# FORK1005 Preparatory Course in Mathematics 2015/16 Lecture 4 

July 30, 2015

## 1 Introduction to Integration

### 1.1 Background: Calculating Areas

Given a function such as the one in Figure 4.1, how do you calculate the area between the curve $y=f(x)$ and the $x$-axis from $x=a$ to $x=b$ ? In the history of mathematics, this was one of the biggest questions mathematicians struggled with. The were plenty of reasons for wanting a mathematical method of calculating the area of irregular shapes:

- It would allow mathematicians to calculate the area of disks, ellipses, and other curved geometric objects;
- Physicists and engineers could better calculate the area and volume of physical objects they were working with;
- Governments and land-owners could determine the size of land they owned, in order to estimate the value of their property.


Figure 1.1: How do you calculate the area under a curve?

In spite of the huge interest, thousands of years would pass before mathematicians developed the mathematical framework to calculate the area of curved shapes. Finally, in the middle of the 17th century, mathematicians Gottfried Leibniz and Isaac Newton invented Calculus. Calculus is concerned with differentiation and integration, and they discovered that there is a deep link between finding the area below a curve and differentiating a function. This is often regarded is the most important discovery in all of science. Briefly put, they discovered that calculating the area below a function is the opposite of differentiating a function.

### 1.2 Idea of Integration

Recall that in differentiation, we are given a function $f(x)$ and we calculate its derivative, $f^{\prime}$. Integration is the inverse operation: In integration we are given a derivative $f^{\prime}$ and we calculate the original function $f$.

- So differentiation takes you from $f$ to $f^{\prime}$.
- And integration takes you from $f^{\prime}$ to $f$.

Theorem 1.1 (Fundamental Theorem of Calculus). Newton and Leibniz' famous discovery is called the Fundamental Theorem of Calculus. What it essentially says is that while differentiating a function gives you its rate of change, integrating a function gives you the area under the function's curve.

## 2 Integration

### 2.1 Antiderivative

The first concept we need for integration is the antiderivative:
Definition 2.1 (Antiderivative). Given a function $f$, a function $F$ is called the antiderivative of $f$ if

$$
F^{\prime}(x)=f(x)
$$

So the antiderivative is naturally the opposite of a derivative:
Example 2.2. Every function $f$ is an antiderivative to $f^{\prime}$.
Example 2.3. Let $C$ be any constant. If $F$ is an antiderivative to $f$, then $F+C$ is also an antiderivative to $f$. Why? When you differentiate a function, constant terms disappear, since they have a zero slope:

$$
(F(x)+C)^{\prime}=F^{\prime}(x)+(C)^{\prime}=F^{\prime}(x)+0=F^{\prime}(x)=f(x) .
$$

Example 2.4. What is the antiderivative of $2 x$ ?

1. We know that the derivative of $x^{2}$ is $2 x$. Therefore, $x^{2}$ is an antiderivative of $2 x$.
2. However, the derivative of $x^{2}+1$ is also $2 x$ so $x^{2}+1$ is another antiderivative of $2 x$.

Example 2.5. What is the antiderivative of $x^{2}$ ?

1. We know that the derivative of $x^{3}$ is $3 x^{2}$. So by dividing by 3 , we see that the derivative of $\frac{1}{3} x^{3}$ is $x^{2}$. So $\frac{1}{3} x^{3}$ is an antiderivative of $x^{2}$.
2. $\frac{1}{3} x^{3}+20$ is another antiderivative of $x^{2}$.

### 2.2 Integral

The integral is the mathematical operator that takes a function $f$ as an argument, and returns its antiderivative $F$ :

Definition 2.6 (Integral). The (indefinite) integral of $f$ is defined to be

$$
\int f(x) \mathrm{d} x:=F(x)+C
$$

where $F$ is an antiderivative of $f$ and $C$ is an unspecified constant. The $\int$-symbol is called the integral sign, the function $f$ appearing between $\int$ and $\mathrm{d} x$ is called the integrand. We include the $\mathrm{d} x$ to specify that $x$ is the variable of integration. We call this type of integral an indefinite integral

Example 2.7. - $\int 2 x \mathrm{~d} x=x^{2}+C$

- $\int x^{2} \mathrm{~d} x=\frac{1}{3} x^{3}+C$
- $\int \frac{1}{x} \mathrm{~d} x=\ln x+C$


### 2.3 Integration Rules

## Table of Rules:

| Addition | $\int f(x)+g(x) \mathrm{d} x=\int f(x) \mathrm{d} x+\int g(x) \mathrm{d} x$ |
| :---: | :---: |
| Scalar multiplication | $\int c f(x) \mathrm{d} x=c \int f(x) \mathrm{d} x$ |
| $x^{n}$ | $\int x^{n} \mathrm{~d} x=\frac{1}{n+1} x^{n+1}+C$ |
| $1 / x$ | $\int \frac{1}{x} \mathrm{~d} x=\ln x+C$ |
| $e^{x}$ | $\int e^{x} \mathrm{~d} x=e^{x}+C$ |

Example 2.8. Calculate

$$
\int x^{2}+2 x+4 \mathrm{~d} x:
$$

1. We use the addition and scalar multiplication rules to separate the integral:

$$
\int x^{2}+2 x+4 \mathrm{~d} x=\int x^{2} \mathrm{~d} x+2 \int x \mathrm{~d} x+4 \int 1 \mathrm{~d} x
$$

2. Each of these integrals are easy to calculate: We get

$$
\begin{aligned}
\int x^{2} \mathrm{~d} x+2 \int x \mathrm{~d} x+4 \int 1 \mathrm{~d} x & =\left(\frac{1}{3} x^{3}\right)+2\left(\frac{1}{2} x^{2}\right)+4(x)+C \\
& =\frac{1}{3} x^{3}+x^{2}+4 x+C
\end{aligned}
$$

3. So the answer is

$$
\int x^{2}+2 x+4 \mathrm{~d} x=\frac{1}{3} x^{3}+x^{2}+4 x+C
$$

## 3 Integration Techniques

Now we will look at three integration techniques, to help us calculate various integrals.

### 3.1 Integration by Parts

Recall the product rule for differentiation:

$$
\frac{\mathrm{d}}{\mathrm{~d} x}(u \cdot v)=\left(\frac{\mathrm{d} u}{\mathrm{~d} x}\right) \cdot v+u \cdot\left(\frac{\mathrm{~d} v}{\mathrm{~d} x}\right) .
$$

This can be rewritten as

$$
\left(\frac{\mathrm{d} u}{\mathrm{~d} x}\right) \cdot v=\frac{\mathrm{d}}{\mathrm{~d} x}(u \cdot v)-u \cdot\left(\frac{\mathrm{~d} v}{\mathrm{~d} x}\right) .
$$

By taking the integral on each side, we get the equality

$$
\begin{aligned}
\int\left(\frac{\mathrm{d} u}{\mathrm{~d} x}\right) \cdot v \mathrm{~d} x & =\int\left(\frac{\mathrm{d}}{\mathrm{~d} x}(u \cdot v)-u \cdot\left(\frac{\mathrm{~d} v}{\mathrm{~d} x}\right)\right) \mathrm{d} x \\
& =\int \frac{\mathrm{d}}{\mathrm{~d} x}(u \cdot v) \mathrm{d} x-\int u \cdot\left(\frac{\mathrm{~d} v}{\mathrm{~d} x}\right) \mathrm{d} x
\end{aligned}
$$

By the definition of the integral, we have

$$
\int \frac{\mathrm{d}}{\mathrm{~d} x}(u \cdot v) \mathrm{d} x=u \cdot v+C
$$

This gives us an integration identity called Integration by Parts:
Lemma 3.1. Let $u$ and $v$ be two functions. We then have the integration by parts identity

$$
\int u^{\prime} \cdot v \mathrm{~d} x=u \cdot v-\int u \cdot v^{\prime} \mathrm{d} x .
$$

(The $C$ constant of $u \cdot v$ is absorbed into $\int u \cdot v^{\prime} \mathrm{d} x$ )

The integration by parts identity helps us solve a variety of otherwise difficult integrals: Example 3.2. Using integration by parts, calculate the integral

$$
\int x \cdot \ln x \mathrm{~d} x
$$

1. We want to write $\int x \cdot \ln x \mathrm{~d} x$ in the form

$$
\int u^{\prime} \cdot v \mathrm{~d} x .
$$

So we need to choose suitable functions $u^{\prime}(x)$ and $v(x)$.
2. We try setting $u^{\prime}=x$ and $v=\ln x$. We then have the four expressions

$$
\begin{array}{ll}
u=\frac{1}{2} x^{2} & u^{\prime}=x \\
v=\ln x & v^{\prime}=\frac{1}{x}
\end{array}
$$

Furthermore

$$
\int x \cdot \ln x \mathrm{~d} x=\int u^{\prime}(x) \cdot v(x) \mathrm{d} x .
$$

3. The integration by parts formula is given by

$$
\int u^{\prime} \cdot v \mathrm{~d} x=u \cdot v-\int u \cdot v^{\prime} \mathrm{d} x
$$

Plugging in for our functions, we get the equation

$$
\begin{aligned}
\int x \cdot \ln x \mathrm{~d} x & =\frac{1}{2} x^{2} \cdot \ln (x)-\int \frac{1}{2} x^{2} \cdot \frac{1}{x} \mathrm{~d} x \\
& =\frac{1}{2} x^{2} \ln x-\frac{1}{2} \int x \mathrm{~d} x \\
& =\frac{1}{2} x^{2} \ln x-\frac{1}{2} \cdot \frac{1}{2} x^{2}+C \\
& =\frac{1}{2} x^{2} \ln x-\frac{1}{4} x^{2}+C
\end{aligned}
$$

which is our final answer.

### 3.2 Integration by Substitution

The next integration technique we will look at is Integration by Substitution. This technique is best understood in practice so we start with an example:
Example 3.3. Calculate the integral

$$
\int \frac{1}{1-x} \mathrm{~d} x:
$$

1. We don't know what the antiderivative of $\frac{1}{1-x}$ is. However, we do know that the antiderivative of $\frac{1}{x}$ is $\ln x$.
2. We would therefore like to rewrite the integral so that the integrand is of the form $1 / x$ instead of $\frac{1}{1-x}$.
3. Define a function $u:=1-x$. Then our integral can be written in the form

$$
\int \frac{1}{1-x} \mathrm{~d} x=\int \frac{1}{u} \mathrm{~d} x
$$

4. The integral $\int \frac{1}{u} \mathrm{~d} x$ cannot be solved in this form, because we need the variable of the integrand to match the variable of integration. Right now, the variable of integration is $x$ as in $\mathrm{d} x$, while the variable of the integrand is $u$. So we need to replace $\mathrm{d} x$ with $\mathrm{d} u$.
5. Since $u=1-x$, differentiating $u$ with respect to $x$ gives us

$$
\frac{\mathrm{d} u}{\mathrm{~d} x}=-1
$$

Cheating with algebra, we rearrange to get

$$
\mathrm{d} x=-\mathrm{d} u .
$$

6. So we replace $\mathrm{d} x$ by $(-\mathrm{d} u)$.

$$
\int \frac{1}{u} \mathrm{~d} x=\int \frac{1}{u}(-\mathrm{d} u)=-\int \frac{1}{u} \mathrm{~d} u .
$$

7. Since the antiderivative of $\frac{1}{u}$ is $\ln u$, we get that

$$
-\int \frac{1}{u} \mathrm{~d} u=-\ln u+C
$$

In other words,

$$
\int \frac{1}{1-x} \mathrm{~d} x=-\ln u+C
$$

8. We want the answer in terms of $x$, so as $u=1-x$, we get the final answer

$$
\int \frac{1}{1-x} \mathrm{~d} x=-\ln u+C=-\ln (1-x)+C
$$

Example 3.4. What is $\int \frac{1}{x+k} \mathrm{~d} x$ for a general $k$ ?

1. We do integration by substitution again. We want the integrand in the form $\frac{1}{u}$, so define $u:=x+k$.
2. We then rewrite the integral as

$$
\int \frac{1}{x+k} \mathrm{~d} x=\int \frac{1}{u} \mathrm{~d} x
$$

3. We differentiate $u$ to get

$$
\frac{\mathrm{d} u}{\mathrm{~d} x}=1
$$

so

$$
\mathrm{d} x=\mathrm{d} u
$$

4. We substitute into the integral to get

$$
\int \frac{1}{u} \mathrm{~d} x=\int \frac{1}{u} \mathrm{~d} u=\ln u+C .
$$

5. We plug $x+k$ back in for $u$ to get the answer

$$
\int \frac{1}{x+k} \mathrm{~d} x=\ln u+C=\ln (x+k)+C
$$

So

$$
\int \frac{1}{x+k} \mathrm{~d} x=\ln (x+k)+C
$$

## for all $k$.

### 3.3 Integration by Partial Fractions

The last technique we will look at is Integration by Partial Functions.
We have already integrated some fractions. We saw for example that

$$
\int \frac{1}{x+k} \mathrm{~d} x=\ln (x+k)+C
$$

But what about an expression like

$$
\int \frac{1}{x^{2}+4 x+3} \mathrm{~d} x ?
$$

Example 3.5. Solve the integral

$$
\int \frac{1}{x^{2}+4 x+3} \mathrm{~d} x
$$

1. The first thing we will do is to factorize the quadratic term $x^{2}+4 x+3$. We find that

$$
x^{2}+4 x+3=(x+3)(x+1) .
$$

2. So

$$
\int \frac{1}{x^{2}+4 x+3} \mathrm{~d} x=\int \frac{1}{(x+3)(x+1)} \mathrm{d} x .
$$

3. Now we want to write the integrand as partial fractions. That is, we want to find constants $A$ and $B$ that satisfies:

$$
\frac{A}{x+3}+\frac{B}{x+1}=\frac{1}{(x+3)(x+1)}
$$

We solve this algebraically:

$$
\begin{aligned}
\frac{A}{x+3}+\frac{B}{x+1} & =\frac{1}{(x+3)(x+1)} \\
\frac{A(x+1)}{(x+3)(x+1)}+\frac{B(x+3)}{(x+1)(x+3)} & =\frac{1}{(x+3)(x+1)} \\
\frac{A(x+1)+B(x+3)}{(x+3)(x+1)} & =\frac{1}{(x+3)(x+1)} \\
\frac{(A+B) x+(A+3 B)}{(x+3)(x+1)} & =\frac{0 x+1}{(x+3)(x+1)} \\
(A+B) x+(A+3 B) & =0 x+1 .
\end{aligned}
$$

We need the $x$-coefficients on each side to match each other, and the constant terms on each side to match:

$$
\begin{array}{r}
A+B=0 \\
A+3 B=1
\end{array}
$$

We subtract the first equation from the first:

$$
\begin{aligned}
A+3 B-(A+B) & =1-0 \\
2 B & =1 \\
B & =\frac{1}{2},
\end{aligned}
$$

and

$$
\begin{aligned}
& A+B=0 \\
& A+\frac{1}{2}=0 \\
& A=-\frac{1}{2}
\end{aligned}
$$

4. So we have that

$$
-\frac{1}{2(x+3)}+\frac{1}{2(x+1)}=\frac{1}{(x+3)(x+1)}
$$

We plug this into the integral:

$$
\begin{aligned}
\int \frac{1}{x^{2}+4 x+3} \mathrm{~d} x & =\int-\frac{1}{2(x+3)}+\frac{1}{2(x+1)} \mathrm{d} x \\
& =-\frac{1}{2} \int \frac{1}{x+3} \mathrm{~d} x+\frac{1}{2} \int \frac{1}{x+1} \mathrm{~d} x \\
& =-\frac{1}{2} \ln (x+3)+\frac{1}{2} \ln (x+1)
\end{aligned}
$$

and we are done.

## 4 Using Integration to Calculate Area

Definition 4.1 (Definite Integral). Let $f$ be a function with antiderivative $F$, and fix $a<b$. The definite integral of $f$ from $a$ to $b$ is defined to be

$$
\int_{a}^{b} f(x) \mathrm{d} x:=F(b)-F(a) .
$$



Figure 4.1: How do you calculate the area under a curve?
Definition 4.2 (Area of Curve). For a function $f(x)$, the area below the function's curve from $a$ to $b$ means the area between the $x$-axis and the function $f(x)$ for $a \geq x \geq b$. See picture above.

Theorem 4.3 (Fundamental Theorem of Calculus). Let $f$ be a function and fix $a<b$. Then the area below the curve of $f$ from $a$ to $b$ is equal to

$$
\int_{a}^{b} f(x) \mathrm{d} x .
$$

Example 4.4. Find the area created by the curve of $f(x)=3 x^{2}+2 x+4$ for $x=1$ to 3 .

$$
\begin{aligned}
A & =\int_{1}^{3} 3 x^{2}+2 x+4 \mathrm{~d} x \\
& =\int_{1}^{3} 3 x^{2} \mathrm{~d} x+\int_{1}^{3} 2 x \mathrm{~d} x+\int_{1}^{3} 4 \mathrm{~d} x \\
& =\left[x^{3}\right]_{1}^{3}+\left[x^{2}\right]_{1}^{3}+[4 x]_{1}^{3} \\
& =\left(3^{3}-1^{3}\right)+\left(3^{2}-1^{2}\right)+(4 \cdot 3-4 \cdot 1) \\
& =27-1+9-1+12-4 \\
& =42 .
\end{aligned}
$$

The area is 42 units.

# FORK1005 <br> Preparatory Course in Mathematics 2015/16 Lecture 5: Functions of Several Variables 

August 10, 2015

## 1 Introduction

Example 1.1. A company has to decide on the amount of money they spend on research and on advertising. They know that for a given year, their profit function is given by

$$
P(x, y)=-x^{2}-2 y^{2}+30 x+15 y-50
$$

where $x$ is amount spent on research and $y$ is amount spent on advertising. $x, y$ and $P(x, y)$ are all given in million USD. How does the company choose $x$ and $y$ so that profits $P(x, y)$ is maximized?
$P(x, y)$ is an example of a function of several variables. In order to solve maximization problems of this nature, we will need to extend our theory from functions of one variable. Things become more complicated when dealing with multiple variables, but most concepts are still the same:

- Instead of ordinary differentiation, we do partial differentiation: The function has derivatives with respect to each of its variables.
- When maximizing functions of one variable, we set the derivative to zero. In higher dimensions, we do the same, just that we have to set each partial derivative to zero.
- Second-order conditions follow the same principle as for one variable, but instead of just one ordinary second-derivative, there are many partial second-derivatives, which complicates things.

Notation. We will write $f(x, y)$ or $f(x, y, z)$ for ease of notation, but we could do all the same mathematics with functions of $n$ variables, $f\left(x_{1}, x_{2}, \ldots, x_{n}\right)$.

## 2 Partial Differentiation

Partial derivative is just the same as regular differentiation, just that we differentiate with respect to one specific variable, and hold the other variables constant.

Definition 2.1 (Partial derivative). Consider a function of two variables, $f(x, y)$. Its partial derivative with respect to $x$ is given by

$$
f_{x}^{\prime}(x, y):=\lim _{h \rightarrow 0} \frac{f(x+h, y)-f(x, y)}{h}
$$

and its partial derivative with respect to $y$ is given by

$$
f_{y}^{\prime}(x, y):=\lim _{h \rightarrow 0} \frac{f(x, y+h)-f(x, y)}{h}
$$

Example 2.2. If

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

then

$$
f_{x}^{\prime}(x, y)=2 x \quad \text { and } \quad f_{y}^{\prime}(x, y)=2 y
$$

Similarly, if

$$
g(x, y)=x^{2} y
$$

then

$$
g_{x}^{\prime}(x, y)=2 x y \quad \text { and } \quad g_{y}^{\prime}(x, y)=x^{2}
$$

When you compute $f_{x}^{\prime}$, think of $x$ as the variable and $y$ as a constant.
Example 2.3. Consider the function $f(x, y)=2 x^{2}+y^{2}$ evaluated at the point $(x, y)=(1,1)$. We have

$$
\begin{array}{rlrl}
f(x, y) & =2 x^{2}+y^{2} & f(1,1) & =2+1=3 \\
f_{x}^{\prime}(x, y) & =4 x & f_{x}^{\prime}(1,1) & =4 \\
f_{y}^{\prime}(x, y) & =2 y & f_{y}^{\prime}(1,1) & =2
\end{array}
$$

This tells us two things:

- $f$ is increasing at a rate of 4 units in the direction of $x$ at the point $(1,1)$;
- $f$ is increasing at a rate of 2 units in the direction of $y$ at the point $(1,1)$.


## 3 First-Order Conditions

Theorem 3.1 (First-order conditions). Suppose $\left(x^{*}, y^{*}\right)$ is a maximum point or a minimum point of $f(x, y)$. Then

$$
f_{x}^{\prime}(x, y)=f_{y}^{\prime}(x, y)=0
$$

Definition 3.2 (Stationary point). A stationary point $\left(x_{*}, y_{*}\right)$ of $f$ is any point where

$$
f_{x}^{\prime}(x, y)=f_{y}^{\prime}(x, y)=0
$$

Corollary 3.3. So all maximum and minimum points are also stationary points.
Example 3.4. Take the function $f(x, y)=x^{2}+y^{2}$. Suppose we want to find all its maximum and minimum points. We find its partial derivatives, $f_{x}^{\prime}(x, y)=2 x$ and $f_{y}^{\prime}(x, y)=2 y$ and set them equal to zero:

$$
f_{x}^{\prime}(x, y)=2 x=0 \quad \text { and } \quad f_{y}^{\prime}(x, y)=2 y=0
$$

so $(x, y)=(0,0$ is the only stationary point. And as you can see from the following graph, $(0,0)$ is a global minimum:


## 4 Second-Order Derivatives

### 4.1 Saddle Points

So the first-order conditions (FOCs) tell us that any maximum or minimum point is a stationary point. But this is not enough to be able to maximize or minimize a function. Firstly, we don't know if a stationary point is a minimum point or a maximum point, or a saddle point.

Example 4.1. The function $f(x, y)=-x^{2}+y^{2}$ has a stationary point at $(x, y)=(0,0)$. But as you can see from the graph below, it is neither a maximum or a minimum point. If you move in the $x$-direction, it decreases, but if you move in the $y$-direction, it increases. Such a point is called a saddle point.


We can classify stationary points into three types:

- Maxima
- Minima
- Saddle points


### 4.2 Second-Order Partial Derivatives

To determine whether a stationary point is a maximum, minimum or neither, we need second-order partial derivatives. This is just a (partial) derivative of a partial derivative.

Definition 4.2 (Second-order partial derivative). Let $f(x, y)$ be a function of two variables. Then its four second-order partial derivatives are

$$
\begin{aligned}
f_{x x}^{\prime \prime}(x, y) & :=\lim _{h \rightarrow 0} \frac{f_{x}^{\prime}(x+h, y)-f_{x}^{\prime}(x, y)}{h} \\
f_{x y}^{\prime \prime}(x, y) & :=\lim _{h \rightarrow 0} \frac{f_{x}^{\prime}(x, y+h)-f_{x}^{\prime}(x, y)}{h} \\
f_{y y}^{\prime \prime}(x, y) & :=\lim _{h \rightarrow 0} \frac{f_{y}^{\prime}(x, y+h)-f_{y}^{\prime}(x, y)}{h}
\end{aligned}
$$

and

$$
f_{y x}^{\prime \prime}(x, y):=\lim _{h \rightarrow 0} \frac{f_{y}^{\prime}(x+h, y)-f_{y}^{\prime}(x, y)}{h} .
$$

Example 4.3. Consider the function $f(x, y)=x^{2}+y^{2}$. We have the partial derivatives

$$
f_{x}^{\prime}(x, y)=2 x, \quad \text { and } \quad f_{y}^{\prime}(x, y)=2 y
$$

and the second-order partial derivatives

$$
\begin{aligned}
& f_{x x}^{\prime \prime}(x, y)=2 \\
& f_{x y}^{\prime \prime}(x, y)=0 \\
& f_{y y}^{\prime \prime}(x, y)=2
\end{aligned}
$$

and

$$
f_{y x}^{\prime \prime}(x, y)=0
$$

Example 4.4. Consider the function $g(x, y)=e^{x} y+2 x y-x y^{2}$. We have the partial derivatives

$$
g_{x}^{\prime}(x, y)=e^{x} y+2 y-y^{2}, \quad \text { and } \quad f_{y}^{\prime}(x, y)=e^{x}+2 x-2 x y
$$

and the second-order partial derivatives

$$
\begin{aligned}
g_{x x}^{\prime \prime}(x, y) & =e^{x} y \\
g_{x y}^{\prime \prime}(x, y) & =e^{x}+2-2 y \\
g_{y y}^{\prime \prime}(x, y) & =e^{x}+2-2 y
\end{aligned}
$$

and

$$
g_{y x}^{\prime \prime}(x, y)=-2 x
$$

Notice in the two previous examples that $f_{x y}^{\prime \prime}=f_{y x}^{\prime \prime}$ and $g_{x y}^{\prime \prime}=g_{y x}^{\prime \prime}$. This is not a coincidence:

Theorem 4.5 (Schwarz' theorem). For (almost ${ }^{1}$ ) all functions $f$, we have the equality

$$
f_{x y}^{\prime \prime}(x, y)=f_{y x}^{\prime \prime}(x, y)
$$

So this simplifies the calculations of our second-order partial derivatives a bit: We only need to compute three second-order partial derivatives, not all four.

Notation. There are many notations in use for expressing partial derivatives:

$$
f_{x y}^{\prime \prime}(x, y)=f_{x y}(x, y)=\partial_{x y} f(x, y)=\frac{\partial^{2} f}{\partial x y}(x, y)
$$

Here we will stick to $f_{x y}^{\prime \prime}(x, y)$, but in other literature/courses/websites you might find different notation.

[^0]
## 5 Second Partial Derivative Test

The second partial derivative test is a test that will tell us if any stationary point is a minimum, a maximum or a saddle point:

Theorem 5.1. Suppose $\left(x_{*}, y_{*}\right)$ is a stationary point of $f(x, y)$. Define

$$
D(x, y):=f_{x x}^{\prime \prime}\left(x_{*}, y_{*}\right) f_{y y}^{\prime \prime}\left(x_{*}, y_{*}\right)-\left(f_{x y}^{\prime \prime}\left(x_{*}, y_{*}\right)\right)^{2} .
$$

Then

1. If

$$
D(x, y)>0 \quad \text { and } \quad f_{x x}^{\prime \prime}\left(x_{*}, y_{*}\right)>0
$$

then $\left(x_{*}, y_{*}\right)$ is a local minimum of $f$;
2. If

$$
D(x, y)>0 \quad \text { and } \quad f_{x x}^{\prime \prime}\left(x_{*}, y_{*}\right)<0
$$

then $\left(x_{*}, y_{*}\right)$ is a local maximum of $f$;
3. If

$$
D(x, y)<0
$$

then $\left(x_{*}, y_{*}\right)$ is a saddle point of $f$.
4. If

$$
D(x, y)=0
$$

then the test is inconclusive.
Example 5.2. Consider the function $f(x, y)=x^{2}+y^{2}$. It has a stationary point at $(0,0)$. Its second order partial derivatives are

$$
f_{x x}^{\prime \prime}(x, y)=2, \quad f_{x y}^{\prime \prime}(x, y)=0 \quad \text { and } \quad f_{y y}^{\prime \prime}(x, y)=2
$$

Therefore

$$
D(0,0)=f_{x x}^{\prime \prime}(0,0) f_{y y}^{\prime \prime}(0,0)-\left(f_{x y}^{\prime \prime}(0,0)\right)^{2}=2^{2}-0=4>0
$$

and

$$
f_{x x}^{\prime \prime}(x, y)=2>0
$$

so by the second partial derivative test, $(0,0)$ is a local minimum.
Example 5.3. Consider the function $g(x, y)=-x^{2}+y^{2}$. It has a stationary point at $(0,0)$. Its second order partial derivatives are

$$
g_{x x}^{\prime \prime}(x, y)=-2, \quad g_{x y}^{\prime \prime}(x, y)=0 \quad \text { and } \quad g_{y y}^{\prime \prime}(x, y)=2
$$

Therefore

$$
g_{x x}^{\prime \prime}(0,0) g_{y y}^{\prime \prime}(0,0)-\left(g_{x y}^{\prime \prime}(0,0)\right)^{2}=-2 \cdot 2-0=-4<0
$$

so by the second partial derivative test, $(0,0)$ is a saddle point.

### 5.1 The Hessian Matrix

With a very basic knowledge of $2 \times 2$ matrices and determinants, we can rephrase the second partial derivative test in a more concise way.
Definition 5.4 (Determinant). Let $A=\left[\begin{array}{ll}a & b \\ c & d\end{array}\right]$ be a $2 \times 2$ matrix. Then its determinant $\operatorname{det}(A)$ is the value

$$
\operatorname{det}(A)=a d-b c
$$

Definition 5.5 (Hessian matrix). For a function $f(x, y)$, the Hessian matrix, $H f$, is the matrix containing the second-order partial derivatives of $f$ :

$$
H f(x, y)=\left[\begin{array}{ll}
f_{x x}^{\prime \prime}(x, y) & f_{x y}^{\prime \prime}(x, y) \\
f_{y x}^{\prime \prime}(x, y) & f_{y y}^{\prime \prime}(x, y)
\end{array}\right]
$$

Proposition 5.6. The determinant of $H f(x, y)$, denoted $D(x, y)$, is given by

$$
D(x, y)=f_{x x}^{\prime \prime}(x, y) f_{y y}^{\prime \prime}(x, y)-\left(f_{x y}^{\prime \prime}(x, y)\right)^{2}
$$

This is exactly the expression we calculate in the second partial derivative test. So we can rephrase the test:

Theorem 5.7. Suppose $\left(x_{*}, y_{*}\right)$ is a stationary point of $f(x, y)$. Then

1. If

$$
D(x, y)>0 \quad \text { and } \quad f_{x x}^{\prime \prime}\left(x_{*}, y_{*}\right)>0
$$

then $\left(x_{*}, y_{*}\right)$ is a local minimum of $f$;
2. If

$$
D(x, y)>0 \quad \text { and } \quad f_{x x}^{\prime \prime}\left(x_{*}, y_{*}\right)<0
$$

then $\left(x_{*}, y_{*}\right)$ is a local maximum of $f$;
3. If

$$
D(x, y)<0
$$

then $\left(x_{*}, y_{*}\right)$ is a saddle point of $f$.
4. If

$$
D(x, y)=0
$$

then the test is inconclusive.

## 6 Maximizing/Minimizing functions

Going back to our initial problem:
Example 6.1. You want to maximize a company's profit function, given by

$$
P(x, y)=-x^{2}-2 y^{2}+30 x+15 y-50
$$

where $x$ is amount spent on research and $y$ is amount spent on advertising, all given in millions USD.

1. First you want to locate stationary points, so you compute partial derivatives and set them equal to 0 :

$$
\begin{aligned}
& P_{x}^{\prime}(x, y)=-2 x+30=0 \quad \Longrightarrow \quad x=15 \\
& P_{y}^{\prime}(x, y)=-4 y+15=0 \quad \Longrightarrow \quad y=\frac{15}{4} .
\end{aligned}
$$

So the only stationary point is $(x, y)=(15,15 / 4)$.
2. We are looking for a maximum, so we compute the second-order partial derivatives to do the second partial derivative test:

$$
\begin{aligned}
& P_{x x}^{\prime \prime}(x, y)=-2, \\
& P_{y y}^{\prime \prime}(x, y)=-4
\end{aligned}
$$

and

$$
P_{x y}^{\prime \prime}(x, y)=0
$$

This means that
$D(15,15 / 4))=P_{x x}^{\prime \prime}(15,15 / 4) P_{y y}^{\prime \prime}(15,15 / 4)-\left(P_{x y}^{\prime \prime}(15,15 / 4)\right)^{2}=(-2) \cdot(-4)-0=8>0$,
and

$$
P_{x x}^{\prime \prime}(15,15 / 4)=-2<0
$$

so by the second partial derivative test, $(15,15 / 4)$ is a maximum.
So to maximize profits, the company should invest $15,000,000$ USD in research, and $15,000,000 / 4=3,750,000$ USD in advertising.

## 7 Global vs Local Maximum \& Minimum

The second partial derivative test locates local maxima and minima, but not necessarily global maxima or minima. A local maximum of a function maximizes the function within a bounded area, but not necessarily everywhere. For example, the function in the figure below has a local maximum at $x=0$, because in a neighborhood around $x=0$ (the red square), it maximizes the function. But outside of the red square, there are points where $f(x)$ is larger than $f(0)$. Therefore, $x=0$ is not a global maximum.


Figure 7.1: This function has a local maximum at $x=0$, but it is not a global maximum.

In contrast, the maximum below is also a global maximum:


Figure 7.2: This function has a global maximum at $x=0$.

Definition 7.1 (Local maximum). A point $\left(x_{*}, y_{*}\right)$ is a local maximum of $f$ if $\left(x_{*}, y_{*}\right)$ maximizes $f$ for all points $(x, y)$ sufficiently close ${ }^{2}$ to $\left(x_{*}, y_{*}\right)$.

Definition 7.2 (Global maximum). A point $\left(x_{*}, y_{*}\right)$ is a global maximum of $f$ if $\left(x_{*}, y_{*}\right)$ maximizes $f$ for all possible points $(x, y)$.

The second partial derivative test tells you if a point is a local maximum or minimum. It does not tell you if this maximum or minimum is global.

[^1]
## 8 Convex vs Concave Functions

If a function is convex or concave, we can determine whether a local maximum or minimum is also a global maximum/minimum.

Definition 8.1. A function $f(x, y)$ is convex if

$$
D(x, y)=f_{x x}^{\prime \prime}(x, y) f_{y y}^{\prime \prime}(x, y)-\left(f_{x y}^{\prime \prime}(x, y)\right)^{2} \geq 0
$$

and

$$
f_{x x}^{\prime \prime}(x, y) \geq 0
$$

for all points $(x, y)$.
Definition 8.2. A function $f(x, y)$ is concave if

$$
D(x, y)=f_{x x}^{\prime \prime}(x, y) f_{y y}^{\prime \prime}(x, y)-\left(f_{x y}^{\prime \prime}(x, y)\right)^{2} \geq 0
$$

and

$$
f_{x x}^{\prime \prime}(x, y) \leq 0
$$

for all points $(x, y)$.

## Theorem 8.3.

- If $f(x, y)$ is convex, all stationary points of $f$ are global minima.
- If $f(x, y)$ is concave, all stationary points of $f$ are global maxima

Example 8.4. Consider the function $f(x, y)=x^{2}+y^{2}$. We saw earlier that it had a local minimum at $(0,0)$. We note that

$$
f_{x x}^{\prime \prime}(x, y)=2, \quad f_{y y}^{\prime \prime}(x, y)=2 \quad \text { and } \quad f_{x y}^{\prime \prime}(x, y)=0
$$

Therefore

$$
D(x, y)=f_{x x}^{\prime \prime}(x, y) f_{y y}^{\prime \prime}(x, y)-\left(f_{x y}^{\prime \prime}(x, y)\right)^{2}=2 \cdot 2-0=4>0
$$

and

$$
f_{x x}^{\prime \prime}(x, y)=2>0
$$

for all $(x, y)$, so $\boldsymbol{f}$ is convex. Therefore, $(0,0)$ is a global minimum.

Example 8.5. Now we return to the profit maximization problem

$$
P(x, y)=-x^{2}-2 y^{2}+30 x+15 y-50
$$

We found that $P$ had a local maximum at $(15,15 / 4)$, but we did not verify that it was a global maximum. We have

$$
\begin{aligned}
& P_{x x}^{\prime \prime}(x, y)=-2, \\
& P_{y y}^{\prime \prime}(x, y)=-4
\end{aligned}
$$

and

$$
P_{x y}^{\prime \prime}(x, y)=0
$$

This means that

$$
D(x, y))=P_{x x}^{\prime \prime}(x, y) P_{y y}^{\prime \prime}(x, y)-\left(P_{x y}^{\prime \prime}(x, y)\right)^{2}=(-2) \cdot(-4)-0=8>0
$$

and

$$
P_{x x}^{\prime \prime}(x, y)=-2<0
$$

so $P$ is concave. Therefore, $(15,15 / 4)$ is a global maximum.


[^0]:    ${ }^{1}$ It is true for all functions you will see in this course. The requirement is that $f$ must have continuous second-order partial derivatives.

[^1]:    ${ }^{2}$ 'Sufficiently close' meaning all points within a radius $r>0$ to $\left(x_{*}, y_{*}\right)$, where you may set $r$ to be as small as you like.

