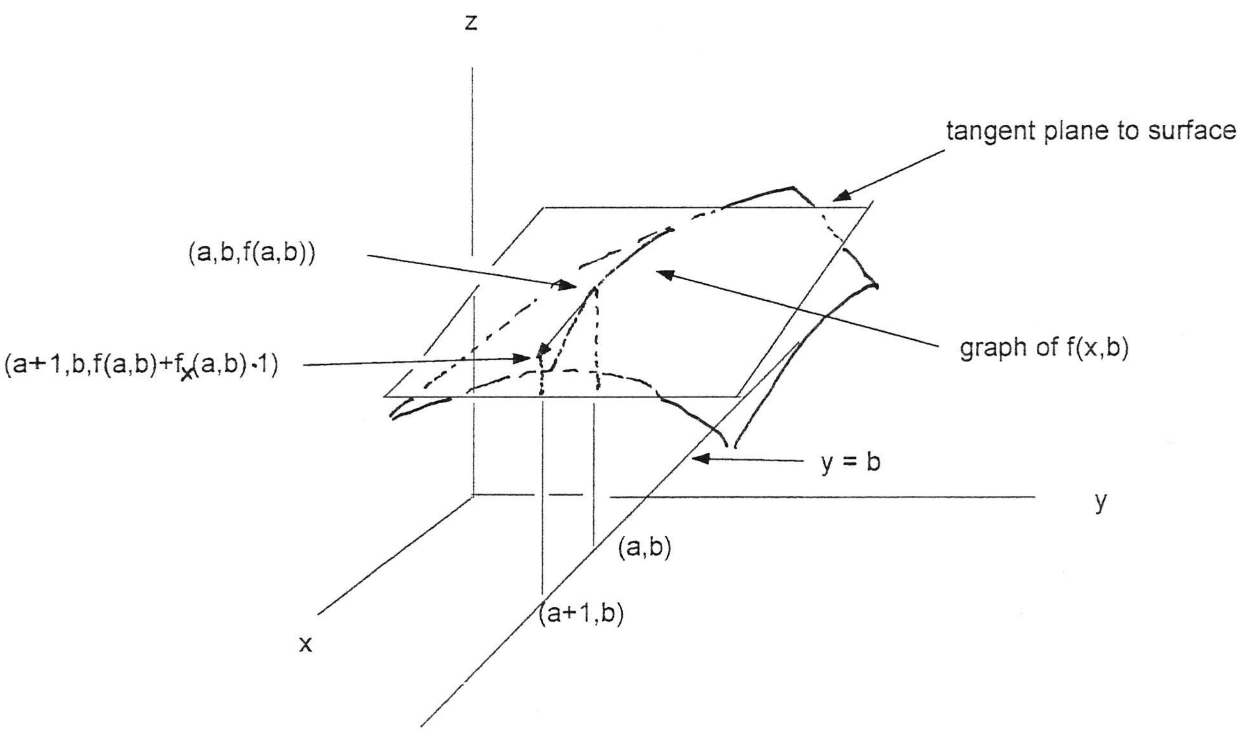

z
tangent plane to surface
(a,b,f(a,b))
graph of f(x,b)
(a+1,b,f(a,b)+f$_x$(a,b)·1)
y = b
y
(a,b)
(a+1,b)
x

Figure 4(b)