Chapter 2
Basic Mathematical Concepts and Methods

This chapter and the next two have three objectives. First, to introduce the reader to some basic concepts and formulas that will be needed in later chapters. Second, to serve as an introduction to computation and numerical methods and the use of Excel and Matlab procedures. The present chapter is devoted to mathematics and Chap. 3 is an introduction to computation and Chap. 4 will concentrate on probability theory and statistics. Those who are familiar with the material may want to glance through these chapters and move on. A third function of the chapters is to provide a handy reference for readers who, in reading later chapters, might feel a need to refresh their understanding of a concept or to check a formula.

2.1 Functions of Real Variables

In studying economic phenomena, we frequently come across cases in which variation in one variable induces variation in another. For example, an increase in income increases consumption, and an increase in price of a good or service reduces its demand. In other words, one variable, say y, depends on another, say x. Such dependencies are not confined to economics; they are observed in physical sciences and in everyday life. For example, the area of a circle, denoted by A, depends on its radius R, that is, $A = \pi R^2$. Similarly, the distance traveled by a car depends on the speed and time traveled.

If the relationship is such that every value of x leads to a unique value of y, then we can write $y = f(x)$ and say that y is a function of x. Note that the same y can be attached to more than one x, but that each x should be attached to only one y. Functions of real variables can be written as a mapping from the extended real line to itself. In other words, every real number in the domain corresponds to a unique real number in the range.

$$f : \Re \rightarrow \Re \tag{2.1}$$

Needless to say, a function need not be confined to one argument. We can write y as a function of x and z or as a function of $x_1, \ldots, x_k$. We can write them as

K. Dadkhah, *Foundations of Mathematical and Computational Economics,* 2nd ed.,
DOI 10.1007/978-3-642-13748-8_2,

$$f : \Re^2 \to \Re \qquad \text{or} \qquad y = f(x, z) \tag{2.2}$$

and

$$f : \Re^k \to \Re \qquad \text{or} \qquad y = f(x_1, \ldots, x_k) \tag{2.3}$$

We will encounter many kinds of functions in this chapter, and more functions yet throughout this book (examples of functions appear in Fig. 2.1a–c). Among them are polynomial functions, which are of the general form

$$y = \sum_{i=0}^{k} a_i x^i \tag{2.4}$$

Letting k =0, 1, 2, 3, we have

$y = a_0$	Constant function
$y = a_0 + a_1 x$	Linear function
$y = a_0 + a_1 x + a_2 x^2$	Quadratic function
$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$	Cubic function

Example 2.1 (Utility Function). The utility function is an important tool of economic analysis. But as a function, it has a special feature that we would like to emphasize. The function attaches a real number to any bundle of goods and services. For instance, if the amount of each good or service is denoted by $x_i, \; i = 1, \cdots, n$, then

$$U = U(x_1, \cdots, x_n)$$

This function is such that if a particular bundle, say, bundle a, is preferred to another bundle b, then the utility, U^{a}, attached to a, is a bigger number than the utility attached to b. That is, $U^a > U^b$. But the numbers themselves do not have any significance in the following sense. Suppose $U^a = 10$ and $U^b = 5$. Clearly, the bundle a is preferred to bundle b. But we could also assign $U^b = 9.5$ to the bundle b and it would make no difference, in the sense that it conveys the same information as $U^b = 5$. The only important consideration is that $U^a > U^b$. Because of the property just described, utility is an *ordinal* number and utility function is an ordinal function. An ordinal number is different from a *cardinal* one like the amount of income. If a person makes \$50,000 a year and another person \$25,000, then there is a \$25,000 difference between their incomes, and the first one makes twice as much as the second. But the difference between $U = 10$ and $U = 5$ does not convey any information, nor does it mean that one bundle is preferred twice as much as the other. Another important property of the utility function is that if we keep all x_i's constant and increase only one of them, then the new bundle is preferred to the old one. To put it simply, the utility function is based on the idea that more is preferred to less.

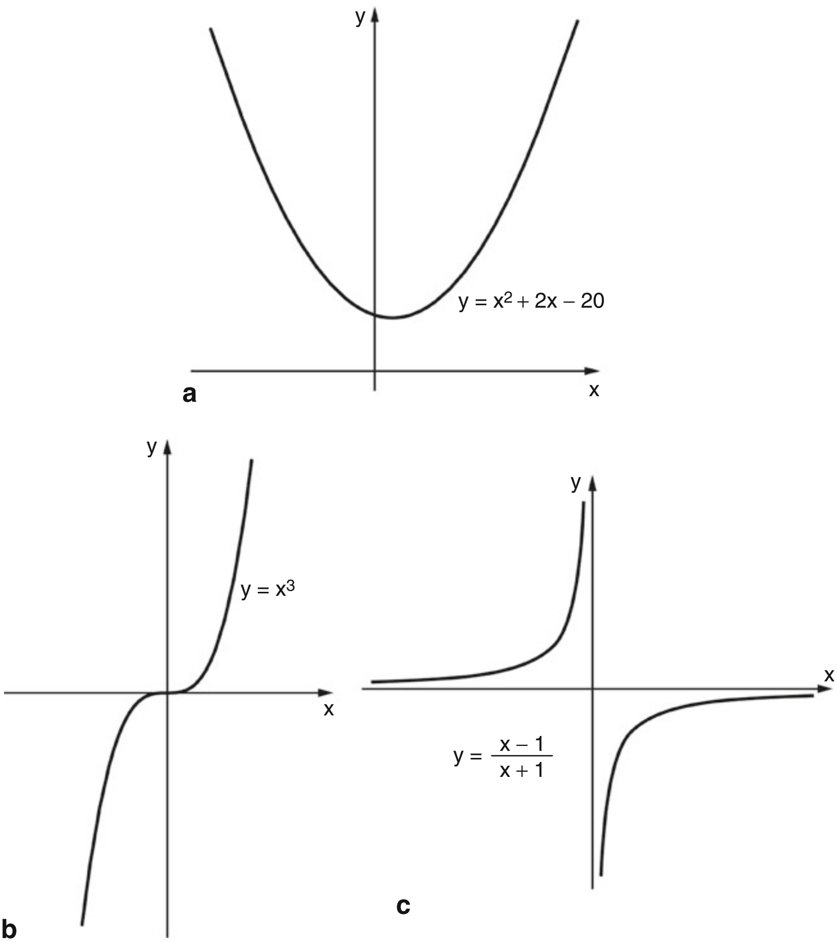

Fig. 2.1 (**a**) Quadratic function or parabola; (**b**) cubic function; (**c**) hyperbola

Example 2.2 (Cobb Douglas Production Function). A production function relates services of labor (L) and capital (K) to the maximum amount of output (Q) attainable from their combination. There are a number of production functions, which we shall discuss in Chap. 9. An important production function is the Cobb Douglas, which has the form

$$Q = AK^{\alpha}L^{\beta}$$

Graphing functions with Matlab is straightforward

Matlab code

```
% Define the domain of the function from -3 to 3 with
% increments of 0.1
x = -3:0.1:3;
% Define the function
y = x.^3;
% Plot the function
plot(x, y)
% You can plot more than one function on the
% same graph
hold on
for k=1:3
      y = x.^3 + 3*k;
      plot(x, y);
end
hold off
```

If you plot the function $y = x^3$, you will get a graph similar to Fig. 2.1b above. The same can be accomplished using Excel. To graph the function $y = x^3 + 3$, create the following on an Excel sheet. Highlight column B and use
>Insert → Line
You can use Select Data and use column A as the horizontal axis.

A	B
−3.0	=A1^3+3
−2.9	=A2^3+3
−2.8	=A3^3+3
...	...
2.9	=A60^3+3
3.0	=A61^3+3

Plotting three-dimensional graphs is slightly different. In Matlab use the following code:

Matlab code

```
% Define the domain of the function
[x, y] = meshgrid(-2:0.1:2,-2:.1:2);
% Define the function
z = x.*exp(-x.^2 - y.^2);
% Plot the function
mesh(x, y, z)
```

You will get the following graph (Fig. 2.2):
Similarly, we can plot the Cobb-Douglas production function.

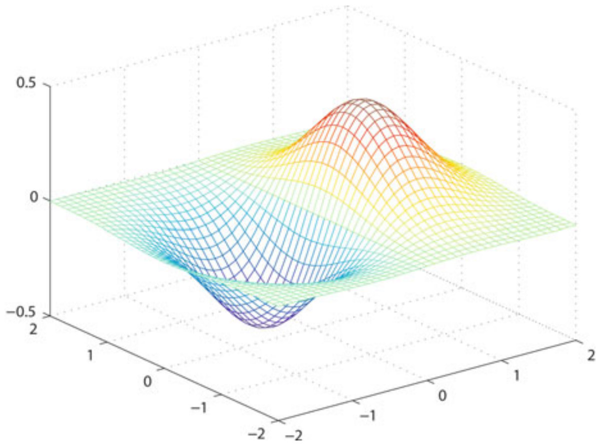

Fig. 2.2 Graph of the function $z = x\exp(-x^2 - y^2)$.

A	B	C	...	AQ
	−2.0	−1.9	...	2.0
−2.0	= A2*exp(– A2^2 – B 2^2)	= A2*exp(– A2^2 – C 2^2)	...	= A2*exp(– A2^2 – AQ 2^2)
−1.9	= A3*exp(– A2^2 – B 2^2)	= A3*exp(– A2^2 – C 2^2)	...	= A3*exp(– A2^2 – AQ 2^2)
−1.8	= A4*exp(– A2^2 – B 2^2)	= A4*exp(– A2^2 – C 2^2)	...	= A4*exp(– A2^2 – AQ 2^2)
...				
1.9	= A41*exp(– A2^2 – B 2^2)	= A41*exp(– A2^2 – C 2^2)	...	= A41*exp(– A2^2 – AQ 2^2)
2.0	= A42*exp(– A2^2 – B 2^2)	= A42*exp(– A2^2 – C 2^2)	...	= A42*exp(– A2^2 – AQ 2^2)

Matlab code

```
% Define the domain of the function
[K, L] = meshgrid(0:0.1:2, 0:0.1:4);
% Define the production function
Q = 5.*(K.^0.4).*(L.^0.6);
% Plot the function
mesh(K, L, Q)
```

Creating a three dimensional graph in Excel is a bit more time consuming and the result not as expressive as that of Matlab. Create the above worksheet in Excel. Highlight the square containing the computed numbers but not the values assigned to x and y. Then click Insert → Other Charts and choose one of the options.

2.1.1 Variety of Economic Relationships

In economics, we encounter three types of relationships:

1. Identities or definitions
2. Causal relationships
3. Equilibrium conditions

Identities or *definitions* are relationships that are true by definition or because we constructed them as such. Examples are the national income identity, in which GDP (Y) is defined as the sum of consumption (C), investment (I), government expenditures (G), and the difference between exports and imports, that is, net exports ($X - M$):

$$Y = C + I + G + X - M$$

Similarly, we define profit as revenues (quantity sold times price) net of cost:

$$\pi = PQ - C$$

Given their nature, such identities are not subject to empirical verification; they are always true. If we estimate the national income identity above using data from any country, we get an $R^2 = 1$ and coefficients that are highly significant and are usually 0.99999 (or -0.99999) and 1.00001 (or -1.00001). More important, since identities do not posit any hypothesis, no amount of algebraic manipulation of them will result in new insights into the workings of an economy.

Causal relationships are the mainstay of economics. They incorporate hypotheses regarding the behavior of economic agents, or technical and legal characteristics of the economy. Therefore, they are subject to empirical testing. Examples of behavioral relationships are consumption function, demand function, demand for imports, production functions, and tax revenues as a function of aggregate income.

By writing one variable as a function of a set of other variables, we implicitly declare that causation runs from the right-hand side (RHS) or explanatory variables to the left-hand side (LHS) or dependent variable. But how do we know this? How could we substantiate such a statement? Unlike physics and chemistry where experiments are the main source for accepting or rejecting a hypothesis, experiments play a very limited role in economics. Economics is an observational science.

Having been denied experiments and knowing well that correlation does not imply causation, econometricians have devised statistical tests of causality. The most widely used test of causality is due to Clive Granger (Nobel Laureate 2003). The test is for the necessary, but not sufficient, condition of the existence of causality in the strict sense. Thus, failing to reject the null hypothesis of no causality via the Granger test shows that x does not cause y in the strict sense. On the other hand, rejecting the null hypothesis establishes the necessary, but not sufficient, condition for causation. Most economic variables mutually affect each other. Money supply affects prices, which in turn affect the demand for money and indirectly the supply of money. Given that many economic variables are measured over arbitrary intervals of a month, quarter, or year, we may observe the mutual causation in the form of simultaneity. Of course, we may also observe simultaneity among economic variables because they are simultaneously determined through interdependent processes.

Equilibrium conditions describe the situation or condition when two or several variables are in such configuration that they need not change. Unlike identities, equilibrium conditions do not always hold. On the other hand, equilibrium conditions differ from causal relations in that a change in one variable does not automatically bring a change in another. Only if equilibrium is restored would a change in one variable bring about a change in the other. An equilibrium condition, if stable, implies that any deviation from equilibrium sets in motion forces that will bring back equilibrium. Therefore, stable equilibrium conditions are subject to statistical testing. Such tests, referred to as tests of cointegration, were proposed by Robert Engle (Noble Laureate 2003) and Granger, and by Søren Johansen. Furthermore, because there must be a force to restore the equilibrium, cointegration implies an error-correction mechanism. Thus, a delayed causation arises through error correction if the equilibrium is stable. We will encounter all these types of economic relationships in this book, and the reader will get a better sense of them after working with several specific examples.

2.1.2 Exercises

E.2.1. Graph the following functions for $-5 < x < 5$.

$$\begin{aligned} &i.\ \ y = 10 + 2x \\ &ii.\ \ y = 5 + 2x + 3x^2 \\ &iii.\ \ y = 7x^3 - 14x + 5 \\ &iv.\ \ y = \frac{1-x}{1+x} \end{aligned}$$

E.2.2 Make a list of economic relationships that you recall from economics courses and classify them as identities, causal relationships, and equilibrium conditions.

2.2 Series

The sequence of numbers

$$x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_{n-1} \quad x_n \tag{2.5}$$

is called a series.

Example 2.3 The following are examples of series

$$\begin{array}{llllllll} i. & 1 & 2 & 3 & \ldots & n & \\ ii. & 2 & 2^2 & 2^3 & \ldots & 2^n & \\ iii. & \frac{1}{2} & \frac{1}{2^2} & \frac{1}{2^3} & \ldots & \frac{1}{2^{n-1}} & \frac{1}{2^n} \end{array}$$

Two issues are of importance here. First, could we write a series in a more compact format instead of enumerating its members? This can be done in two ways: by writing a general expression for its n-th term or by writing its recurrence relation. For instance, the n-th terms of series in Example 2.3 can be written as

$$i. \quad x_n = n \qquad ii. \quad x_n = 2^n \qquad iii. \quad x_n = \frac{1}{2^n}$$

Not all series can be written in this format. An alternative is to write their recurrence relation. For the above series the recurrence relations are

$$\begin{array}{ll} i.\ x_1 = 1, & x_n = x_{n-1} + 1 \\ ii.\ x_1 = 2, & x_n = 2x_{n-1} \\ iii.\ x_1 = \frac{1}{2}, & x_n = \frac{1}{2}x_{n-1} \end{array}$$

Example 2.4 Consider the Fibonacci sequence

$$1 \quad 1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \quad 21 \quad \ldots$$

It starts with 1 and 1 and then each term is the sum of its two previous numbers. Thus, the recurrence relation is

$$x_0 = x_1 = 1 \qquad \text{and} \qquad x_n = x_{n-1} + x_{n-2} \qquad n = 2, 3, \ldots,$$

We cannot always find recurrence relations for a series. For example, if the series is the realization of a random variable, we would not be able to find such a formula.

The second question is whether the sum of a series exists and if so, how we could calculate it. Note that mathematically speaking, when we say something exists, we mean that the entity in question has a finite value. Thus, here the question is whether the sum of a series is finite or tends to infinity. Before discussing these questions, however, we need to learn about the *summation notation* Σ and the concept of *limit*.

2.2.1 Summation Notation Σ

We are all familiar with summing a set of specific numbers. But suppose we would like to talk of the sum of $x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7$, and x_8. Of course, we can always write it as

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8$$

But such a formula is cumbersome and inefficient. It is even more cumbersome when we have 20 or 100 values to add. Even worse is when we want to represent the sum of an infinite series of numbers. We make the following convention:

$$\sum_{i=1}^{n} x_i \tag{2.6}$$

by which we mean the sum of numbers x_1 to x_n inclusive. Note that i is simply a counter and can easily be exchanged with j or k or any other symbol, although by usage, i, j and k are the most commonly used letters for counters. Other examples of summation are

$$\sum_{i=0}^{\infty} x_i, \qquad \sum_{j=-n}^{n} x_j, \qquad \sum_{t=1}^{T} x_{2t}$$

A few properties of sums should be noted:

$$\sum_{i=1}^{n} a = (\underbrace{a + a + \ldots + a}_{n}) = na \qquad \text{where } a \text{ is a constant} \tag{2.7}$$

$$\begin{aligned} \sum_{i=1}^{n} ax_i &= ax_1 + ax_2 + \ldots + ax_n = a(x_1 + x_2 + \ldots + x_n) \\ &= a \sum_{i=1}^{n} x_i \end{aligned} \tag{2.8}$$

$$\begin{aligned} \left(\sum_{i=1}^{n} x_i\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j = \sum_{i=1}^{n} x_i^2 + \sum_{i\neq j}^{n}\sum_{j=1}^{n} x_i x_j \\ &= \sum_{i=1}^{n} x_i^2 + 2\sum_{i<j}^{n}\sum_{j=2}^{n} x_i x_j \end{aligned} \tag{2.9}$$

2.2.2 Limit

Consider the series (*ii*) in Example 2.3. As n increases, the last term of the series gets increasingly large. As n tends to ∞, so does the last term of the series. In such cases we say that the series has no limit. Note that ∞ is not a number. On the other hand, as n increases, the last term in (*iii*) in the same example becomes smaller and smaller as depicted in Table 2.1:

It can be seen that as n increases, $1/2^n$ tends to zero and, for all practical purposes, we can take it to be zero. In such cases, we say that the limit of the series exists and as n tends to ∞, $1/2^n$ tends to zero, and we write

$$\lim_{n\to\infty} \frac{1}{2^n} = 0 \tag{2.10}$$

Table 2.1 Approaching limit

N	$1/2^n$
1	0.5
2	0.25
...	...
10	0.0009765
11	0.0004882
...	...
100	7.8886×10^{-31}
101	3.9443×10^{-31}
...	...

Note that the limit needs not always be zero. It can be any number $L < \infty$. Now that we have an intuitive notion of a limit, let us present a formal definition.

Definition 2.1 Let $x_1, x_2, x_3, \ldots$ be a sequence of points on the real line. L is called the limit of this sequence if, for any number $\varepsilon > 0$, we could find a number N such that $|x_n - L| < \varepsilon$ if $n > N$.

If we apply the above definition to series (*ii*) in Example 2.3, we can reason that the series does not have a limit, because no matter what values we choose for L and N and no matter how large or small ϵ is, we cannot have $|2^n - L| < \varepsilon$ for all $n > N$. The reason: As n gets larger, so does 2^{n} and there is no limit to how large it can get. For the series (*iii*) the story is different. Let $L = 0$ and set $\varepsilon = 0.001$, then for all $n > 9$ we have $1/2^n < 0.001$. For example, $1/2^{10} = 0.0009765$. We can set $\varepsilon = 10^{-33}$, that is the decimal point followed by 32 zeros and then one. Now for all $n > 109$, we have $1/2^n < 10^{-33}$. The following properties of limits will prove quite useful.

Property 2.1 Let x_n and y_n represent two series and assume that both

$$\lim_{n\to\infty} x_n \qquad \text{and} \qquad \lim_{n\to\infty} y_n$$

exist. Then

$$\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n \tag{2.11}$$

The proposition is also true for the sum of any finite number of series. If c is a constant, it follows from (2.11) that

$$\lim_{n\to\infty}(c + x_n) = c + \lim_{n\to\infty} x_n \tag{2.12}$$

and

$$\lim_{n\to\infty} cx_n = c \lim_{n\to\infty} x_n \tag{2.13}$$

It is evident that the limit of the series

$$c \quad c \quad c \quad \ldots$$

is c.

Property 2.2 Let the series x_n and y_n be as in Property 2.1, then

$$\lim_{n\to\infty} x_n y_n = \lim_{n\to\infty} x_n \lim_{n\to\infty} y_n \tag{2.14}$$

Also

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n} \tag{2.15}$$

provided $\lim\limits_{n\to\infty} y_n \neq 0$.

2.2.3 Convergent and Divergent Series

Consider the sum of the first n terms of a series

$$S_n = \sum_{i=1}^{n} x_i \tag{2.16}$$

Clearly, for every value of n we have a different sum. These sums, referred to as partial sums, form a series themselves. The question is whether the sum S_n exists as $n \to \infty$. In other words, is the following statement true?

$$S = \lim_{n\to\infty} S_n < \infty \tag{2.17}$$

The answer is that the sum exists if

$$\lim_{n\to\infty} x_n = 0 \tag{2.18}$$

If S exists, then the series is called *convergent*, or else it is called *divergent*.

Example 2.5 The sum

$$S = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{2^i}$$

exists because $\lim\limits_{n\to\infty} (1/2^n) = 0$. Later in this chapter we will show how such sums can be calculated.

Example 2.6 Is the sum

$$\frac{1}{3} + \frac{2}{5} + \frac{3}{7} + \ldots + \frac{n}{2n+1} + \cdots$$

convergent or divergent? Because

$$\lim_{n\to\infty} \frac{n}{2n+1} = \lim_{n\to\infty} \frac{1}{2+1/n} = \frac{1}{2} \neq 0$$

we conclude that the series is divergent.

An alternative way of determining if a series is convergent or divergent when all terms are positive is the d'Alembert[1] test.

Property 2.3 (d'Alembert test). The sequence of positive numbers

$$x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_n \quad \ldots \tag{2.19}$$

is convergent and the limit

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} x_n \tag{2.20}$$

exists, if

$$\lim_{n\to\infty} \frac{x_{n+1}}{x_n} < 1 \tag{2.21}$$

If the above limit is greater than one, then the series is divergent. The case of the limit being equal to one is indeterminate.

Let us apply this test to some of the series we have encountered in this section. Note that all terms in these series are positive.

Example 2.7 The series in (*ii*) in Example 2.3 is divergent because:

$$\lim_{n\to\infty} \frac{2^{n+1}}{2^n} = 2 > 1$$

---

[1] Probably the most dramatic event in the life of the French mathematician Jean Le Rond d'Alembert (1717–1783) was that as a newborn he was left on the steps of a church. He was found and taken to a home for homeless children. Later, his father found him and provided for his son's living and education. D'Alembert made contributions to mathematics, mechanics, and mathematical physics. The eighteenth century was the age of European enlightenment and nothing represented the spirit of that age better than the *Encyclopédistes*, a group of intellectuals gathered around Diderot including Voltaire, Condorcet, and d'Alembert. They published the 28-volume *Encyclopedia* that contained articles on all areas of human knowledge including political economy.

But, the series in (*iii*) is convergent because

$$\lim_{n\to\infty} \frac{1/2^{n+1}}{1/2^n} = \frac{1}{2} < 1$$

For the series in Example 2.6 we have:

$$\lim_{n\to\infty} \frac{\dfrac{n+1}{2(n+1)+1}}{\dfrac{n}{2n+1}} = 1$$

Thus, in this case the d'Alembert test cannot resolve the issue.

In the following two subsections we will discuss two examples of series: arithmetic and geometric progressions.

2.2.4 Arithmetic Progression

The series

$$a \quad a+d \quad a+2d \quad a+3d \quad \ldots \quad a+(n-1)d \tag{2.22}$$

is called *arithmetic progression*. We can write it more compactly as

$$x_n = a+(n-1)d \qquad n = 1, 2, \ldots \tag{2.23}$$

or

$$x_1 = a \qquad x_n = x_{n-1} + d \qquad n = 2, 3, \ldots \tag{2.24}$$

Thus, every member of the series is equal to its predecessor plus a constant number.

Example 2.8 The following are arithmetic series:

$$\begin{array}{ll} i. & 1 \;\; 2 \;\; 3 \;\; 4 \;\; \ldots \;\; 20 \\ ii. & 5 \;\; 8 \;\; 11 \;\; 14 \;\; \ldots \end{array}$$

To calculate the sum of arithmetic series in (*i*), above, we can write

$$\begin{array}{l} S \;\, = 1 \;\;\, + \;\; 2 + \;\; 3 + \cdots + 20 \\ \underline{S \;\, = 20 + 19 + 18 + \cdots + \;\; 1} \\ 2S = 21 + 21 + 21 + \cdots + 21 \end{array}$$

Thus,

$$2S = 20 \times 21$$

and

$$S = \frac{20 \times 21}{2} = 210$$

This can be generalized to the sum of any n consecutive integers starting with 1.

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2} \tag{2.25}$$

Following the same line of reasoning for the general case, the sum of n consecutive terms in an arithmetic progression is,

$$S = n\left[a + \frac{(n-1)d}{2}\right] \tag{2.26}$$

Example 2.9 For the sum of the first 20 integers, we have $a = d = 1$ and $n = 20$. Plugging the numbers into (2.26), we get the sum of 210.

Example 2.10 For the sum of the first 10 integers divisible by 3, we have $a = d = 3$ and $n = 10$. Plugging the numbers into (2.26), we get the sum of 165.

These formulas can be programmed in Matlab in two ways. First, we can simply write a procedure that adds up, one by one, the n terms in a particular series. Alternatively, we can use (2.26) to evaluate the sum of the series.

Matlab code

```
% Initialize n, a, d, and S
n = 20;
a = 1;
d = 3;
S = 0;
% Compute S by adding the 20 terms
for i = 1:n
      S = S + a + (i-1)*d;
end
% Call S
S
% Alternatively you can write
S = n*(a + (n-1)*d/2)
```

Note that you can change n, a, and d to any number and run the procedure again and again.

The same can be accomplished using Excel. By replacing n, a, and d with the desired values, you will get the sum of the arithmetic series in two different ways.

Note that in the illustration below and in subsequent Excel illustrations we make reference to cell numbers. In the next chapter we learn how to name variable in Excel and to refer to them by name.

A	B	C	D	E
=C1		1	3	20
=A1+D$1				
=A2+D$1		=E1*(C1+(E1−1)*D1/2)		
⋮				
=SUM(A1:A20)				

2.2.5 Geometric Progression

The series

$$a \quad aq \quad aq^2 \quad aq^3 \quad \ldots \quad aq^{n-1} \quad \ldots \tag{2.27}$$

is called a geometric progression. The recurrence relation is

$$x_1 = a \qquad x_n = qx_{n-1} \qquad n = 2, 3, \ldots \tag{2.28}$$

We are interested in finding the sum of the first n terms of this series. Let

$$S = \sum_{i=0}^{n-1} aq^i = a + aq + aq^2 + aq^3 + \ldots + aq^{n-1} \tag{2.29}$$

Multiplying S by q and subtracting it from S, we have

$$\begin{array}{rl} S = & a + aq + aq^2 + aq^3 + \cdots + aq^{n-1} \\ -Sq = & \quad\;\, -aq - aq^2 - aq^3 - \cdots - aq^{n-1} - aq^n \\ \hline S - Sq = & a - aq^n \end{array}$$

Thus,

$$S = a\frac{1 - q^n}{1 - q} \tag{2.30}$$

Example 2.11 Find the following sum:

$$S = 2 + 6 + 18 + 54 + 162 + 486 + 1458$$

Because $a = 2, q = 3$, and $n = 7$, we have

$$S = 2\frac{1 - 3^7}{1 - 3} = 2186$$

Geometric progression finds a few applications in macroeconomics including aggregate demand multiplier, money multiplier, and present value.

Example 2.12 (Keynesian multiplier) When discussing the effect of an increase in government expenditures on aggregate demand and income, the following argument is offered. Suppose the government increases its expenditures by \$100 billion. These additional expenditures by the government will become the income of individuals who provide the goods and services to the government. Assuming a marginal propensity to consume of 0.92, the additional consumption will be \$92 billion. This consumption, in turn, forms the income of those who produce consumer goods and services. But then they will spend 0.92×92 or \$84.64 billion on consumption which in turn will be the income of those who produce consumption goods and services. You get the idea. The stream of income generated in different stages is shown in Table 2.2:

The sum of the first 20 terms of the addition to national income can be calculated as

$$S = 100\frac{1 - 0.92^{20}}{1 - 0.92} = 1014.13$$

If we repeat the same calculation for the first 40 terms, we get a total of \$1205 billion. The second 20 terms add less than a quarter of the first 20. The sum of the first 100 terms equals 1249.7. The reason: $0.92 < 1$ and when a number whose absolute value is less than one is raised to increasing exponents, it becomes smaller and smaller. The smaller the absolute value of the number, the sooner it reaches zero. For example, if the marginal propensity to consume was 0.5 instead of 0.92, the sum of the first 44 terms would be \$250 billion and additional terms would have no effect. Indeed, terms beyond the first 20 would have no practical significance. Thus, if we allow the process to continue indefinitely, that is, letting $n \to \infty$, we will have

$$S = 100\frac{1}{1 - 0.92} = 1250$$

Note 0.92 is marginal propensity to consume, and $1/(1-0.92)$ is our good old multiplier.

Table 2.2 Multiplier effect at work: the effect of government expenditures on income

Steps	Increase in income	
0	100	$= 100$
1	92	$= 100 \times 0.92$
2	84.64	$= 100 \times 0.92^2$
3	77.8688	$= 100 \times 0.92^3$
...	...	
...	...	

We can generalize the results of the last example by noting that

$$\lim_{n\to\infty} q^n = 0 \qquad \text{if} \qquad |q| < 1 \tag{2.31}$$

It follows that

$$\lim_{n\to\infty} S = \lim_{n\to\infty} a\frac{1-q^n}{1-q} = a\frac{1}{1-q} \qquad \text{if} \qquad |q| < 1 \tag{2.32}$$

As in arithmetic progression we can use Matlab to carry out the necessary calculations:

Matlab code

```
% Initialize n, a, q, and S
n = 7;
a = 2;
q = 3;
S = 0;
% Compute S by adding the 20 terms
for i = 1:n
      S = S + a.*q.^(i-1);
end
% Call S
S
% Alternatively you can write
S = a.*(1-q.^n)./(1-q)
```

If you use the second method, you may want to define a function and call it when needed. First you create an M-file in Matlab containing the function.

Matlab code

```
function G = Geoprog(v);
n = v(1);
a = v(2);
q = v(3);
G = a.*(1-q.^n)./(1-q);
```

Then you can call this function for different values of n, a, and q.

Matlab code

```
v = [7 2 3];
S = Geoprog(v);
```

We can perform these computations in Excel as illustrated below:

A	B	C	D	E
=C1		2	3	7
=A1*D$1				
=A2*D$1		=C1*(1−D1^E1)/(1−D1)		
⋮				
=SUM(A1:A7)				

2.2.6 Exercises

E.2.3 Find the sum of all odd numbers from 1 to 451.

E.2.4 Find the sum of all even numbers from 2 to 450.

E.2.5 Find the sum of the following geometric series:

$$1 \quad \frac{1}{2} \quad \frac{1}{4} \quad \frac{1}{8} \quad \dots$$

$$1 \quad \frac{1}{3} \quad \frac{1}{9} \quad \frac{1}{27} \quad \dots$$

E.2.6 The *present value* (PV) of a stream of income is $D_t,\ t = 0, 1, \dots, T$ is defined as

$$\mathrm{PV} = \sum_{t=0}^{T} \frac{D_t}{(1+r)^t}$$

where $t = 0$ is the current year and r is the rate of interest.

i. Compute the present value of a winning lottery ticket that will pay \$200,000 per year for 20 years starting in the present year. Assume an interest rate of 12%. Solve the same problem assuming interest rates of 15% and 20%. [*Hint*: For interest rate of 12%, $r = 0.12$.]
ii. Compute the value of a government bond that pays one dollar every year in perpetuity (i.e., forever) given the interest rate of r.

E.2.7 Show that

$$\sum_{i=0}^{\infty} (i+1)\lambda^i = \frac{1}{(1-\lambda)^2}, \qquad |\lambda| < 1$$

[*Hint*: $\lim_{n\to\infty} n\lambda^n = 0$.]

2.3 Permutations, Factorial, Combinations, and the Binomial Expansion

Counting rules discussed in this section are the elementary building blocks of combinatorics, a branch of mathematics that has applications in many areas including cryptography, computer science, probability theory, statistics, econometrics, and economics. Consider a collection of n items denoted by $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$. Suppose we choose $r \leq n$ items from $\mathbf{A}$ and arrange them in the order they are chosen. A typical arrangement will look like

$$\overbrace{A_3, A_7, \ldots, A_{r+2}}^{r}$$

How many such collections can we form that are different from each other at least in one item or in the position of one item? We can argue as follows. For the first item, we can choose from n items; for the second place, from among the remaining $n - 1$ items, because one item has already been taken for the first place. Continuing in this way, for the r-th item we can choose from among the remaining $n - (r-1)$ items. Thus, the total possible arrangements are

$$n \times (n-1) \times \ldots \times (n-r+1) \tag{2.33}$$

For example if we have five objects, we can make $5 \times 4 \times 3 = 60$ different arrangements containing three elements. If we allow $r = n$, then we have

$$1 \times 2 \times 3 \times \ldots \times n = \prod_{i=1}^{n} i = n! \tag{2.34}$$

$n!$ is called "n factorial," and its meaning is quite obvious. Π is similar to the summation notation, except that it stands for the product of a set of numbers or variables. Note that $0!=1$. The reason: We can arrange or permute in only one way the elements of a null set (the set with zero elements). Using the convention of (2.34) we can write (2.33) as

$$\frac{n!}{(n-r)!} \tag{2.35}$$

Now suppose we ask, in how many ways can we pick r elements from the set containing n elements? The number of combinations is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{2.36}$$

Example 2.13 Suppose five soccer teams are playing in a tournament. How many games will be played? Let us designate our teams by letters A, B, C, D, and E. Here is the list of games to be played:

AB, AC, AD, AE, BC, BD, BE, CD, CE, DE

which makes a total of 10 games. Note that we do not have both AB and BA because when A has played against B, the reverse is also true. The problem is the same as choosing two out of a set of five. Based on (2.36) we have,

$$\binom{5}{2} = \frac{5!}{2!3!} = \frac{5 \times 4}{2} = 10$$

Example 2.14 A mutual fund is a portfolio consisting of a number of equities held in different proportions. For example, it may have 5% of its assets in IBM stock, 6% in Verizon, 10% in Microsoft, and so on. Assume that 1000 stocks in the market are deemed to be appropriate for inclusion in such funds. Further suppose that each fund consists of 30 stocks. How many different portfolios can one form from the 1000 stocks?

$$\binom{1000}{30} = \frac{1000!}{30!970!} = 242960819217375 \times 10^{43}$$

A very large number indeed. As can be seen, the precision of these numbers is 15 digits; that is, the first 15 digits are accurate and the rest give the order of magnitude. What is interesting is that the number of potential mutual funds far exceeds the number of stocks. Note that a mutual fund needs not consist of exactly 30 stocks; it can have 40, 50, 100, 200, or any other number of stocks. For each of those numbers, a large number of funds could be formed. Thus, the total number of potential mutual funds is astronomical.

Two functions in Matlab allow calculations of $n!$ and $\binom{n}{r}$.

Matlab code

```
% for n!
factorial(n)
% for
nchoosek(n,r)
```

Excel has a function for factorial and for combination. In Formulas choose Insert Function and then choose Math & Trig, Finally choose FACT for factorial and COMBIN for combination. Alternatively you could type in:

=FACT(n)	=COMBIN(n,r)

Combinations prove useful in writing the binomial expansion.

$$(a+b)^n = \binom{n}{0} a^n + \binom{n}{1} a^{n-1}b + \binom{n}{2} a^{n-2}b^2 + \cdots + \binom{n}{n-1} ab^{n-1} + \binom{n}{n} b^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i}b^i \tag{2.37}$$

Example 2.15 We can illustrate the general formula in (2.37) by applying it to $n = 2, 3, 4$.

$$(a+b)^2 = \binom{2}{0} a^2 + \binom{2}{1} ab + \binom{2}{2} b^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = \binom{3}{0} a^3 + \binom{3}{1} a^2b + \binom{3}{2} ab^2 + \binom{3}{3} b^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = \binom{4}{0} a^4 + \binom{4}{1} a^3b + \binom{4}{2} a^2b^2 + \binom{4}{3} ab^3 + \binom{4}{4} b^4$$
$$= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

2.3.1 Exercises

E.2.8 Evaluate the following expressions using a calculator, the Excel function `COMBIN`, and the Matlab function `nchoosek`.

$$\binom{14}{3}, \quad \binom{9}{6}, \quad \binom{23}{8}, \quad \binom{33}{12}$$

E.2.9 There are 50 delegates at a convention, 32 men and 18 women. In how many ways can we choose a committee of eight equally divided between men and women?

E.2.10 There are 9 potential judges for a contest including 5 women and four men. In how many ways can we choose 5 judges provided that at least two of them are women?

E.2.11 Show that

$$\sum_{j=0}^{n} \binom{n}{j} = \binom{n}{0} + \ldots + \binom{n}{n} = 2^n$$

[*Hint*: Consider the binomial expansion of $(1+1)^n$.]

2.4 Logarithm and Exponential Functions

Most likely you are already familiar with logarithm and exponential functions, because both play important roles in mathematics. In addition, they find many uses in economics, especially in dynamic models and growth theory. Many econometric models involve logarithms of both dependent (endogenous) and explanatory (exogenous) variables.

2.4.1 Logarithm

Suppose

$$y = a^x \qquad\qquad y > 0, \quad a > 1 \tag{2.38}$$

then x is the logarithm[2] of y in the base a, which we denote as

$$x = \log_a y \tag{2.39}$$

Note that both a and y are positive real numbers. Logarithms of negative numbers are complex numbers. In this book we confine ourselves to logarithm of positive numbers. Whereas a could be any positive real number, the three important bases are 2, 10, and e. Base 2 is used in information science and communication. Base 10 is convenient for certain calculations; note that the logarithm of 1, 10, 100, 1000, ... in base 10 are 0, 1, 2, 3, ...

The base we will be dealing with in this book is e, an irrational number approximately equal to 2.7182818285. This unusual number somewhat like π, will prove quite useful and will play a significant role in mathematics and computation. In the next section, we have more to say about e, but for the time being consider it a number. The logarithm in base e is referred to as the natural logarithm and sometimes (to avoid confusion) is denoted by ln—a practice we will adopt in this book.

A basic property of logarithm that makes manipulation and calculations easier is that

$$\ln(xy) = \ln x + \ln y \tag{2.40}$$

Let

$$x = e^{\alpha}, \qquad y = e^{\beta}$$

then

[2] John Napier (1550–1617), a Scottish nobleman, conceived the idea of the logarithm. The first tables using base 10 were calculated by Henry Briggs (1561–1631), a professor of geometry at Gresham College.

$$xy = e^{\alpha} e^{\beta} = e^{\alpha+\beta}$$

and

$$\ln(xy) = \alpha + \beta = \ln x + \ln y$$

Repeated application of (2.40) results in

$$\ln\left(x^n\right) = n \ln x \tag{2.41}$$

Combining (2.40) and (2.41), we have

$$\ln\left(\frac{x}{y}\right) = \ln(xy^{-1}) = \ln x + \ln(y^{-1}) = \ln x - \ln y \tag{2.42}$$

Thus, logarithm turns multiplication into addition, division into subtraction, raising to a power into multiplication, and finding the roots of a number into division.

Logarithmic functions are programmed in every calculator and in software such as Excel. In Matlab one can get the logarithm of a positive number in three bases.

Matlab code

```
% Natural logarithm
log(x)
% In base 10
log10(x)
% In base 2
log2(x)
```

Excel has three functions for logarithm: LN for the natural logarithm, LOG10 for the base 10, and LOG for any base the user specifies.

You will hardly ever need the logarithm of a number in any other base, but should such a need arise, the calculation is simple. Suppose you are interested in finding the logarithm of y in the arbitrary base of $b > 1$. Let x, and z be, respectively, logarithms of y in bases e and b. We can write

$$y = e^x = b^z$$

and

$$\ln y = x = z \ln b$$

Therefore,

$$\log_b y = z = \frac{x}{\ln b} = \frac{\ln y}{\ln b} \tag{2.43}$$

Example 2.16

$$\ln 45 = 3.8066625 \quad \ln 10 = 2.3025851$$

$$\log_{10} 45 = \frac{\ln 45}{\ln 10} = \frac{3.8066625}{2.3025851} = 1.6532125$$

Example 2.17

$$\log_2 1024 = \frac{\ln 1024}{\ln 2} = \frac{6.9314718}{0.6931471} = 10$$

2.4.2 Base of Natural Logarithm, e

Next to π, the base of natural logarithm, e, is the most famous irrational number among mathematicians and those who apply mathematics. It is approximately equal to 2.71828182845905 and more precisely

$$e = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x \tag{2.44}$$

We need not dwell on the origin and the logic behind this number. Rather, we can gain an intuitive understanding of it through an example from economics. Suppose you deposit \$1000 in an interest-bearing account with an interest rate of 12%. After a year, your money would be \$1000 (1+0.12) = \$1120. But the underlying assumption in this calculation is that the interest accrues to your money at the end of the year. Why should it be that way? Suppose at the end of 6 months you receive half of the annual interest and increase your account to \$1060. For the next 6 months you earn interest on this new amount and, at the end of the year, your balance would be

$$\$1000\left(1+\frac{0.12}{2}\right)^2 = \$1123.60$$

Why should we stop there? Why not ask for the interest to accrue every season, every month, or even instantaneously? Table 2.3 shows the amount of principal plus interest when interest accrues at different frequencies. The last amount is approximately equal to $\$1000 \times e^{0.12} = \$\ 1127.50$. This is the amount you would have had if interest accrued every second. As a matter of notation, sometimes e^x is written as $\exp(x)$. Matlab has a ready-made function for $\exp(x)$.

Matlab code

```
% Exponential function
```

Table 2.3 Effect of the frequency of interest accrual on the total amount of interest

Frequency	Total interest
Annual	$1000(1+0.12)=1120.00$
Semiannual	$1000\left(1+\frac{0.12}{2}\right)^{2}=1123.60$
Seasonal	$1000\left(1+\frac{0.12}{4}\right)^{4}=1125.51$
Monthly	$1000\left(1+\frac{0.12}{12}\right)^{12}=1126.83$
Daily	$1000\left(1+\frac{0.12}{365}\right)^{365}=1127.47$

```
exp(x)
```

The Excel function for exponential is EXP.

2.4.3 Exercises

E.2.12 Graph the following functions for $0.1 < x < 6$:

$$y = \ln(x) \qquad y = e^{x} \qquad y = e^{-2x}$$

E.2.13 For the annual interest rates 20, 18,15, 12, 10, 8, 5, and 2%,

i. Compute the corresponding daily rates.
ii. Compute the corresponding effective annual rates if the interest is compounded daily.
iii. Compute the corresponding effective annual rates if the interest is compounded instantaneously.
iv. How close are the results in *ii* and *iii*?

E.2.14 Given the following equations, find x and y.

$$\begin{aligned} y^{x} &= x^{y} \\ y &= 2x \\ x &> 0 \end{aligned}$$

2.5 Mathematical Proof

In many math books proofs of theorems end with the abbreviation QED that stands for the Latin phrase "Quod Erat Demonstrandum," meaning "which was to be shown." But what is to be shown and what do we mean by a mathematical proof?

2.5.1 Deduction, Mathematical Induction, and Proof by Contradiction

Mathematicians prove their propositions in one of three ways: deduction or direct proof, mathematical induction, or by contradiction. As we mentioned in Chap. 1, mathematical propositions are tautologies, although the connection between the assumptions (the starting point) and proposition (the end point) may not be easy to see. The goal of Mathematics is to find and substantiate such connections. The genius of a great mathematician is in discerning an important proposition and in proving how it can be derived from a minimal set of assumptions. On many occasions it is easier to start from a proposition and work backward. Other times, mathematicians must refine the assumptions or add to or subtract from them. In still other cases, the proposition may need adjustment. Once a proposition is proved, others may find easier proofs, discover that the proposition needs less strict assumptions, or that the proposition is simply a special case of a more general theorem. Finding implications of a general proposition, finding interesting applications and special cases for it, and discovering its connections to other propositions provide avenues for further research.

Proof by Deduction. Direct proofs or deductions start with assumptions and lead to the proposition. We have to show that every statement follows logically from the previous one. In other words, we have to show that each step is implied by what we knew in the previous step. In this process we can use any theorem or lemma that has already been proved because by having proved them, we know they are logically correct. Our derivation of the formula for the sums of arithmetic and geometric series, although elementary, are examples of direct proof. Similarly, many propositions you remember from high school geometry are proved by direct reasoning.

Proof by Induction. Another way of proving a proposition is by induction, in which we first prove the validity of a proposition for the case of $n = 1$; then assuming that the proposition is true for the case of $n - 1$, we show that it is also true for n. Since we already know that the theorem or lemma is true for $n = 1$, then it should be true for $n = 2$, and therefore, $n = 3$, and indeed for any n.

Example 2.18 We have already seen that the sum of integers from 1 to n is equal to $n(n+1)/2$. We can verify that this formula is correct for $n = 1$ and indeed for $n = 2$ and $n = 3$. Suppose we know that the formula is true for the sum of $n - 1$ numbers, that is,

$$S_{n-1} = \frac{(n-1)n}{2}$$

Then for the sum of n consecutive integers we have

$$S_n = S_{n-1} + n = \frac{(n-1)n + 2n}{2} = \frac{n(n+1)}{2}$$

Example 2.19 Similarly we showed that the sum of n terms of geometric progression is

$$S = a\frac{1-q^n}{1-q}$$

We can verify that this sum is correct for the first term, and the sum of the first two terms. Now let us assume that the formula is correct for the first $n-1$ terms. Then

$$\begin{aligned} S_n &= S_{n-1} + aq^{n-1} \\ &= a\tfrac{1-q^{n-1}}{1-q} + aq^{n-1} \\ &= a\tfrac{1-q^{n-1}+q^{n-1}-q^n}{1-q} \\ &= a\tfrac{1-q^n}{1-q} \end{aligned}$$

Proof by Contradiction. In proving a proposition by contradiction, we first assume that the proposition is false. Then, deriving the implications of the proposition being false, we show that they contradict some proven theorems or known facts. The conclusion is that the proposition cannot be false.

Example 2.20 One of Euclid's theorems states that the number of primes is infinite. Recall that a prime number is divisible only by one and itself. To prove the theorem, we assume the contrary, that prime numbers are finite. Therefore, we can write them as

$$p_1\, p_2\, p_3 \ldots\, p_n$$

But now consider

$$p_{n+1} = p_1 \times p_2 \times p_3 \times \ldots \times p_n + 1$$

This number is not divisible by other primes because the division will have a remainder of one, therefore, it is a prime. Thus, no matter how many numbers we have, we can add one more and then another. This contradicts the assumption that there are only a finite number of primes.[3]

[3] Alternatively, the fact that p_{n+1} is not divisible by any prime contradicts the fundamental theorem of arithmetic that states that any integer $k > 1$ has a unique factorization of the form

2.5.2 Computer-Assisted Mathematical Proof

Proofs of mathematical propositions rely on logical steps that are convincing, if not to all ordinary mortals, at least to all who have the proper training. Moreover, such proofs are general in the sense that they apply to a class of problems. For instance, the proof of the solution of the quadratic equation does not rely on any particular values of the coefficients. Rather it is correct for all equations with real coefficients. But suppose we make a statement about all integers less than a particular finite number n. Could we use the computer and check the statement for all such numbers and show that it is true and call it a mathematical proof? Well, it did not happen exactly like that, but in 1976 computers made their first nontrivial appearance in the realm of mathematical proofs. It involved the famous *four-color problem*.

Consider a map of the world on a flat surface or on a globe.[4] We want to color the map with the condition that no two countries with a common border are of the same color. To make the problem more specific, the areas of all countries have to be contiguous, and no common boundary can be only a single point. The first condition rules out countries with two or more pieces; for instance, the United States (because Alaska and Hawaii are not attached to the mainland). The question is, how many colors do we need? Cartographers have dealt with this problem for ages. It was conjectured that the feat could be accomplished with four or fewer colors. But proof of this conjecture seemed to be out of reach.

In 1976 Kenneth Appel and Wolfgang Haken[5] proved the theorem with partial help from a computer. The proof relies on old-fashioned mathematical work, but 1200 hours of computer time were used to check certain difficult cases.

Computer-assisted mathematical proofs are still exceptions and most mathematicians go about their work in the old-fashioned way. It is said that computer proofs are uncertain and cannot be checked and verified. The uncertainty arises because there may be faults in the hardware, problems with the operating system, or bugs in the program. Assuming that these issues have been thoroughly checked, we can be sure with a high probability of the validity of the proof. This is different from traditional proofs that are offered with certainty and there can be no doubt about them, even an infinitesimal one. If computer-assisted proofs become the prevalent mode of work, mathematics would resemble physics, in which laws are tested and either rejected or not rejected, but never 100% accepted. Furthermore, it is said that

---

$$k = p_1 \times p_2 \times \cdots \times p_r$$

Therefore, p_{n+1} must be a prime.

[4]Technically, we are talking about a *planar* map or graph. Suppose we represent every country by a node and connect each pair of the nodes representing adjacent countries by a line. If we are able to draw such a graph without the lines crossing, then the graph is planar.

[5]For a better idea of the problem and its solution you may want to check Appel and Haken's article in *Scientific American* (October 1977) or their book *Every Planar Map is Four Colorable* (1989). A more technical understanding of the subject could be gained from textbooks on graph theory or discrete mathematics.

computer-assisted proofs that involve thousands of lines of codes cannot be verified; no one would spend her energy on the thankless job of checking a complicated computer program.

One may think of these issues as a matter of degree. After all, complicated proofs such as Gödel's theorem or Fermat's last theorem (conjecture)[6] cannot easily be checked even by many mathematicians. On the other hand, if computers make headway in proving mathematical theorems, we can imagine that in the future mathematical proofs will not be checked but confirmed through independent replications and then held to be true with a high probability. There is another way that computers could help in the advancement of mathematics. A computer program could be written to carry out all logical steps necessary for the proof of a theorem. This doesn't mean that the computer is proving the theorem. Rather it is carrying out the instructions of the mathematician. Such a step-by-step operation would be time consuming and too tedious for human beings, but computers don't mind. The procedure would be especially beneficial when the proof runs into tens and perhaps hundreds of pages. Such activities are already under way, but the role of computers in the mathematics of the future is a matter of speculation.

2.5.3 Exercises

E.2.15 Use mathematical induction to show that

$$
\begin{array}{ll}
i. & 1^2 + 2^2 + 3^2 + \ldots + n^2 = \dfrac{n(n+1)(2n+1)}{6} \\
ii. & n! \geq n^2, \qquad \forall n \geq 4 \\
iii. & \sum_{j=1}^{n} (2j - 1) = n^2
\end{array}
$$

---

[6] Consider the equation $x^n + y^n = z^n$. If $n = 2$, we can find integers satisfying the equation $3^2 + 4^2 = 5^2$. But could the same be done for $n \geq 3$? French mathematician Pierre de Fermat (1601–1665) claimed that he could prove that no such solutions could be found. But because he was writing on the margin of a book, he said he could not write it out. In all likelihood he did not have such a proof. Over the years, many contributed to the solution of the problem. In 1993, the British mathematician Andrew John Wiles (1953) (now at Princeton University in the United States) announced that he had proved the theorem. But there was a significant gap in the proof that took Wiles and a co-worker one and a half years to fill. There are two books written for the public on this subject: *Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem* (1996) by Amir Aczel, and *Fermat's Enigma: The Epic Quest to Solve the World's Greatest Mathematical Problem* (1997) by Simon Singh. Both are available in paperback.

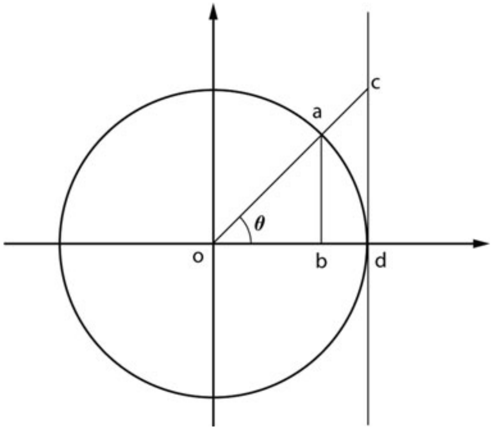

Fig. 2.3. Geometric representation of trigonometric functions

2.6 Trigonometry

Trigonometry, one of the most fun areas of mathematics,[7] has many practical applications in engineering, statistics, and econometrics, as well as in everyday life. What is more, it requires learning only a few basic relationships and the rest is a matter of deduction. Consider the circle in Fig. 2.3. It has a radius of unity. We define the following functions[8] of the angle θ:

$$\sin\theta = \frac{\text{ab}}{\text{oa}} = \text{ab}, \qquad \cos\theta = \frac{\text{ob}}{\text{oa}} = \text{ob}, \qquad \tan\theta = \frac{\text{dc}}{\text{od}} = \text{dc} \tag{2.45}$$

Thus, for the angle θ and the point a on the unit circle, $\cos\theta$ and $\sin\theta$ are, respectively, the coordinates of the point a on the x-and y-axes. If we consider the right-angle triangle oab, then $\sin\theta$ is the ratio of the side opposing the angle to the hypotenuse. Similarly, $\cos\theta$ is the ratio of the side forming the angle to the hypotenuse. This definition applies to all right-angle triangles regardless of the length of the hypotenuse. In the case depicted in Fig. 2.3, the hypotenuse has a length of one and, therefore, we can ignore the denominator of the ratios.

A graph of $\sin(x)$ is shown in Fig. 2.4. As Figs. 2.3 and 2.4 show, both sine and cosine functions take values between -1 and 1. The tangent function, however, is bounded neither from below nor from above. If we multiply sine or cosine functions by ρ, the range of the functions is changed from $[-1, 1]$ to $[-\rho, \rho]$ and ρ is referred to as the *amplitude*.

[7] The interested reader is referred to *Trigonometric Delights* by Eli Maor (1998).

[8] Other trigonometric functions exist, but we will not discuss them here because economists rarely if ever come across them and, therefore, we have no reason to clutter the subject with many unfamiliar notations.

Using elementary geometry the following relationships can be deduced:

$$\tan\theta = \frac{\sin\theta}{\cos\theta} \tag{2.46}$$

$$\sin^2\theta + \cos^2\theta = 1 \tag{2.47}$$

The first is based on the Thales theorem[9] and the second on the Pythagoras theorem.[10] Observe the notation for the square of a trigonometric function. It is written $\sin^2\theta$ and not $\sin\theta^2$, as the latter means the angle θ is raised to the power 2. Of course, we could write it as $(\sin\theta)^2$, but we prefer the economy in the universally accepted convention. Recall that angles can be measured in terms of degrees, radians, and grads. A circle spans 360 degrees, 2π radians, and 400 grads. Thus, a right angle would be 90 degrees, $\pi/2$ radians, and 100 grads. In this book and in most mathematics books, angles are measured in radians.

From Fig. 2.3 it is evident that

$$\begin{aligned} &\sin 0 = \sin\pi = \sin 2\pi = 0, \\ &\sin\frac{\pi}{2} = 1, \\ &\sin\frac{3\pi}{2} = -1 \end{aligned} \tag{2.48}$$

Similarly

$$\begin{aligned} &\cos 0 = \cos 2\pi = 1, \\ &\cos\frac{\pi}{2} = \cos\frac{3\pi}{2} = 0, \\ &\cos\pi = -1 \end{aligned} \tag{2.49}$$

In addition, using well-drawn circles and a ruler, the reader should convince herself of the following identities:

$$\begin{aligned} \sin\left(\theta + \frac{\pi}{2}\right) &= \cos\theta, & \cos\left(\theta + \frac{\pi}{2}\right) &= -\sin\theta, \\ \sin(\theta+\pi) &= -\sin\theta & \cos(\theta+\pi) &= -\cos\theta, \\ \sin(\theta+2\pi) &= \sin\theta, & \cos(\theta+2\pi) &= \cos\theta \\ \sin(-\theta) &= -\sin\theta, & \cos(-\theta) &= \cos(\theta) \end{aligned} \tag{2.50}$$

Trigonometric functions are programmed in all scientific calculators. In addition, software such as Excel, Matlab, and Maple also have these functions. Matlab's trigonometric functions follow.

---

[9] The theorem is named after Thales de Miletos (624 B.C.–547 B.C.) although the germ of the idea dates back to 1650 B.C. and the building of the Pyramids.

[10] This is the famous Pythagoras theorem that the square of hypotenuse is equal to the sum of the squares of the other two sides of a right-angle triangle. Egyptians who built the Pyramids clearly had an empirical understanding of this theorem. Pythagoras (569 B.C.–475 B.C.), for whom the theorem is named, is one of the great mathematicians of antiquity and pioneers of mathematics.

Matlab code

```
% sin, cos and tan of x are obtained using
sin(x)
cos(x)
tan(x)
% Matlab assumes that x is expressed in terms of
% radians. Thus, the sin of π/6 or 30° can be
% calculated in one of the the following ways
sin(pi./6)
% or
x = 30;
sin(x.*pi./180)
% Both return
ans =
        0.5000
```

In Excel the sine, cosine, and tangent of an angel, say $\pi/4$, can be obtained as

=SIN(PI()/4)	=COS(PI()/4)	=TAN(PI()/4)

The fact that $\sin(\theta + \pi/2) = \cos(\theta)$ shows that the sine and cosine functions are out of phase by $\pi/2$ radians. In other words, it takes $\pi/2$ angle rotations for the sine function to catch up with the cosine function. Similarly, the two functions $y_1 = \sin(\theta)$ and $y_2 = \sin(\theta + \pi/4)$ (see Fig. 2.4) are out of phase by $\pi/4$. In general, when we have $\sin(\phi + \theta)$ or $\cos(\phi + \theta)$ with ϕ being a constant, then ϕ is referred to as the *phase*.

In many applications we need to find trigonometric functions of sums or differences of two or more angles. The following relationships exist between trigonometric functions of sums and differences of angles, and the trigonometric functions of the angles themselves.

$$\sin(\theta \pm \phi) = \sin\theta\cos\phi \pm \sin\phi\cos\theta \tag{2.51}$$

$$\cos(\theta \pm \phi) = \cos\theta\cos\phi \mp \sin\theta\sin\phi \tag{2.52}$$

Letting $\phi = \theta$, we have

$$\sin 2\theta = 2\sin\theta\cos\theta \tag{2.53}$$

and

$$\cos 2\theta = \cos^2\theta - \sin^2\theta \tag{2.54}$$

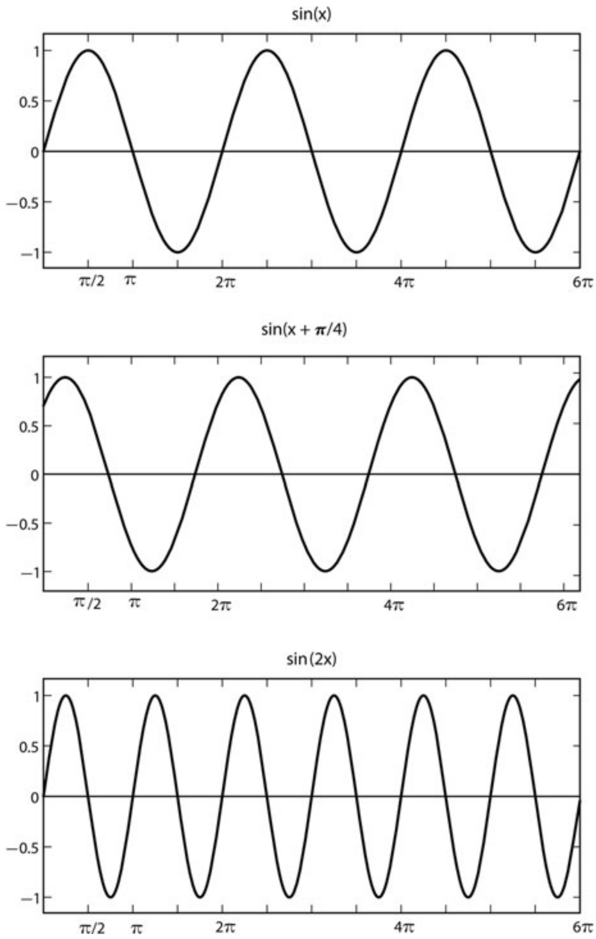

Fig. 2.4 Sin functions with different phases and frequencies

Recalling (2.47), (2.54) can be written as

$$\cos 2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta \tag{2.55}$$

Example 2.21

$$\sin\left(\frac{\pi}{5}\right) = 0.588 \qquad \cos\left(\frac{\pi}{5}\right) = 0.809$$

$$\sin\left(2\frac{\pi}{5}\right) = 2\sin\left(\frac{\pi}{5}\right)\cos\left(\frac{\pi}{5}\right) = 2 \times 0.588 \times 0.809 = 0.951$$

$$\cos\left(2\frac{\pi}{5}\right) = \cos^2\left(\frac{\pi}{5}\right) - \sin^2\left(\frac{\pi}{5}\right) = 0.809^2 - 0.588^2 = 0.309$$

2.6.1 Cycles and Frequencies

Trigonometric functions are cyclical because as the point on the circle travels counterclockwise, it comes back to the same point again and again (see Fig. 2.4). As a result, the sine and cosine functions assume the same values for angles $\theta, 2\pi + \theta$, and, in general, $2k\pi + \theta$. Similarly, the tan function has the same value for θ and $\theta + \pi$. These functions are called *periodic*.[11] Compare the two functions

$$y_1 = \sin(x) = \sin(x + 2k\pi)$$

and

$$y_2 = \sin(2x) = \sin(2x + 2k\pi) = \sin[2(x + k\pi)]$$

Clearly y_2 returns to the same value—or completes a cycle—twice as fast as y_1. In general,

$$y = \sin(fx) \tag{2.56}$$

completes a cycle f times faster than $\sin(x)$. We call f the frequency of the function. Alternatively we can write (2.56) as

$$y = \sin\left(\frac{x}{p}\right) \tag{2.57}$$

Because $p = 1/f$, it is clear that every p periods the function will have the same value. In other words, the function completes a cycle in p periods or the cycle length is p. These concepts are better understood if we take the argument of the function to be time, measured in discrete values for a given time interval, that is, $t = 1, \ldots, T$. Let

$$y = \sin\left(\frac{2\pi t}{p}\right) \tag{2.58}$$

If $p = T$, then it takes T time periods to complete the cycle and the frequency is $1/T$. On the other hand, if frequency is $4/T$, then the length of the cycle is $T/4$. As an example, let time be measured in months and the period under consideration be a year, that is, $T = 12$. If $p = 3$, then we have four cycles per year and the frequency is 1/3 of a cycle per month. On the other hand, if $p = 1/2$, there are 24 cycles in a year and the frequency per month is two.

---

[11] A function $y = f(x)$ is called periodic if $f(x) = f(x + c), \quad c \neq 0$.

2.6.2 Exercises

E.2.16 Find the numerical value of the following:

$$i.\ \sin\frac{3}{2}\pi,\qquad ii.\ \frac{\cos^2\frac{7}{2}\pi-\sin\frac{2}{3}\pi}{2\cos\frac{5}{4}\pi},\qquad iii.\ \frac{\sin^3\frac{5}{3}\pi-\tan\frac{3}{4}\pi}{\sin\frac{3}{4}\pi}$$

E.2.17 Graph the following functions in the interval $0\le x\le 2\pi$.

$$\begin{array}{llll} i. & y=\cos\left(x+\frac{\pi}{2}\right), & ii. & y=\sin\left(x+\frac{\pi}{2}\right) \\ iii. & y=\cos\left(x+\frac{\pi}{2}\right)+\sin\left(x+\frac{\pi}{2}\right), & iv. & y=\cos\left(x+\frac{\pi}{2}\right)-\sin\left(x+\frac{\pi}{2}\right) \\ v. & y=\tan\left(x+\frac{\pi}{2}\right), & vi. & y=\sin x+\sin 2x+\sin 3x \\ vii. & y=\sin x+0.5\sin 2x+0.25\sin 3x & & \end{array}$$

E.2.18 Show that

$$\tan 2\theta=\frac{2\tan\theta}{1-\tan^2\theta}$$

E.2.19 Write $\sin 3x$ in terms of $\sin x$ and its powers.
E.2.20 Write $\sin 4x$ in terms of $\sin x$ and its powers.
E.2.21 Show that

$$\begin{array}{llll} i. & \dfrac{1}{\tan\theta+\dfrac{1}{\tan\theta}}=\sin\theta\cos\theta & ii. & \dfrac{1-\sin x}{\cos x}=\dfrac{\cos x}{1+\sin x} \\ iii. & \dfrac{1+\sin\theta}{1-\sin\theta}-\dfrac{1-\sin\theta}{1+\sin\theta}=4\dfrac{1}{\cos\theta}\tan\theta & iv. & \dfrac{1-\sin\theta}{1+\sin\theta}=\left(\tan\theta-\dfrac{1}{\cos\theta}\right)^2 \end{array}$$

2.7 Complex Numbers

Complex numbers are two-dimensional numbers where one dimension is on the real axis and the other on the imaginary axis.[12] We are already familiar with real numbers and the real line. The imaginary number is

$$i=\sqrt{-1} \tag{2.59}$$

[12] It is customary to introduce complex numbers in the context of the solution to quadratic equations involving the square root of a negative number. This practice has the unfortunate consequence that students may get the impression that somewhere among the real numbers or along the real line there are caves where complex numbers are hiding and once in a while show their faces.

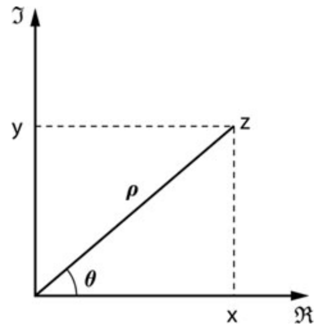

Fig. 2.5 Point z in the complex plane

i is an imaginary number because $i^2 = -1$, and there is no real number whose square is a negative number. Geometrically, a complex number is a point in the two-dimensional complex space. Any function of complex variables maps these variables into the two-dimensional complex plane.

Figure 2.5 depicts point z in the complex plane where the horizontal axis is the real line and the vertical axis the imaginary line. Thus, we can write z as

$$z = x + iy \tag{2.60}$$

Two complex numbers are equal if they are equal in both real and imaginary dimensions. That is, $z_1 = x_1 + iy_1$ is equal to $z_2 = x_2 + iy_2$ if $x_1 = x_2$ and $y_1 = y_2$. Real numbers are a special case of complex numbers when the imaginary dimension is set equal to zero. Similarly, an imaginary number is a complex number with its real dimension set equal to zero.

Example 2.22 The following are examples of complex numbers:

$$z_1 = 3 + i, \qquad z_2 = 5 - 3i, \qquad z_3 = 6 + 0.5i$$

Complex numbers come in pairs. Every complex number has its twin, called a *conjugate*. If $z = x + iy$, then its conjugate complex number is $\bar{z} = x - iy$. It follows that $\bar{\bar{z}} = z$. In other words, z is the conjugate of $\bar{z}$.

Example 2.23 The conjugates of the complex numbers in Example 2.22 are

$$\bar{z}_1 = 3 - i, \qquad \bar{z}_2 = 5 + 3i, \qquad \bar{z}_3 = 6 - 0.5i$$

Operations of addition, subtraction, and multiplication of complex variables are defined as

$$
\begin{aligned}
z_1 + z_2 &= (x_1 + x_2) + (y_1 + y_2)\, i \\
z_1 - z_2 &= (x_1 - x_2) + (y_1 - y_2)\, i \\
z_1 z_2 &= (x_1 + iy_1)\,(x_2 + iy_2) \\
&= (x_1x_2 - y_1y_2) + (x_1y_2 + x_2y_1)\, i
\end{aligned}
\tag{2.61}
$$

Example 2.24

$$
\begin{aligned}
(3 - i) + (5 + 3i) &= 8 + 2i \\
(5 + 3i) - (6 + 0.5\, i) &= -1 + 2.5\, i \\
(6 - 0.5i)\,(3 + i) &= 18.5 + 4.5\, i
\end{aligned}
$$

Addition, subtraction, and multiplication of a complex number by its conjugate result in

$$
\begin{aligned}
z + \bar{z} &= 2x \\
z - \bar{z} &= 2iy \\
z\bar{z} &= x^2 + y^2 = \rho^2
\end{aligned}
\tag{2.62}
$$

where the last equality refers to Fig. 2.5 and is based on the Pythagoras theorem. Division of complex numbers is a bit more involved:

$$
\frac{z_1}{z_2} = \frac{z_1\bar{z}_2}{z_2\bar{z}_2} = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i\frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2}
\tag{2.63}
$$

Example 2.25

$$
\frac{6 - 0.5i}{5 + 3i} = \frac{28.5}{34} - i\frac{20.5}{34} \approx 0.838 - 0.603i
$$

Referring again to Fig. 2.5, we observe that

$$
x = \rho\cos\theta \qquad \text{and} \qquad y = \rho\sin\theta
$$

which implies that

$$
\begin{aligned}
\rho^2 &= x^2 + y^2 \\
\tan\theta &= \tfrac{y}{x}
\end{aligned}
$$

These relationships enable us to write a complex variable either in terms of its Euclidean coordinates or in terms of ρ and θ, that is, its *polar coordinates*:

$$
z = x + iy = \rho(\cos\theta + i\sin\theta)
\tag{2.64}
$$

where

$$
\begin{aligned}
\rho &= \sqrt{x^2 + y^2} \\
\theta &= \tan^{-1}\tfrac{y}{x}
\end{aligned}
\tag{2.65}
$$

Example 2.26 Let us rewrite complex numbers in Example 2.22 using polar coordinates.[13] For

$$z_1 = 3 + i$$

we have

$$\rho = \sqrt{3^2 + 1^2} = \sqrt{10}$$
$$\theta = \tan^{-1}\frac{1}{3} = 0.32175 = 0.102416\pi$$

therefore,

$$z_1 = \sqrt{10}(\cos 0.102416\pi + i \sin 0.102416\pi)$$

Similarly,

$$\begin{aligned} z_2 &= \sqrt{34}(\cos 0.54042 - i \sin 0.54042) \\ &= \sqrt{34}(\cos 0.17202\pi - i \sin 0.17202\pi) \end{aligned}$$

$$\begin{aligned} z_3 &= \sqrt{36.25}(\cos 0.08314 + i \sin 0.08314) \\ &= \sqrt{36.25}(\cos 0.02646\pi + i \sin 0.02646\pi) \end{aligned}$$

Note that because $\sin(-\theta) = -\sin\theta$, (2.64) implies

$$\bar{z} = \rho(\cos\theta - i \sin\theta)$$

We have a third way to write complex numbers. For this, we state without proof the following relationships[14]:

$$\begin{aligned} \exp(i\theta) &= \cos\theta + i \sin\theta \\ \exp(-i\theta) &= \cos\theta - i \sin\theta \end{aligned} \tag{2.66}$$

Therefore,

$$\begin{aligned} x + iy &= \rho \exp(i\theta) \\ x - iy &= \rho \exp(-i\theta) \end{aligned} \tag{2.67}$$

Where ρ and θ are as defined in (2.65).

[13] Angles are measured in radians. If you use a calculator, you need to set it in the radian mode to get the same numbers as in the text. If your calculator is in the degree mode, then in order to get the same numbers as in the text, $\theta = \tan^{-1}(x/y)$ needs to be converted into radians by multiplying it by $\pi/180$.

[14] We shall provide a proof of these relationships in Chap. 10.

Example 2.27 Again using complex numbers in Example 2.22, we have

$$
\begin{aligned}
z_1 &= \sqrt{10}\exp(0.32175i)\\
z_2 &= \sqrt{34}\exp(-0.54042i)\\
z_3 &= \sqrt{36.25}\exp(0.08314i)
\end{aligned}
$$

Using the following program, the reader could check the validity of the formulas in (2.66) for different values of the angle t.

Matlab code

```
% Set the value of the angle
t = pi./3;
% Trigonometric version
cos(t)+i*sin(t)
% Exponential version
exp(i*t)
% Trigonometric version
cos(t)-i*sin(t)
% Exponential version
exp(-i*t)
```

The idea of the equivalence of circular sine and cosine functions with the exponential function may bother the intuitive sense of some readers. But $\exp(i\theta)$ is indeed a circular function in the complex plane that traces a circle as θ changes from 0 to 2π. On the other hand, ρ determines the distance of the point from the origin. Indeed, we can define trigonometric functions in terms of the exponentials of complex numbers.

$$
\begin{aligned}
\cos\theta &= \frac{\exp(i\theta)+\exp(-i\theta)}{2}\\
\sin\theta &= \frac{\exp(i\theta)-\exp(-i\theta)}{2i}
\end{aligned}
\tag{2.68}
$$

An important consequence of (2.66) is De Moivre's theorem.[15]

Theorem 2.1

$$
\begin{aligned}
z^k &= [\rho(\cos\theta + i\sin\theta)]^k\\
&= \rho^k(\exp(i\theta))^k\\
&= \rho^k\exp(ik\theta)\\
&= \rho^k(\cos k\theta + i\sin k\theta)
\end{aligned}
\tag{2.69}
$$

[15] Abraham De Moivre (1667–1754), a French mathematician who spent most of his life in England, was a pioneer in the development of probability theory and analytic geometry. He was appointed to the commission set up to examine Newton's and Leibnitz's claims for the discovery of calculus.

Example 2.28

$$(\cos\theta + i\sin\theta)^3 = \cos 3\theta + i\sin 3\theta$$
$$(\cos\theta - i\sin\theta)^3 = \cos 3\theta - i\sin 3\theta$$

Example 2.29 *Let* $\theta = \pi/4$, *then*

$$\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)^5 = (0.7071 + 0.7071i)^5$$
$$= -0.7071 - 0.7071i$$

$$\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4} = \cos\left(\pi + \frac{\pi}{4}\right) + i\sin\left(\pi + \frac{\pi}{4}\right)$$
$$= -\cos\frac{\pi}{4} - i\sin\frac{\pi}{4}$$
$$= -0.7071 - 0.7071i$$

2.7.1 Exercises

E.2.22 Write the following complex numbers in alternative forms of (2.64) and (2.67)

i. $z_1 = 1 + i$ *ii.* $z_2 = 1 - i$
iii. $z_3 = 5i$ *iv.* $z_4 = 3.5 - 2.6i$
v. $z_5 = 7 + 4i$

E.2.23 Referring to E.2.21, compute

i. $z_1 z_2$ *ii.* $\frac{z_1}{z_2}$ *iii.* $z_3 z_4$
iv. $\frac{z_4}{z_5}$ *v.* $\bar{z}_3 z_5$ *vi.* $\frac{\bar{z}_5}{\bar{z}_3}$

E.2.24 We already know that $e^a e^b = e^{a+b}$ where a and b are real numbers. Show that for real numbers a and b,

$$e^{ai+bi} = e^{ai} e^{bi}$$

Foundations of Mathematical and Computational Economics
Dadkhah, K.
2011, XVI, 542 p., Hardcover
ISBN: 978-3-642-13747-1