DOCUMENT RESUME

ED 218 127 SE 038 237

AUTHOR Gale, David; And Others
TITLE Optimal Assignment Problem Applications of Finite Mathematics to Business and Economics. [and] Difference Equations with Applications. Applications of Difference Equations to Economics and Social Sciences. [and] Selected Applications of Mathematics to Finance and Investment. Applications of Elementary Algebra to Finance. [and] Force of Interest. Applications of Calculus to Finance. Modules and Monographs in Undergraduate Mathematics and Its Applications Project. UMAP Units 317, 322, 381, 382.
INSTITUTION Education Development Center, Inc., Newton, Mass.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 80
GRANT SED-76-19615; SED-76-19615-A02
NOTE 160p.

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Business; *Business Education; Calculus; *College Mathematics; Economics; *Economics Education; Higher Education; Instructional Materials; Interest (Finance); *Learning Modules; *Mathematical Applications; Matrices; Problem Solving; *Social Sciences; Supplementary Reading Materials
IDENTIFIERS *Finance

ABSTRACT

Four units make up the contents of this document. The first examines applications of finite mathematics to business and economies. The user is expected to learn the method of optimization in optimal assignment problems. The second module presents applications of difference equations to economics and social sciences, and shows how to: 1) interpret and solve elementary difference equations and 2) understand how difference equations can be used to model certain problems. The next unit looks at applications of elementary algebra to finance. The goal is to help students understand business and financial concepts and formulas, and teach them how to apply these concepts and formulas in practical circumstances. Appreciation of the power and usefulness of mathematical concepts and techniques is also promoted. The final module examines applications of calculus to finance. The user is taught to: 1) appreciate a natural connection between exponential functions and compound interest; 2) apply the definition of the derivative in a non-science situation; and 3) understand terms used in finance such as simple interest, compound interest, yield rate, nominal rate, and force of interest. Exercises and answers come in all units. The last module contains a model exam. Solutions are provided. (MP)

umap

UNIT 317

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT

THE OPTIMAL ASSIGNMENT PROBLEM

by David Gale

		Profits			
		9	6	5	0
Wages	2	10	⑧ ✓	⑦	0
	0	8	4	3	⓪ ✓
	0	⑨ ✓	⑥	4	⓪
	0	-2	3	⑤ ✓	⓪

APPLICATIONS OF FINITE MATHEMATICS
TO BUSINESS AND ECONOMICS

edc/umap /55chapel st./newton.mass 02160

THE OPTIMAL ASSIGNMENT PROBLEM

by

David Gale
Department of Mathematics
University of California
Berkeley, California 94720

TABLE OF CONTENTS

Intermodular Description Sheet: UMAP Unit 317

Title: THE OPTIMAL ASSIGNMENT PROBLEM

Author: David Gale
Department of Mathematics
University of California
Berkeley, CA 94720

Review Stage/Date: III 9/29/78

Classification: APPL FINITE MATH/BUSINESS & ECON

Suggested Support Materials:

Prerequisite Skills:
Familiarity with the notion of permutation, and with elementary row and column operations on matrices.

Output Skills:

1. Understand and be able to use the method of opti.
2. Achieving an optimal assignment of personnel described in the unit.

Other Related Units:

MODULES AND MONOGRAPHS IN UNDERGRADUATE

MATHEMATICS AND ITS APPLICATIONS PROJECT (UMAP)

The goal of UMAP is to develop, through a community of users and developers, a system of instructional modules in undergraduate mathematics and its applications which may be used to supplement existing courses and from which complete courses may eventually be built.

The Project is guided by a National Steering Committee of mathematicians, scientists and educators. UMAP is funded by a grant from the National Science Foundation to Education Development Center, Inc., a publicly supported, nonprofit corporation engaged in educational research in the U.S. and abroad.

PROJECT STAFF

NATIONAL STEERING COMMITTEE

The Project would like to thank members of the UMAP Finite Mathematics Panel for their reviews, and all others who assisted in the production of this unit.

This material was prepared with the support of National Science Foundation Grant No. SED76-19615. Recommendations expressed are those of the author and do not necessarily reflect the views of the NSF nor of the National Steering Committee.

THE OPTIMAL ASSIGNMENT PROBLEM

1. Description of the Problem

The optimal assignment problem is concerned with situations of the following kind: a certain company has a number of job openings and there are a number of applicants available for the jobs. The company's problem is to decide how to assign applicants to jobs in a way which will maximize the benefit to the company. Before making the assignment each applicant is given a set of tests which are designed to measure his aptitude for each of the jobs. These test *scores* can be displayed by means of a table as illustrated below for the case of three applicants and three jobs

(1.1)

	J_1	J_2	J_3
A_1	10	8	7
A_2	4	8	3
A_3	6	4	9

The number in row i and column j of the table which we call *position* (i,j) gives the score of the i^{th} applicant for the j^{th} job. One may think of these aptitudes as measuring the value of the applicant to the company when assigned to the given job. Thus, in the example A_1 is worth, say, 10 dollars per hour when assigned to J_1 but only 7 dollars per hour when assigned to J_3. If the scores are interpreted in this way then clearly the company will achieve maximum benefit by assigning applicants in such a way as to maximize the sum of the scores. Such an assignment is called an *optimal assignment.* In the example the optimal assignment can be found by inspection, for observe that A_1 is best at J_1, A_2 is best at J_2 and A_3 is best at J_3. Therefore the optimal assignment is obtained by assigning each person to the job he does best. It is convenient to indicate the assignment by checking the corresponding entries in the table as indicated below:

(1.2)

	J_1	J_2	J_3
A_1	10✓	8	7
A_2	4	8✓	3
A_3	6	4	9✓

Without making any further calculation we know that the assignment checked in (1.2) is optimal, for since every applicant is assigned to the job he does best, no other assignment could raise the score of any applicant, hence the total score can not be raised.

Of course there was considerable luck in the above example, for it was simply fortuitous that each applicant was best at a different job. In general one cannot expect this to happen. Here is another example:

(1.3)

	J_1	J_2	J_3
A_1	10✓	8	6
A_2	8	4	7✓
A_3	6	9✓	4

Notice that in this case it is no longer possible to assign each person to the job he does best, for both A_1 and A_2 are best at J_1. Nevertheless the optimal solution is easily found by changing the point of view. Instead of trying to assign each applicant to the job he does best we try to pick the best man for each job. The best man for J_1 is A_1, for J_2 is A_3 and for J_3 is A_2. Therefore the optimal assignment is the one given by the checked entries in (1.3).

The procedure used in these examples can be described concisely as follows: the tables like (1.1), (1.2), and (1.3) are called *assignment matrices*. In the first example we chose the largest entry in each row. It turned out that each of these entries was in a different column so that an assignment was obtained which was therefore optimal. In such cases we will say that the matrix in question has a *row-max* assignment. In the second example the matrix did not have a row-max assignment but it did have a *column-max* assignment, meaning that the maximum entries in each column were all in different rows. Now in general a matrix need not have either a row- or column-max assignment as the following example illustrates:

$$(1.4) \qquad \begin{array}{c|c|c|c|} & J_1 & J_2 & J_3 \\ \hline A_1 & 10 & 8 & 7\checkmark \\ \hline A_2 & 8\checkmark & 4 & 3 \\ \hline A_3 & 9 & 6\checkmark & 4 \\ \hline \end{array}$$

One sees at once that neither of the foregoing methods works, for if each person was assigned to the job he does best, then everyone would be assigned to J_1, and if each job was given to the best man for the job, then A_1 would be assigned to all three jobs. It is claimed that the optimal assignment is given by the checked entries, which give the total *value* of the assignment to be 7 + 8 + 6 = 21. The reader should verify that this is optimal by simply trying each of the other five possible assignments and observing that they give a lower total score. The reader should also note the following facts. In this optimal assignment only one applicant, A_2, is assigned to the job he does best and in fact, A_1 is assigned to J_3, the job he does worst. Likewise, only J_3 is assigned to the best man for the job. Nevertheless, as we have seen, this assignment is optimal, and this suggests that "common sense methods" will not be very helpful in solving such problems and that some kind of "theory" is needed.

Of course for small 3×3 examples, one can always find a solution by "trial and error," meaning one simply lists all possible assignments and chooses the one with the largest value. This procedure, however, is clearly impractical for even moderate sized problems. For example, if one were to use the method on a 5×5 problem it would require listing 120 possibilities each of which would involve performing 4 additions so that 480 additions would be required. The method to be presented in these notes will enable the reader to solve 8×8 problems by hand in a very moderate amount of time. If one were to do this by listing all possibilities it would require over 250,000 additions.

In a general assignment problem the number of jobs and applicants need not be equal. There may be more applicants than jobs ($m > n$) in which case the company will hire the n applicants giving it the highest total value, or there may be more jobs than applicants ($n > m$) in which case the company will fill the m jobs which give the highest total value. However, there is a simple trick whereby all problems can be reduced to the "square" case in which the number of jobs and applicants are equal. If, say, $m > n$ then the assignment matrix has more rows than columns. Then one simply augments the matrix by adding $m - n$ additional columns all of whose entries are zero. This gives an m×m matrix with $m - n$ additional "dummy" jobs which may be thought of as the job of being unemployed. Now clearly if one finds the solution to this problem one has also obtained the solution to the original problem, for assigning a person to a dummy job means not assigning him at all in the original problem. In a similar way if $m < n$ one adds $n - m$ rows of zeros. The reader should convince himself that a solution to this $n \times n$ problem also solves the original $m \times n$ problem. From now on we will restrict ourselves to the square case. This permits us to formalize the notion of an assignment in the following way.

DEFINITION. An $n \times n$ assignment problem consists of an $n \times n$ matrix A. An assignment α consists of a *permutation* $[i_1, i_2, \ldots, i_n]$ of the integers from 1 to n. This is to be read as follows: in the assignment α applicant A_1 is assigned to job J_{i_1}, applicant A_2 is assigned to job J_{i_2}, etc. The *value* of the assignment α, denoted by $v(\alpha)$ is the number $a_{1i_1} + a_{2i_2} + \ldots + a_{ni_n}$. An assignment α having the maximum value among all possible assignments is called an *optimal assignment*.

Using the above notation the optimal assignment for (1.1) is [1, 2, 3] with value $a_{11} + a_{22} + a_{33} = 27$. The optimal assignment for (1.3) is [1, 3, 2] with value $a_{11} + a_{23} + a_{32} = 26$ and the optimal assignment for (1.4) is [3, 1, 2] with the value $a_{13} + a_{21} + a_{32} = 21$.

Exercises

1. How many possible assignments are there in an n×n problem?
2. For the matrix below calculate the value of each of the following assignments:
[1, 2, 3, 4] [4, 3, 2, 1] [3, 1, 2, 4] [2, 3, 4, 1].

$$\begin{pmatrix} 9 & 8 & \textcircled{8} & 7\checkmark \\ 8\checkmark & 7 & 7 & 5 \\ 7 & \textcircled{7} & 6\checkmark & \textcircled{5} \\ \textcircled{7} & 6\checkmark & 5 & 4 \end{pmatrix}$$

3. For the above matrix write in the form $[i_1, i_2, i_3, i_4]$ the assignment corresponding to the checked entries of the matrix. Do the circled entries above correspond to an assignment? Why?

2. A Paradox

Before proceeding to develop the theory for the assignment problem it is worthwhile pointing out one further property of such problems. We return to the third example given by the table

(2.1)

	J_1	J_2	J_3
A_1	10	8	7✓
A_2	8✓	4	3
A_3	9	6✓	4

where the entries in the optimal assignment have been checked. Now suppose a fourth applicant A_4 appears, is tested, and obtains the following scores

(2.2)

	J_1	J_2	J_3
A_4	2	3	5

The question which now arises is whether A_4 should replace any of the three applicants of the original assignment. Observe that his test scores for each job are lower by 6, 3, and 2 points respectively than the scores of the people assigned to the jobs in (2.1). At first glance one might conclude therefore that A_4 is less qualified than the present work force and should not be hired, but this turns out to be a wrong conclusion, for if A_4 is assigned to J_3, A_1 to J_2 and A_3 to J_1 the value of the assignment is $5 + 8 + 9 = 22$ which is an improvement over the previous assignment whose value was 21. Thus, it is optimal to replace A_2 by A_4 even though A_2's overall score is well above that of A_4. Once again, we see that common sense does not seem to be very helpful in attacking these problems.

Exercises

4. Find by listing all possibilities the optimal assignment for the table given by

$$\begin{array}{c|c|c|c|}
 & J_1 & J_2 & J_3 \\ \hline
A_1 & 6 & 5 & 2 \\ \hline
A_2 & 4 & 2 & 0 \\ \hline
A_3 & 2 & 2 & 1 \\ \hline
\end{array}$$

5. Suppose a_{11} above is increased from 6 to 7. Will the value of the optimal assignment increase? In general for which values of a_{ij} in Exercise 4 will an increase of one unit produce an increase in the value of the optimal assignment?
6. For which a_{ij} in Exercise 4 will a decrease of one unit produce a decrease in the value of the optimal assignment?
7. If a fourth applicant A_4 becomes available with scores (1, 4, 0) for the three jobs will it be possible to find an improved assignment in Exercise 4? If so who will A_4 replace?
8. Same as Exercise 7 if A_4 has scores (4, 3, 1).
9. Recall in the example given by (1.4) the optimal assignment assigned only one person to the job he does best and only one job to the person who does it best. Show by giving an example that one can have an n × n assignment problem in which this same thing occurs.
10. Is it possible to have an optimal assignment in which no one is assigned to the job he does best? If so give an example. If not give a proof of the impossibility. (This is a somewhat more difficult exercise.)

---

3. Preliminary Theory - (Wages and Profits)

This section is concerned with a simple but fundamental property of assignment problems which will play the key role in the theory to follow.

Two $n \times n$ matrices A and A' are called *equivalent* if they have the same set of optimal assignments meaning that every assignment which is optimal for one is optimal

for the other. Suppose now that A' is the matrix obtained from A by adding or subtracting a constant number w_i to every entry in the i^{th} row of A. The claim is that the matrices A and A' are equivalent. To see this let *a* be any assignment and let $v(a)$ and $v'(a)$ be the value of *a* for the matrices A and A' respectively. Then $v'(a) = v(a) + w_1$ so that the effect of adding w_i simply changes the value of all assignments by the constant amount w_i, and it is clear that adding such a constant to all assignments will not change their *comparative* values. That is, if one assignment gives a higher value than another on A it will also do so on A'. In particular then, the optimal assignments on A are the same as those on A'. The same argument applies if a constant p_j is added to all entries in the j^{th} column of A. We can state this formally as follows:

<u>Theorem 1</u>: *If a constant is added to all entries in any row or column of an assignment matrix, the new matrix obtained is equivalent to the original one.*

Of course one can add constants to any number of the rows or columns of A and all the matrices obtained will be equivalent to A. The idea of our algorithm for solving the assignment problem is to perform a sequence of additions (or subtractions) of constants to the rows of A until we obtain a matrix A' equivalent to A for which the optimal assignment is "obvious" in that the matrix A' will have a column-max assignment as described in Section 1. To illustrate this consider again the matrix given by

$$(3.1) \qquad \begin{array}{|c|c|c|} \hline
10 & 8 & 7\checkmark \\ \hline
8\checkmark & 4 & 3 \\ \hline
9 & 6\checkmark & 4 \\ \hline
\end{array}$$

in which the optimal assignment has been checked. (Henceforth we omit the row and column headings A_i and J_j.) The optimal assignment here is definitely not column-max, but the question is whether one can add or subtract constants to the rows in such a way as to obtain an equivalent matrix for which the given assignment is a column-max. By trying various things one discovers that by subtracting 3 from row 1 and 1 from row 3 one gets the matrix

(3.2)

7	⑤	④✓
⑧✓	4	3
⑧	⑤✓	3

where the entries which are maximal in their columns have been circled. Observe that the checked entries in this matrix provide a column-max assignment and therefore the optimal assignment is [3, 1, 2]. But by Theorem 1 this matrix is equivalent to the original one and hence this must also be an optimal assignment for the original problem. *Notice that we now have a proof of the optimality of this assignment which does not require finding the values of the other five assignments.*

Next consider the case of (3.1) with the fourth applicant with scores (2, 3, 5). By introducing a dummy job we get a 4 × 4 problem whose matrix is given by

(3.3)

10	8✓	7	0
8	4	3	0✓
9✓	6	4	0
2	3	5✓	0

where the optimal assignment [2, 4, 1, 3] has been marked. In order to verify directly that the assignment is optimal, one would have to calculate the value of the other 23 assignments. Instead let us again try to find numbers which when subtracted from the rows make the given assignment a column-max assignment. In trying various things one finds that by subtracting 2 from row 1 we get

(3.4)

8	⑥✓	⑤	-2
8	4	3	⓪✓
⑨✓	⑥	4	⓪
2	3	⑤✓	⓪

where again maximal entries in each column have been circled, and we see for this matrix the given assignment is a column-max, and so we have a proof that the given assigment is indeed optimal.

The process of subtracting constants from the rows of the assignment matrix has an interesting economic interpretation which will be used in developing the later theory. Let us think of the number w_i as the *wage* paid to applicant A_i. Recall now that a_{ij} may be interpreted as the value to the company, say in dollars, when A_i is assigned to J_j. But if the company must pay A_i the wage w_i then the company's *profit* p_j from assigning A_i to J_j is $a_{ij} - w_i$. In other words, after subtracting the constants w_i from rows of A the entries in the new matrix obtained may be thought of as the profits the company makes in filling each of the jobs. It is then natural for the company to assign each job to the applicant which will give it the greatest profit and this corresponds exactly to a column-max assignment in this matrix.

We will now introduce these ideas into our computation. It turns out to be inconvenient to have to rewrite the whole assignment matrix every time one subtracts a constant from the rows. Instead of doing this therefore we will simply list the wage constants w_i next to the appropriate row of the matrix. In addition we list at the head of column j the maximum p_j of the numbers $a_{ij} - w_i$, that is the maximum of the profits obtained by assigning J_j to the various applicants. Finally, we circle the positions in the matrix which yield these maximum profits. Using this notation for example (3.1) one would have instead of (3.2) the following array called a *display* of the problem:

(3.5)
$$\begin{array}{cc|c|c|c|}
 & & \multicolumn{3}{c}{\text{Profits}} \\
 & & 8 & 5 & 4 \\
\hline
 & 3 & 10 & \textcircled{8} & \textcircled{7}\checkmark \\
\text{Wages} & 0 & \textcircled{8}\checkmark & 4 & 3 \\
 & 1 & 9 & \textcircled{6}\checkmark & 4 \\
\hline
\end{array}$$

In the same way instead of rewriting matrix (3.3) as (3.4) we would write:

(3.6)
$$\begin{array}{cc|c|c|c|c|}
 & & \multicolumn{4}{c}{\text{Profits}} \\
 & & 9 & 6 & 5 & 0 \\
\hline
 & 2 & 10 & \textcircled{8}\checkmark & \textcircled{7} & 0 \\
\text{Wages} & 0 & 8 & 4 & 3 & \textcircled{0}\checkmark \\
 & 0 & \textcircled{9}\checkmark & \textcircled{6} & 4 & \textcircled{0} \\
 & 0 & 2 & 3 & \textcircled{5}\checkmark & \textcircled{0} \\
\hline
\end{array}$$

Once again these displays are just a kind of shorthand way of saying that if each of the wage constants are subtracted from the corresponding rows the column maxima will be the profits listed at the head of each column. In view of our present economic interpretation we will refer to the assignments corresponding to the checked entries in (3.5) and (3.6) as *profit-max assignments* rather than column-max assignments.

We now give the general definition of a display for an n × n assignment problem. It consists of (1) the original assignment matrix A; (2) 2n constants $w_1, \ldots, w_n, p_1, \ldots, p_n$, and (3) certain circled entries in the matrix satisfying the conditions

(A) $w_i + p_j \geq a_{ij}$ for all positions (i, j),
(B) if $w_i + p_j = a_{ij}$ then a_{ij} is circled.
(C) there is at least one circle in every column of A.

Conditions (A), (B), and (C) are called the *feasibility conditions*. The reader should check that they are just another way of saying that p_j is the maximum of the numbers $a_{ij} - w_i$ for i = 1, 2, ..., n.

An assignment which uses only circled entries of a display is called a *profit-max assignment*, and the ideas of this section can be summarized in the following theorem:

Theorem 2: *A profit-max assignment is optimal.*

The proof is just a matter of retracing our steps. A profit maximizing assignment with constants w_i and p_j is equivalent to a column-max assignment on the matrix obtained by subtracting w_1 from the i^{th} row of A, but from Theorem 1 subtracting constants from rows of A does not change the optimal assignment.

From now on our objective will be to find the "right" wage constants w_i (and profits p_j) so as to get a display with a profit maximizing assignment. Of course we have not proved that the desired constants exist, much less have we provided a systematic method for finding them. For the present, for small problems the reader should try by experiment and guess work to find the w_i's which work. The sections which follow will show how they can be found in an efficient manner for problems of any size.

Exercises

11. In Exercise 4 find the constants w_i and p_j and give the display showing that the assignment you obtained for that problem is profit maximizing.

12. Set up the problems of Exercise 8 and 9 as 4 × 4 assignment problems. Find constants w_i and p_j and the display with a profit-max solution.

13. Show by finding the appropriate w_i and p_j and giving the display that the checked assignment below is optimal.

9	8	8	7✓
8	7	7✓	5
7	7✓	6	5
7✓	6	5	4

---

4. The Simple Assignment Problem

In order to use the method of Section 3 to solve the assignment problem it is necessary to be able to recognize when an assignment matrix possesses a column-max or profit-max assignment. The following example shows that this may not always be easy.

(4.1)

⑩	8	⑧	⑩	7
9	⑩	4	6	⑧
5	⑩	7	9	⑧
⑩	9	⑧	⑩	5
6	⑩	7	9	⑧

As usual the maximal entries in each column have been circled, and one must now decide whether it is possible to choose a set of five circled entries so that there will be exactly one entry in each row and column. Note that the question has nothing to do with the numbers in the matrix. One could as well consider the following display

(4.2)

0		0	0	
	0			0
	0			0
0		0	0	
	0			0

in which all numbers have been deleted. This problem can also be interpreted in terms of jobs and applicants where instead of each applicant having a test score for each job he merely gets a grade of *pass* or *fail:* If A_i passes the test for J_j a circle is entered in position (i, j) and we say that A_i *qualifies* for J_j. The problem is then to assign as many applicants as possible to jobs for which they are qualified. This is called the *simple assignment* problem. (There is a second somewhat more picturesque interpretation of this problem in which the rows and columns of the matrix correspond to men and women rather than jobs and applicants and a circle in position (i, j) means that man i and woman j are *compatible.* The so-called *marriage problem* then asks that we pair off the men and women in as many compatible pairs as possible.)

Returning to (4.2), the question is whether a *complete assignment*, i.e., assigning all five applicants, is possible. A little experimentation may convince the reader that there seems to be no way of filling all the jobs with qualified applicants. What is needed then is some sort of "proof" that in fact there is no complete assignment. Now it turns out that one can give such a proof, for notice that if one considers jobs J_1, J_3, and J_4 there are only two applicants A_1 and A_4 who qualify for them. It follows that there is no way of filling all three of these jobs because of the shortage of qualified applicants, hence a complete assignment is impossible.

The situation which occurs here is typical and very important for what follows. We will say that a simple assignment problem has a *bottleneck* if there is some set of r jobs for which fewer than r applicants are qualified (or in the marriage problem r women who are compatible with fewer than r men). In the example above the bottleneck consists of the set $\{A_1, A_4; J_1, J_3, J_4\}$. We now state the result which plays the key role in the method for solving assignment problems.

Main Theorem: *Every simple assignment problem has either a complete assignment or a bottleneck (but not both).*

The proof of the theorem will emerge as we present the solution method for the simple assignment problem which will be illustrated by means of the following example.

We are given below the qualification matrix for a 10 × 10 simple assignment problem.

(4.3)

	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}
A_1			0✓			0			0	
A_2	0✓		0		0					
A_3		0✓		0			0	0		0
A_4			0		0✓				0	
A_5	0		0			0✓			0	
A_6		0		0✓				0		0
A_7		0					0✓	0		
A_8		0		0	0		0			0
A_9	0					0			0	
A_{10}			0			0			0	

The objective is to find either a complete assignment or a bottleneck. The procedure for starting out is very simple. We run along the list of jobs and assign each to the first "available" applicant. Thus J_1 is assigned to A_2, J_2 to A_3, J_3 to A_1, J_4 to A_6. (Note that A_3 was also qualified for J_4 but is not available since he has already been assigned to J_2.) J_5 to A_4 (since A_2 is already assigned) J_6 to A_5 (why not to A_1?) J_7 to A_7 (why not to A_3?). These assignments have been indicated by checking the appropriate position in (4.3).

The situation changes abruptly when we seek to assign J_8 for we observe that all three qualified applicants A_3, A_6, and A_7 have already been assigned. This does not mean, however, that there is no way to assign J_8. We will see that it is possible to assign J_8 if jobs $J_1, \ldots, J_7$ are *reassigned* properly. The method for finding such a reassignment if it exists constitutes the heart of our computational procedure. It is called the *labeling method* because it involves attaching numerical labels to jobs and applicants in the following way:

1) Label with the numeral 0 the job to be assigned, in this case J_8. This is done by writing the label 0 at the bottom of column 8. We will refer to J_8 as the *0-job*.

2) Label with a 1 all applicants who qualify for the 0-job J_8, in this case A_3, A_6, and A_7. To do this we look for *circled positions* in column 8 and place a 1 at the right end of each row having a circle in column 8. These will be called *1-rows* and the corresponding applicants *1-applicants*.

3) Label with a 1 all jobs to which 1-applicants have been assigned. In this case the 1-applicants A_3, A_6, and A_7 have been assigned to J_2, J_4, and J_7. Accordingly we write a 1 at the bottom of columns 2, 4, and 7. This is done by searching all the 1-rows for *checked* positions and where we find one we enter the label 1 at the bottom of the corresponding column. The display with the labels is now

(4.4)

	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	labels
A_1			0✓			0			0		
A_2	0✓		0		0						
A_3		0✓		0			0	0		0	1
A_4			0		0✓				0		
A_5	0		0			0✓			0		
A_6		0		0✓				0		0	1
A_7		0					0✓	0			1
A_8		0		0	0		0			0	2
A_9	0					0			0		
A_{10}			0			0			0		
labels		1		1		1		0			

4) This is like step (2). Look for all applicants who qualify for 1-jobs and if *they are not already labeled* label them 2. In this case A_6, A_7, and A_8 qualify for the 1-job J_2 but A_6 and A_7 have already been labeled. However, A_8 is not yet labeled so it gets the label 2, etc. The exact computational procedure once again is to look for circled positions in the 1-columns and label the corresponding rows with 2 provided they are not already labeled.

5) We now make the important observation that in labelling A_8 we have labeled an applicant who has not been previously assigned. When this happens we say that *breakthrough* has occurred and this means that it is now possible to find a reassignment which includes J_8. The method is the following: A_8 has been labeled 2. This means he qualifies for some 1-job, in this case J_2, J_4, and J_7. Assign him to one of these, say J_2. Now J_2 is a 1-job which means it was previously assigned to some 1-applicant, in this case A_3. We therefore "unassign" A_3 from J_2. Finally A_3 being a 1-applicant qualifies for the 0-job J_8 so he is assigned to it. This gives the new assignment. The only changes are A_8 to J_2 and A_3 to J_8. The new display is then

(4.5)

	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	labels
A_1			0✓			0			0		
A_2	0✓		0		0						
A_3		0		0			0	0✓		0	1
A_4			0		0✓				0		
A_5	0		0			0✓			0		
A_6		0		0✓				0		0	1
A_7		0					0✓	0			2
A_8		0✓		0	0		0			0	1
A_9	0✓					0			0✓		
A_{10}			0			0			0		
labels		1		1			2	1		0	

We now proceed with the assignments noting that J_9 can be assigned to A_9. On the other hand J_{10} cannot be assigned, at least for the present, since all qualified applicants A_3, A_6, and A_8 are already assigned to other jobs. We proceed with the labeling method (refer to 4.5 for the picture).

1) J_{10} gets the label 0.

2) All applicants qualified for J_{10}, namely A_3, A_6, and A_8 get label 1.

3) The jobs assigned to 1-applicants A_3, A_6, and A_8, namely J_8, J_4, and J_2 get label 1.

4) All applicants qualified for 1-jobs J_2, J_4, and J_8 *not already labeled* get label 2. This turns out to be only A_7.

5) All jobs assigned to 2-applicants, namely J_7 get label 2.

6) All applicants qualified for 2-jobs, namely A_3, A_7, and A_8 and not already labeled get label 3. But in this case all three applicants are already labeled, hence no applicant gets label 3 and the labeling procedure *terminates*. Note also that breakthrough has not occurred for the only unassigned applicant, A_{10}, has not been labeled, so we get no reassignment. Instead however, we get a *bottleneck* consisting of the labeled jobs J_2, J_4, J_7, J_8, and J_{10} for which only the labeled applicants A_3, A_6, A_7, and A_8 are qualified as one easily verifies directly from the qualification matrix. We list this bottleneck in the form $\{A_3, A_6, A_7, A_8; J_2, J_4, J_7, J_8, J_{10}\}$. It follows that there does not exist a complete assignment for this matrix, and the problem is solved.

Let us now describe the labeling method in general without reference to specific examples. We assume that the first k jobs have been assigned so that there is a checked position in each of the first k columns. Then,

Step 0 Label column k+1 with 0. Call this the 0-column.

Step 1 Look for all circles in the 0-column and label the corresponding rows 1-rows. If one of these rows contains no check then breakthrough has occurred. Check the circle in this row and the 0-column. This means that J_{k+1} is now assigned. Otherwise there will be exactly one checked position in each 1-row. Label the corresponding column 1.

Step 2 Look for all circles in each 1-column and label the corresponding row 2 provided it has not already been labeled. If break-through occurs (i.e., a row with no check is labeled) a reassignment including J_{k+1} is found according to the method of the example. If not there is a checked position in each 2-row. Label the corresponding columns 2, and so on.

The procedure must terminate in one of two ways: (A) either breakthrough occurs in which case one can get a reassignment including J_{k+1} by the method of the preceeding example or (B) at some step the labeling terminates because all the new rows which are candidates for a label have already been labeled. In this case the claim is that there must be one more labeled column than labeled rows and the corresponding jobs and applicants form a bottleneck. To see why this is so note that at the end of each step there will be one more labeled column than row. This is certainly true at step 0 since there is one 0-column and no 0-row. From then on at each step as long as breakthrough does not occur we label the same number of columns as rows, for each time we label a new row it must contain a checked position (otherwise we would have breakthrough) and this provides a new column with the same label. Now when labeling terminates it means there are no new applicants qualified for the jobs already labeled, hence there is a shortage of one qualified applicant for the labeled jobs. This then provides the proof of the Main Theorem as well as a method for solving the simple assignment problem.

Exercises

14. For each of the following qualification matrices find either a complete assignment or a bottleneck. To specify a bottleneck list a set of r+1 jobs for which there are only r qualified applicants.

	0		0	0	
0		0	0		0
	0		0	0	
	0			0	
0	0	0		0	0
	0		0		

(i)

			0		0
0		0			
		0		0	
0	0				
			0		
	0				0

(ii)

0		0	0	0	0
	0			0	
	0			0	
		0	0	0	0
	0			0	
0	0	0			0

(iii)

15. We reproduce below the final display for the example of this section. The necessary labeling can be done right on this paper without having to recopy the display. Simply write in with pencil the necessary checks and labels and then erase them before going on to the next part of the problem.

	1	2	3	4	5	6	7	8	9	10	labels
1			0✓			0			0		
2	0✓		0		0						
3		0		0			0	0✓		0	
4			0		0✓				0		
5	0		0			0✓			0		
6		0		0✓				0		0	
7		0					0✓	0			
8		0✓		0	0		0			0	
9						0			0✓		
10			0			0			0		
labels											

Can the above assignment be extended to a complete assignment if an additional circle is introduced (i) in position (1,1), (ii) in position (1,2), (iii) in position 2,7)? In each case either give the assignment or list the bottleneck.

16. Fill in the labels from (4.5) in the display of the previous problem. Use the information they provide about bottlenecks to prove that if circles are added in any or all positions of row 3 it will still not be possible to make a complete assignment. Prove the same statement about column 1. Is the statement true of any other rows and columns? Which?

17. Give a brief argument proving the statement "but not both" of the Main Theorem.

18. In the problem of Exercise 15 there is a bottleneck $\{A_3, A_6, A_7, A_9; J_2, J_4, J_7, J_8, J_{10}\}$ meaning that there are too few applicants for the given set of jobs. Show that there is also a *job-bottleneck* by showing that A_1, A_2, A_4, A_5, A_8, and A_{10} are qualified only for J_1, J_3, J_5, J_6, and J_9 so that there are too few jobs for the given set of applicants. Show that this will always be true, i.e., that if a problem has an applicant-bottleneck it must also have a job-bottleneck. For each case in Exercise 15 where you found an applicant-bottleneck find a job-bottleneck.

---

5. The Optimal Assignment Algorithm

The method for solving the optimal assignment problem is now a matter of putting together the material of the two previous sections. We wish to find wages w_i and profits p_j for a given assignment matrix so that it will have a profit maximizing assignment, which by Theorem 2 will then be optimal. We will find the desired constants by solving a sequence of simple assignment problems. We proceed at once to illustrate the method using the third example of Section 1. The idea is to start out with any set of wages w_i and gradually change them until we get a profit-max assignment. A convenient starting point is to set all wages w_i equal to zero. For the example, the initial display is then

(5.1)

		profits		
		10	8	7
	0	⑩	⑧	⑦
wages	0	8	4	3
	0	9	6	4

The procedure is now to ignore the numbers and try to solve the simple assignment problem given by the circles in the above display. In this case, of course, the problem has no complete assignment since there are no circles in rows 2 and 3. The bottleneck in this case is obvious in that there is only one applicant A_1 who qualifies for the three jobs. Thus, there is a severe shortage of qualified labor or, in economic terms, the *demand* for qualified applicants is 3 since there are three job openings, whereas the *supply* of qualified applicants is only 1 since A_1 is the only qualified applicant. We now invoke the fundamental law of economics, the famous *law of supply and demand*, which asserts that if the demand for some good exceeds its supply then its price will rise (and conversely if the supply exceeds the demand the price will fall). For our case this means that the wage w_1 of A_1 must rise. Suppose then w_1 is increased to 1. The new display is then

(5.2)

	9	7	6
1	⑩	⑧	⑦
0	8	4	3
0	⑨	6	4

Notice that increasing w_1 by 1 decreases p_1, p_2, and p_3 by 1 and introduces a new circle in position (3,1). We can describe the operation just performed by the following:

Rule. *If the simple assignment problem has a bottleneck, increase the wages* w_i *of all* A_i *in the bottleneck,* (which will decrease the profits of all jobs J_j in the bottleneck) *by an amount such that at least one new circle appears in the display.*

Let us continue applying the rule. The display (5.2) again has no complete assignment and it has the obvious bottleneck $\{A_1, A_3; J_1, J_2, J_3\}$ since there is no circle in row 2. According to the rule, therefore we should increase w_1 and w_3 and decrease p_1, p_2 and p_3 by 1 for when we do this a new circle appears in position (2,1).

(5.3)

	8	6	5		
2	⑩	⑧	⑦	1	
0	⑧	4	3		labels
1	⑨	6	4		
		1	0		
		labels			

In the display we now go through the labeling process to locate the bottleneck $\{A_1; J_2, J_3\}$ so according to our rule we again increase the wage w_1 and lower p_2 and p_3 by 1 giving

(5.4)

	8	5	4
3	10	⑧	⑦
0	⑧	4	3
1	⑨	⑥	4

Notice that there is a new circle in position (3,2) and also the old circle in position (1,1) has disappeared. The checked positions now give the solution of the simple assignment problem (obtained by the labeling or any other method) and this, by our construction, is a profit-max hence an optimal assignment.

In this procedure the greatest chance for error is failing to fill the circles correctly. The reader should take special care after each application of the rule,

(A) to look for possible *new* circles which can occur only in positions (i,j) where J_j is in the bottleneck and A_i is not;

(B) to look for possible *disappearing* circles which can occur only in positions (i,j) where A_i is in the bottleneck and J_j is not.

Remember that a_{ij} is circled precisely when $a_{ij} = w_i + p_j$.

The appearance and disappearance of the circles has a natural economic meaning. A new circle appeared in position (3,2) above because it has now become profit maximizing to hire A_3 for J_2 in view of the increased wage of A_1. On the other hand the circle in position (1,1) disappears because it is no longer profit-maximizing to hire A_1 for J_1 in view of A_1's increased wage.

Let us now continue with the example, introducing the applicant A_4 with scores (2, 3, 5), and a dummy job J_4. We have

(5.5)

	8	5	4	?
3	10	⑧	⑦✓	0
0	⑧✓	4	3	0
1	⑨	⑥✓	4	0
?	2	3	5	0

One could, of course, start over again with this 4 × 4 problem. However, it is possible to take advantage of the work already done on the 3 × 3 problem provided one can choose suitable values for w_4 and p_4. A good rule is to choose these numbers as small as possible compatible with the feasibility conditions $w_i + p_j \geq a_{ij}$. This means that $w_4 = 1$, for $w_4 < 1$ would violate $w_4 + p_3 \geq a_{43}$, and $p_4 = 0$. This produces circles in positions (2,4) and (4,3).

profits

(5.6) wages

	8	5	4	0	
3	10	⑧	⑦✓	0	3
0	⑧✓	4	3	⓪	1
1	⑨	⑥✓	4	0	2
1	2	3	⑤	0	4
	1	2	3	0	

labels

labels

We proceed to solve the new simple assignment problem by the labeling method which this time leads to breakthrough when A_4 is labeled, and we get the optimal assignment [2, 4, 1, 3]. (The reader should work through the details.)

The problem has now been solved. There is however one last important step which provides an independent *check* on the correctness of the solution. First calculate the value of the assignment which is

(5.7) $8✓ + 0✓ + 9✓ + 5✓ = 22.$

Then compute the sum of all wages and profits which is

(5.8) $(3 + 0 + 1 + 1) + (8 + 5 + 4 + 0) = 22.$

The reason that these two numbers are equal is clear because one only assigns on circled entries, that is, on entries for which $a_{ij} = w_i + p_j$ and since each of the numbers i and j occur exactly once in a complete assignment it follows that the sum of the checked a_{ij} is the same as the sum of the w_i and p_j. Finally, if there is any doubt about the correctness of the answer we should be sure that all the feasibility conditions $a_{ij} \leq w_i + p_j$ are satisfied.

As a final illustration we go through another slightly larger example. The assignment matrix is

Profits

wages

	13	10	10	11	11
0	12	9	⑩	8	⑪
0	8	6	6	5	9
0	⑬	⑩	⑩	⑪	⑪
0	6	2	4	3	5
0	11	7	⑩	9	⑪

and we have taken initial wages to be 0 giving the display as shown.

The display has the obvious bottleneck $\{A_1, A_3, A_5; J_1, J_2, J_3, J_4, J_5\}$ since there are no circles in rows 2 and 4. Applying the rule we can raise w_1, w_3, and w_5 by 2 and lower all profits by 2 producing a new circle in position (2,5).

Profits

Wages

	11	8	8	9	9
2	12	9	⑩	8	⑪
0	8	6	6	5	⑨
2	⑬	⑩	⑩	⑪	⑪
0	6	2	4	3	5
2	11	7	⑩	9	⑪

(The student in making these calculations should definitely not recopy the assignment matrix at each stage. He should always work with the original matrix changing only the w_i and p_j and adding or erasing circles in going from one stage to the next.) The new display shown above again has the obvious bottlenock $\{A_1, A_2, A_3, A_5; J_1, J_2, J_3, J_4, J_5\}$ (since there are no circles in row 4). Raising w_1, w_2, w_3, w_5 by 4 and lowering all p_j by 4 brings in circles in position (4,3) and (4,5) as shown

profits

wages

	7	4	4	5	5
6	12	9	⑩✓	8	⑪
4	8	6	6	5	⑨✓
6	⑬✓	⑩	⑩	⑪	⑪
0	6	2	④	3	⑤
6	11	7	⑩	9	⑪

If one starts the labeling procedure the bottleneck $\{A_3; J_1, J_2, J_4\}$ becomes apparent at once. Increasing w_3 by 1 and decreasing p_1, p_2, and p_4 by 1:

(A) produces new circles at positions (1,1), (1,2) and (4,1)

(B) removes circles at (3,3), and (3,5) giving

profits

wages

labels

	6	3	4	4	5	
6	⑫	⑨✓	⑩	8	⑪	2
4	8	6	6	5	⑨✓	
7	⑬	⑩	10	⑪✓	11	1
0	⑥✓	2	④	3	⑤	3
6	11	7	⑩✓	9	⑪	4
	2	1	3	0	3	

labels

The labeling procedure then produces the final assignment [2, 5, 4, 1, 3].

Check 6✓ + 9✓ + 10✓ + 11✓ + 9✓ = 45 and

$$\underbrace{(6 + 4 + 7 + 0 + 6)}_{\text{wages}} + \underbrace{(6 + 3 + 4 + 4 + 5)}_{\text{profits}} = 23 + 22 = 45.$$

6. Why The Method Works

In the examples of the preceding section it always turned out that after a finite number of steps involving increasing various wages we arrived at a complete assignment for the corresponding simple assignment problem, hence an optimal assignment. Now from the rule for changing wages it is clear that at each stage new circles will appear in the display, but unfortunately it may also happen that old circles may disappear. It is therefore at least conceivable that circles keep appearing and disappearing in such a way that there are never enough of them to make a complete assignment and so the algorithm will go on forever. To show that this cannot occur we use a different approach. Instead of trying to keep track of circles we make use of the feasibility condition

$$w_i + p_j \geq a_{ij} \text{ for all positive } (i,j).$$

At each stage of the calculation let

$$z = w_1 + w_2 + \ldots + w_n + p_1 + p_2 + \ldots + p_n.$$

Now at every stage of the calculation if a complete assignment is not found we get a bottleneck which leads us to raise r of the wages w_i by some amount and lower at least r+1 of the profits by this same amount. This means that the number z will strictly decrease by at least 1 from one stage to the next. On the other hand, by the feasibility condition the value of z can never be smaller than, for example, $a_{11} + a_{22} + \ldots + a_{nn}$ because $w_1 + p_1 \geq a_{11}$, $w_2 + p_2 \geq a_{22}, \ldots, w_n + p_n \geq a_{nn}$. But since z decreases every time the display contains a bottleneck it follows that eventually we must get a display with no bottleneck and hence by The Main Theorem a display with a complete assignment.

Exercises

NOTE: In all problems where you are asked to solve a numerical problem your solution should present the final display with the w_i, p_j and checked and circled entries of the matrix. In actually making the calculations the student should work with pencil and eraser on these sheets recopying only the final display to hand in.

19. Find the optimal assignment for the matrix

$$\begin{array}{|c|c|c|c|}\hline 6 & 5 & 2 & 0 \\ \hline 4 & 2 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 \\ \hline 4 & 3 & 1 & 0 \\ \hline \end{array}$$

20. Below is the final display for the example of Section 5. Insert the suitable circles and checks in pencil without referring back to Section 5.

$$\begin{array}{c|c|c|c|c|c|}\cline{2-6} & 6 & 3 & 4 & 4 & 5 \\ \hline \multicolumn{1}{|c|}{6} & 12 & 9 & 10 & 8 & 11 \\ \hline \multicolumn{1}{|c|}{4} & 8 & 6 & 6 & 5 & 9 \\ \hline \multicolumn{1}{|c|}{7} & 13 & 10 & 10 & 11 & 11 \\ \hline \multicolumn{1}{|c|}{0} & 6 & 2 & 4 & 3 & 5 \\ \hline \multicolumn{1}{|c|}{6} & 11 & 7 & 10 & 9 & 11 \\ \hline \end{array}$$

Now a sixth applicant becomes available with scores (5, 2, 5, 4, 6). Add this and a dummy job (in pencil) in the display above and find the new optimal assignment. (Do your calculations on this sheet recopying only your final display.)

21. Same as Exercise 20 except the scores of the new applicant are (4, 3, 5, 5, 4).

22. Referring to the display of Exercise 20 prove without making any calculations that if any or all uncircled entries are increased by 1 the checked assignment will remain optimal. Show that this is also true if any checked entry is increased by any amount.

23. Referring to the same display increase the entry in position (1,1) from 12 to 13. Which numbers w_i or p_j must be changed to maintain feasibility? After making the change which new circles appear? Which old ones disappear? Find the optimum for this altered problem.

24. Same as Exercise 23 except that

(i) the 9 in position (1,2) is decreased to 8

(ii) the 6 in position (2,2) is increased to 8.

In each case do not start the problem all over from the beginning but adjust the given w_i and p_j so as to restore feasibility and go on from there.

25. Solve the optimal assignment problem whose matrix is given below. Time yourself!

7	6	9	8	5	5	9	6
8	8	8	8	6	7	9	7
5	4	6	5	3	5	7	4
9	8	10	7	7	6	9	6
10	9	10	9	7	8	10	8
5	4	5	6	4	5	6	5
7	7	8	8	5	7	8	6
10	9	10	7	8	9	7	9

26. An additional applicant with scores (1, 2, 3, 4, 5, 4, 3, 2) becomes available. Starting from your solution to Exercise 25 obtain the new optimal assignment.

7. Answers to Exercises

Section 1

1. n!

2. [1 2 3 4] has value 26.
[4 3 2 1] has value 28.
[3 1 2 4] has value 27.
[2 3 4 1] has value 27.

3. The checked assignment is [4 1 3 2] with value 27. The circled entries do not constitute an assignment, since there are two circled entries in row 3 (and no circled entries in row 2).

Section 2

4. The 3! = 6 possible assignments, and their values are:
[1 2 3] v = 9
[1 3 2] v = 8
[2 1 3] v = 10
[2 3 1] v = 7
[3 1 2] v = 8
[3 2 1] v = 6
Therefore [2 1 3] is optimal.

5. Even if a_{11} is increased by 1, the optimal value will not increase (although there will some new optimal assignments). The value of the optimal assignment will increase if a_{12}, a_{21}, or a_{33} is increased.

6. Likewise, a decrease in a_{12}, a_{21}, or a_{33} will lead to a decreased value of the optimal assignment.

7. Yes. Replace A_2 with A_4. The new matrix is:

$$\begin{array}{c|c|c|c|}
 & J_1 & J_2 & J_3 \\
\hline
A_1 & \textcircled{6} & 5 & 2 \\
\hline
A_2 & 1 & \textcircled{4} & 0 \\
\hline
A_3 & 2 & 2 & \textcircled{1} \\
\hline
\end{array}$$

and the assignment [1 2 3] has value 11.

8. In this case, no improved assignment is possible.

9. Choose $r = 2(n-1)$. Form the following matrix:

$$\begin{bmatrix} r & r-1 & r-2 & r-3 & \cdots & r-n+1 \\ r-1 & r-2 & \cdots & & & r-n \\ \vdots & \vdots & & & & \vdots \\ r-n+1 & r-n & \cdots & & & r-2n+2=0 \end{bmatrix}.$$

For example, $n = 3$, $r = 4$, and the matrix is $\begin{bmatrix} 4 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 0 \end{bmatrix}$

Or, for $n = 5$, $r = 8$ and we get $\begin{bmatrix} 8 & 7 & 6 & 5 & 4 \\ 7 & 6 & 5 & 4 & 3 \\ 6 & 5 & 4 & 3 & 2 \\ 5 & 4 & 3 & 2 & 1 \\ 4 & 3 & 2 & 1 & 0 \end{bmatrix}$.

In general, we have a matrix in which each applicant is best qualified for job 1 and each job is done best by applicant 1. Since in any assignment, including the optimal assignment, exactly one person does job 1 and applicant 1 does only one job, we see that only one person does the job for which he is most qualified and only one job is done by the most qualified applicant.

10. This is impossible.

Section 3

11.

p_j w_i	4	3	1
2	⑥	⑤✓	2
0	④✓	2	0
0	2	2	①✓

12.

p_j w_i	5	4	1	0
1	⑥✓	⑤	②	0
0	4	2	0	⓪✓
0	2	2	①✓	⓪
0	1	④✓	0	⓪

p_j w_i	4	3	1	0
2	⑥	⑤✓	2	0
0	④✓	2	0	⓪
0	2	2	①✓	⓪
0	④	③	①	⓪✓

13.

p_j w_i	7	6	6	5
2	⑨	⑧	⑧	⑦✓
1	⑧	⑦	⑦✓	5
1	7	⑦✓	6	5
0	⑦✓	⑥	5	4

Section 4

14. (i) There is a bottleneck $\{A_2, A_5; J_1, J_3, J_6\}$.
(ii) There is an assignment {6, 3, 5, 1, 4, 2}.
(iii) There is a bottleneck $\{A_1, A_4, A_6; J_1, J_3, J_4, J_6\}$.

15. (i) No. There is still the bottleneck $\{A_3, A_6, A_7, A_8; J_2, J_4, J_7, J_8, J_{10}\}$.
(ii) Yes. An assignment now is {2, 1, 8, 5, 6, 4, 7, 10, 9, 3}.
(iii) Yes. One assignment is {6, 7, 10, 5, 1, 4, 8, 2, 9, 3}. Another possibility is {6, 7, 8, 5, 1, 4, 2, 10, 9, 3}.

16. We have the bottleneck $\{A_3, A_6, A_7, A_8; J_2, J_4, J_7, J_8, J_{10}\}$. This means that the only qualified applicants for jobs 2, 4, 7, 8, and 10 are applicants 3, 6, 7, and 8. This will *still* be true even if applicant 3 becomes qualified for other jobs (i.e. - put 0's in row 3) so there will be a bottleneck and no assignment. Likewise, even if other applicants become qualified for job 1 (i.e. - put 0's in column 1) the bottleneck will remain. Additional 0's may be placed in rows 3, 6, 7 or 8 and in columns 1, 3, 5, 6 or 9, and the bottleneck will remain.

In general, 0's may be placed in labeled rows or unlabeled columns without affecting the bottleneck.

17. When breakthrough occurs, there are at least as many labeled rows as labeled columns. When a bottleneck occurs, there is one more labeled column than labeled rows. Hence, both breakthrough and a bottleneck cannot occur simultaneously.

18. One can simply verify from the matrix the existence of the job bottleneck described. In general, suppose there is an applicant bottleneck where the only qualified applicants for some set of jobs J are the set of applicants A, where J has one more element than A. Then consider the sets $\overline{J}$ and $\overline{A}$, the complements of J and A. Since the total number of applicants is equal to the total number of jobs, we have that $\overline{A}$ has one more element than $\overline{J}$. Moreover, the only jobs for which an applicant in $\overline{A}$ is qualified are in $\overline{J}$ (for if the applicant were qualified for a job in J, he would be in A). Thus, the sets $\{\overline{A}; \overline{J}\}$ form a job - bottleneck. Note we could also prove the existence of a job - bottleneck by applying our algorithms to the transposed matrix. In # 1 i) and iii) we have the following job bottlenecks:

i) $\{A_1, A_3, A_4, A_6; J_2, J_4, J_5\}$.
iii) $\{A_2, A_3, A_5; J_2, J_5\}$.

Section 6

19.

	4	3	1	0
2	⑥✓	⑤	2	0
0	④	2	0	⓪✓
0	2	2	①✓	⓪
0	④	③✓	①	⓪

v = 10

20.

	5	2	4	3	5	-1
7	⑫	⑨✓	10	8	11	0
4	8	⑥	6	5	⑨✓	0
8	⑬	⑩	10	⑪✓	11	0
1	⑥✓	2	4	3	5	⓪
6	⑪	7	⑩✓	⑨	⑪	0
1	5	2	⑤	④	⑥	⓪✓

v = 45

21.

	6	3	4	4	5	0
6	⑫✓	⑨	⑩	8	⑪	0
4	8	6	6	5	⑨✓	0
7	⑬	⑩✓	10	⑪	11	0
0	⑥	2	④	3	⑤	⓪✓
6	11	7	⑩✓	9	⑪	0
1	4	3	⑤	⑤✓	4	0

v = 46

22. By the feasibility conditions, $w_i + p_j \geq a_{ij}$ and if a_{ij} is not circled, then $w_i + p_j > a_{ij}$. Hence, if 1 is added to an uncircled entry in position (i, j), we have $w_i + p_j \geq a_{ij} + 1$ and thus the feasibility conditions are still satisfied for these w_i and p_j. Hence, the original checked assignment is still optimal.

Suppose the checked entry in row i is raised by an amount $d_i \geq 0$. Define $w_i^* = w_i + d_i$. Since we had $w_i + p_j \geq a_{ij}$ then certainly $w_i^* + p_j \geq a_{ij}$. Moreover, if entry (i, j) was checked and had d_i added to it, then $w_i^* + p_j = a_{ij} + d_i$. That is, the w_i^* and p_j satisfy the feasibility conditions, and the original checked entries will be circled. Since they form an assignment, it is the optimal assignment.

23. Suppose a_{ij} is changed from 12 to 13. We could change w_i from 6 to 7 (and this would eliminate circles in the (1, 2) (1 ,3) and (1, 5) positions) or we could change p_i from 6 to 7 (which would eliminate circles in the (3, 1) and (4, 1) positions). The original solution is found to still be optimal, but there are alternate optimal solutions now as well.

24. i.

	6	2	4	3	5
6	⑫✓	⑧	⑩	8	⑪
4	8	⑥✓	6	5	⑨
8	13	⑩	10	⑪✓	11
0	⑥	②	④✓	③	⑤
6	11	7	⑩	⑨	⑪✓

v = 44

ii.

	6	4	4	4	5
6	⑫✓	9	⑩	8	⑪
4	8	⑧✓	6	5	⑨
7	⑬	10	10	⑪✓	11
0	⑥	2	④✓	3	⑤
6	11	7	⑩	9	⑪✓

v = 46

25.

	6	6	7	6	4	5	7	5
2	7	6	⑨✓	⑧	5	5	⑨	6
2	⑧	⑧	8	⑧	⑥	⑦	⑨	⑦✓
0	5	4	6	5	3	⑤	⑦✓	4
3	⑨✓	8	⑩	7	⑦	6	9	6
3	10	⑨✓	⑩	⑨	⑦	⑧	⑩	⑧
0	5	4	5	⑥	④✓	⑤	6	⑤
2	7	7	8	⑧✓	5	⑦	8	6
4	⑩	9	10	7	⑧	⑨✓	7	⑨

v = 62

26.

	6	6	7	6	4	5	7	5	0
2	7	6	⑨✓	⑧	5	5	⑨	6	0
2	⑧	⑧	8	⑧	⑥	⑦	⑨	⑦✓	0
0	5	4	6	5	3	⑤	⑦✓	4	⓪
3	⑨✓	8	⑩	7	⑦	6	9	6	0
3	10	⑨✓	⑩	⑨	⑦	⑧	⑩	⑧	0
0	5	4	5	⑥	[illegible]	⑤	6	⑤	⓪✓
2	7	7	8	⑧✓	5	⑦	8	6	0
4	⑩	9	10	7	⑧	⑨✓	7	⑨	0
1	1	2	3	4	⑤✓	4	3	2	0

v = 63

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 1

Request for Help

<u>Student</u>: If you have trouble with a specific part of this unit, please fill out this form and take it to your instructor for assistance. The information you give will help the author to revise the unit.

Your Name______________________________________ Unit No._______

Page______
○ Upper
○ Middle
○ Lower

OR

Section_________
Paragraph________

OR

Model Exam Problem No._____
Text Problem No._____

Description of Difficulty: (Please be specific)

<u>Instructor</u>: Please indicate your resolution of the difficulty in this box.

○ Corrected errors in materials. List corrections here:

○ Gave student better explanation, example, or procedure than in unit. Give brief outline of your addition here:

○ Assisted student in acquiring general learning and problem-solving skills (not using examples from this unit.)

Instructor's Signature____________________________

Please use reverse if necessary.

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 2

Unit Questionnaire

Name________________________________Unit No.____________Date___________

Institution______________________________Course No.______________________

Check the choice for each question that comes closest to your personal opinion.

1. How useful was the amount of detail in the unit?

____Not enough detail to understand the unit
____Unit would have been clearer with more detail
____Appropriate amount of detail
____Unit was occasionally too detailed, but this was not distracting
____Too much detail; I was often distracted

2. How helpful were the problem answers?

____Sample solutions were too brief; I could not do the intermediate steps
____Sufficient information was given to solve the problems
____Sample solutions were too detailed; I didn't need them

3. Except for fulfilling the prerequisites, how much did you use other sources (for example, instructor, friends, or other books) in order to understand the unit?

____A Lot ____Somewhat ____A Little ____Not at all

4. How long was this unit in comparison to the amount of time you generally spend on a lesson (lecture and homework assignment) in a typical math or science course?

Much	Somewhat	About	Somewhat	Much
____Longer	____Longer	____the Same	____Shorter	____Shorter

5. Were any of the following parts of the unit confusing or distracting? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Paragraph headings
____Examples
____Special Assistance Supplement (if present)
____Other, please explain____

6. Were any of the following parts of the unit particularly helpful? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Examples
____Problems
____Paragraph headings
____Table of Contents
____Special Assistance Supplement (if present)
____Other, please explain____

Please describe anything in the unit that you did not particularly like.

Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)

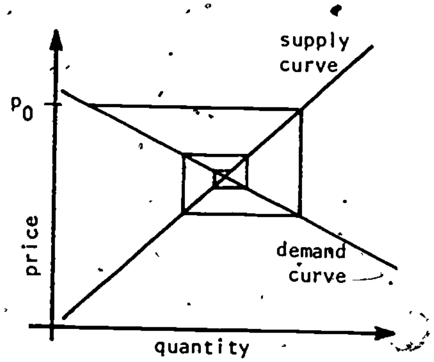
umap

UNIT 322

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT

DIFFERENCE EQUATIONS WITH APPLICATIONS

by Donald R. Sherbert



APPLICATIONS OF DIFFERENCE EQUATIONS
TO ECONOMICS AND SOCIAL SCIENCES

edc/umap /55chapel st./newton, mass. 02160

DIFFERENCE EQUATIONS WITH APPLICATIONS*

by

Donald R. Sherbert
Department of Mathematics
University of Illinois
Urbana, Illinois 61801

TABLE OF CONTENTS

---

*This material is based on work supported by the National Science Foundation under CAUSE Grant No. SER 78-06381 to the University of Illinois at Urbana-Champaign.

Intermodular Description Sheet: UMAP Unit 322

Title: DIFFERENCE EQUATIONS WITH APPLICATIONS

Author: Donald R. Sherbert
Department of Mathematics
University of Illinois
Urbana, Illinois 61801

Review Stage/Date: III 5/30/79

Classification: APPL DIFF EQ ECON/SOC SCI

Prerequisite Skills:

1. Familiarity with solution of linear and quadratic equations.
2. An understanding of complex numbers (for Section 3.2).
3. Knowledge of basic algebraic skills.

Output Skills:

1. To learn how to interpret and solve elementary difference equations.
2. To understand how difference equations can be used to model certain problems in economics and social science.

Other Related Units:

The Dynamics of Political Mobilization I *(Unit 297)*
The Dynamics of Political Mobilization II *(Unit 298)*
Public Support for Presidents I *(Unit 299)*
Public Support for Presidents II *(Unit 300)*
Diffusion of Innovation in Family Planning *(Unit 303)*
Growth of Partisan Support I: Model and Estimation *(Unit 304)*
Growth of Partisan Support II: Model Analytics *(Unit 305)*
Discretionary Review by Supreme Court I *(Unit 306)*
Discretionary Review by Supreme Court II *(Unit 307)*
Budgetary Process: Competition *(Unit 332)*
Budgetary Process: Incrementalism *(Unit 333)*

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT (UMAP)

The goal of UMAP is to develop, through a community of users and developers, a system of instructional modules in undergraduate mathematics and its applications which may be used to supplement existing courses and from which complete courses may eventually be built.

The Project is guided by a National Steering Committee of mathematicians, scientists and educators. UMAP is funded by a grant from the National Science Foundation to Education Development Center, Inc., a publicly supported, nonprofit corporation engaged in educational research in the U.S. and abroad

PROJECT STAFF

Ross L. Finney	Director
Solomon Garfunkel	Associate Director/Consortium Coordinator
Felicia DeMay	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Dianne Lally	Project Secretary
Paula M. Santillo	Administrative Assistant
Carol Forray	Production Assistant
Zachary Zevitas	Staff Assistant

NATIONAL STEERING COMMITTEE

W.T. Martin	MIT (Chairman)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
Ernest J. Henley	University of Houston
William Hogan	Harvard University
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
R. Duncan Luce	Harvard University
George Miller	Nassau Community College
Frederick Mosteller	Harvard University
Walter E. Sears	University of Michigan Press
George Springer	Indiana University
Arnold A. Strassenburg	SUNY at Stony Brook
Alfred B. Willcox	Mathematical Association of America

The Project would like to thank Frances M. Thompson of the Okefenokee Cooperative Educational Services Agency, and J.W. Goldston of Pembroke State University for their reviews, and all others who assisted in the production of this unit.

This material was prepared with the support of National Science Foundation Grant No. SED76-19615 A02. Recommendations expressed are those of the author and do not necessarily reflect the views of the NSF, nor of the National Steering Committee.

DIFFERENCE EQUATIONS WITH APPLICATIONS

1. INTRODUCTION

The main theme of difference equations is that of recursion: computations performed in a recurrent or repeated manner. In fact, difference equations are sometimes referred to as recursion relations. We begin by looking at a familiar sequence from the viewpoint of recursion.

The sequence of numbers $\{1,2,2^2,2^3,\ldots,2^n,\ldots\}$ is a geometric progression with geometric ratio 2. Each term in this sequence, with the exception of the initial term, is obtained from its predecessor by multiplying the predecessor by the number 2. If we introduce the notation $x_n = 2^n$ for $n = 0,1,2,\ldots$, then x_{n+1} and x_n are related by the equation

$$x_{n+1} = 2x_n \tag{1.1}$$

This relation by itself does not uniquely determine the sequence under consideration because it does not provide any information about the value of the initial term x_0. However, once the value

$$x_0 = 1 \tag{1.2}$$

is specified, then the geometric progression is completely determined because there is a starting point (1.2) and a rule (1.1) for calculating each term from the preceding term. That is, beginning with $x_0 = 1$, we get $x_1 = 2x_0 = 2$, then $x_2 = 2x_1 = 2^2$, then $x_3 = 2x_2 = 2^3$, and so on.

Viewed in this manner, the above sequence is said to be determined recursively because the calculation of a particular term is done by a chain of calculations with successive terms linked by (1.1). The relation (1.1) is called a *difference equation* or a recursion relation. The condition (1.2) is called an *initial condition* or boundary condition. By a *solution* to a difference equation is meant a formula for x_n such that x_n may be computed directly without going through a chain of calculations.

In the above example, we started with a given sequence and then described it in the form of a difference equation and initial condition. In what follows, we shall be concerned with the opposite problem. That is, we shall be confronted with a difference equation and wish to solve it in the sense of finding a formula that yields the n^{th} term directly. The following famous sequence of numbers first introduced in the year 1202 by the Italian mathematician known as Fibonacci illustrates this problem.

The sequence of *Fibonacci numbers* is defined recursively as follows. The first two Fibonacci numbers are

$$x_0 = 1,\ x_1 = 1, \tag{1.3}$$

and the remaining numbers are prescribed by the equation

$$x_{n+2} = x_{n+1} + x_n, \quad n = 0,1,2,\ldots. \tag{1.4}$$

Thus the n^{th} Fibonacci number for $n \geq 2$ is obtained by adding the two preceding Fibonacci numbers. The first ten Fibonacci numbers are therefore 1, 1, 2, 3, 5, 8, 13, 21, 34, 55. The formula that gives the n^{th} term directly is by no means obvious. The solution will be derived later after some techniques for solving difference equations have been developed.

The difference equation (1.1) for the geometric progression is a *first order equation* since only one term is needed to obtain the next term. The difference equation (1.4) for the Fibonacci numbers is a *second order equation* since two terms are needed to calculate the next term. Both of these equations are *linear* since the terms of the sequences are not multiplied together or raised to powers.

A general linear difference equation of order r would have the form

$$(1.5)\qquad c_0x_{n+r} + c_1x_{n+r-1} + \dots + c_rx_n = f_n$$

where the values f_n are given along with the coefficients $c_0, c_1, \dots, c_r$. However, in the following discussion we shall restrict our attention to linear difference equations of first and second order. Examples illustrating how difference equations arise in applications will also be given.

Exercises

1. For each of the following sequences, find a difference equation and initial condition that uniquely determines the given sequence.

   a. $3, 6, 9, 12, 15, \dots, 3n, \dots$

   b. $3, 9, 27, 81, \dots, 3^n, \dots$

   c. $2, 5, 9, 17, \dots, 2^n+1, \dots.$

2. Suppose $x_0 = 1$ and $x_{n+1} = 2x_n+1$ for $n = 1,2,\dots.$ Find the values of $x_1, x_2, \dots, x_{10}$.

3. Suppose an initial population of 6 wombats triples each 2 years. Find the population at the end of 14 years.

4. Suppose the recursion relation $x_{n+2} = x_{n+1} + x_n$ has initial conditions $x_0 = 7$, $x_1 = -4$. Find the first ten terms.

---

2. FIRST ORDER DIFFERENCE EQUATIONS

In this section we examine difference equations of the form

$$c_0x_{n+1} + c_1x_n = f_n$$

where c_0, c_1 are nonzero constants and $\{f_n\}$ is given. By dividing by c_0 and changing a sign, we can rewrite this difference equation as

$$(2.1)\qquad x_{n+1} - bx_n = g_n$$

where $b = -c_1/c_0$ and $g_n = f_n/c_0$. We are interested in finding all solutions of (2.1) and in examining the role of initial conditions for these difference equations.

The *associated homogeneous* difference equation is

$$(2.2)\qquad x_{n+1} - bx_n = 0.$$

That is, $g_n = 0$ for all n. If $g_n \neq 0$ for some values of n then Equation (2.1) is called *nonhomogeneous*. Equations (2.1) and (2.2) are closely related. We first concentrate on the homogeneous case (2.2).

2.1 Homogeneous Equations

The homogeneous equation (2.2), which can be written $x_{n+1} = bx_n$, is easy to solve. Starting with the equation corresponding to n = 0, we have

$$x_1 = bx_0.$$

Next, x_2 is obtained by substituting this expression for x_1 into the equation for n = 1. That is,

$$x_2 = bx_1 = b(bx_0) = b^2x_0.$$

Moving to the equation for n = 2, we get

$$x_3 = bx_2 = b(b^2x_0) = b^3x_0.$$

Continuing in this manner, we find

$$(2.3)\qquad x_n = b^nx_0, \quad n = 0,1,2,\dots.$$

This conclusion can be formally established by mathematical induction if desired.

If we set $x_0 = C$, then we have found that any solution of the linear homogeneous difference equation

$$(2.4)\qquad x_{n+1} - bx_n = 0$$

has the form

$$(2.5) \qquad x_n = Cb^n, \quad n = 0,1,2,....$$

This is referred to as the *general solution* of the difference equation. Note that a solution is not uniquely determined unless a value of $C = x_0$ is specified.

For example, the general solution of the homogeneous difference equation $x_{n+1} - 5x_n = 0$ is $x_n = C5^n$ where $C = x_0$. If the condition $x_0 = 2$ is imposed, then we have the solution $x_n = 2 \cdot 5^n$. On the other hand, if we are told that $x_3 = 750$, then $750 = C \cdot 5^3 = 125C$ gives us $C = 6$ so that the solution $x_n = 6 \cdot 5^n$ results.

Example 2.1 (Compound interest). If an initial amount of money P (for principal) is put into an account that bears 6 percent interest per year, then at the end of one year the total amount in the account is $P + 0.06P = 1.06P$. If interest is compounded annually, then this new amount accrues interest during the second year. Thus, at the end of the second year the total amount in the account is $1.06P + (0.06)1.06P = (1.06)^2P$, and so on.

The compounding gives rise to a difference equation as follows. Let x_n denote the amount in the account at the end of the n^{th} year. The relation between x_{n+1} and x_n is then

$$x_{n+1} = x_n + (0.06)x_n = 1.06x_n.$$

That is, the new amount is the old amount x_n plus the interest $0.06x_n$ on that amount. In this situation, $x_0 = P$ is the initial amount in the account. This difference equation has the form $x_{n+1} - bx_n = 0$ with $b = 1.06$. Therefore, the amount in the account at the end of the n^{th} year is

$$x_n = (1.06)^nP.$$

For a general rate of interest r, the amount is

$$(2.6) \qquad x_n = (1+r)^nP.$$

For example, if 200 dollars is put into an account bearing 5 percent interest, then at the end of 10 years there will be

$$x_{10} = (1.05)^{10}200 = 325.78$$

dollars in the account.

In many accounts, compounding is done several times per year. Suppose an account is advertised as having an annual interest rate r compounded monthly. This means that there are 12 interest periods during the year and compounding is done each interest period at an interest rate r/12. The corresponding difference equation has the same form, but now n refers to the number of interest periods and the rate of interest is r/12 instead of r.

For example, suppose D = 200 dollars is put into an account with interest r = 0.05 compounded monthly. Then the monthly interest rate is r/12 = 0.00417 and at the end of 10 years there will have been 120 interest periods. Thus, the account will contain

$$\begin{aligned} x_{120} &= 1.00417^{120}200 \\ &= 329.53 \end{aligned}$$

dollars after 10 years.

Example 2.2 (Population growth). In a population of animals, insects, bacteria, and so on, that has no disturbances to retard population growth, it is reasonable to assume that the *rate* at which the population grows depends only on the size of the population at any time. That is, the number capable of reproduction determines the population rate. Stating this hypothesis precisely leads to a difference equation. First, we consider the size of the population during a sequence of equal time periods; this may be years, months, or minutes depending on the nature of the population. Now let x_n be the population at the beginning of the n^{th} time period. Then the change in size of the population during that time period is $x_{n+1} - x_n$. This population growth is assumed to be proportional to the population at the beginning of the time period, that is,

$$x_{n+1} - x_n = ax_n$$

where a is a constant of proportionality. Thus, we have a first order difference equation

$$x_{n+1} - (1+a)x_n = 0.$$

We see that population grows under these assumptions like money in a compound interest account. In a sense, ignoring population retardants is like ignoring taxes.

The solution of the difference equation is

$$x_n = (1+a)^n P \quad n = 1,2,3,\ldots \tag{2.7}$$

where P is the initial population. For example, if a colony of 100 rabbits increases each month at a rate of 50 percent, then P = 100 and a = 0.50 so that at the end of one year, the population will be

$$\begin{aligned} x_{12} &= (1.5)^{12}100 \\ &= 12,974. \end{aligned}$$

Exercises

1. Find the general solution of the following:

   a. $x_{n+1} - 5x_n = 0$

   b. $3x_{n+1} = 2x_n$

   c. $x_{n+1} = x_n$.

2. Solve the following:

   a. $x_{n+1} = 5x_n$, $x_0 = 1$

   b. $5x_{n+1} + 3x_n = 0$, $x_5 = 3$.

3. A sum of $1,000 is invested at 8 percent interest compounded quarterly. When does the investment double?

4. A man inherits $1,000,000 at age 20, invests it at 6 percent interest compounded annually, and spends 10 percent of the amount each year. If he lives to be 70, how much will his son inherit?

5. The population of a city increases by 25 percent each year. If the population was 100,000 in 1970, what was it in 1950?

6. Radium decays at the rate of 1 percent every 25 years. Let r_n be the amount of radium left after n of the 25-year periods, where r_0 is the initial amount. Find a formula for r_n. How much is left after 100 years? How long does it take for half of it to decay?

2.2 Nonhomogeneous Equations

Let us turn to the nonhomogeneous difference equation

$$x_{n+1} - bx_n = g_n \tag{2.8}$$

where the g_n's may be nonzero. Instead of attempting a direct iterative procedure as before, we shall examine the relation between solutions of this nonhomogeneous equation and the associated homogeneous equation.

Suppose that $\{x_n^{(h)}\}$ is *any* solution of the associated homogeneous equation, so that

$$x_{n+1}^{(h)} - bx_n^{(h)} = 0; \tag{2.9}$$

and suppose that $\{x_n^{(p)}\}$ is one *particular* solution of the nonhomogeneous equation, so that

$$x_{n+1}^{(p)} - bx_n^{(p)} = g_n. \tag{2.10}$$

Adding these two equations together and arranging terms gives us

$$(2.11) \qquad \left[x_{n+1}^{(h)} + x_{n+1}^{(p)}\right] - b\left[x_n^{(h)} + x_n^{(p)}\right] = g_n .$$

This shows us that $x_n^{(h)} + x_n^{(p)}$ is also a solution of the nonhomogeneous equation. The interesting feature is that *every solution of the nonhomogeneous equation can be obtained in this way.*

To see this, let $\{x_n\}$ be the general solution of the nonhomogeneous equation, and let $\{x_n^{(p)}\}$ be a particular solution. That is, $\{x_n^{(p)}\}$ is an explicit solution containing no undetermined coefficient. A particular solution can be thought of as corresponding to a specific initial condition. If we now define $\{x_n^{(h)}\}$ by setting

$$x_n^{(h)} = x_n - x_n^{(p)}$$

we obtain

$$\begin{aligned} x_{n+1}^{(h)} - bx_n^{(h)} &= \left[x_{n+1} - x_{n+1}^{(p)}\right] - b\left[x_n - x_n^{(p)}\right] \\ &= \left[x_{n+1} - bx_n\right] - \left[x_{n+1}^{(p)} - bx_n^{(p)}\right] \\ &= g_n - g_n \\ &= 0 . \end{aligned}$$

Therefore, $\{x_n^{(h)}\}$ is a solution of the associated homogeneous equation. Hence,

$$(2.12) \qquad x_n = x_n^{(h)} + x_n^{(p)}$$

is the sum of the general solution of the associated homogeneous equation and a particular solution of the nonhomogeneous equation.

Example 2.3. Solve the difference equation

$$x_{n+1} + 3x_n = 4$$

with initial condition $x_0 = 5$.

We first find the general solution, saving the initial condition for the end. The associated homogeneous equation

$$x_{n+1} + 3x_n = 0$$

has general solution $x_n^{(h)} = C(-3)^n$. Also, we see by direct inspection that $x_n^{(p)} = 1$ is a particular solution of the nonhomogeneous equation because $1 + 3 = 4$. Therefore, by the above discussion, we conclude that the general solution of the nonhomogeneous equation is

$$\begin{aligned} x_n &= x_n^{(h)} + x_n^{(p)} \\ &= C(-3)^n + 1 . \end{aligned}$$

We now use the initial condition $x_0 = 2$ to find the appropriate value of the constant C. Since

$$5 = x_0 = C(-3)^0 + 1 = C + 1$$

we obtain $C = 4$. Hence, the desired solution is $x_n = (-3)^n + 4$.

Thus the procedure for solving nonhomogeneous linear equations is as follows:

1. Find the general solution $\{x_n^{(h)}\}$ of the associated homogeneous equation.
2. Find a particular solution $\{x_n^{(p)}\}$ of the nonhomogeneous equation.
3. Add $x_n = x_n^{(h)} + x_n^{(p)}$ to get the general solution $\{x_n\}$ of the nonhomogeneous equation.
4. If there are initial conditions, use them to find appropriate values of the constants in the general solution.

This procedure actually applies to any *linear* difference equation. It will be used later when we discuss second order linear difference equations.

The problem in solving nonhomogeneous equations thus hinges on finding particular solutions. There is no universal method for doing this, and if the given $\{g_n\}$ is complicated, it can be a difficult problem. However, for cases of an elementary character, particular solutions can be found with relative ease. We first consider the case $g_n = a$, a constant.

2.3 The Equation $x_{n+1} - bx_n = a$

The difference equation $x_{n+1} - bx_n = a$ where a is a given constant is easily solved. There are two cases to consider:

(i) $b = 1$. It is readily checked that the equation

$$x_{n+1} - x_n = a$$

has a particular solution $x_n^{(p)} = na$ since $(n+1)a - na = a$. The general solution of the homogeneous equation $x_{n+1} - x_n = 0$ is $x_n^{(h)} = C(1)^n = C$, a constant. Thus, the general solution is

$$x_n = C + na. \tag{2.13}$$

(ii) $b \neq 1$. We note that if $\{x_n\}$ is constant, then the difference equation becomes simply a linear equation. Thus, if $x_n^{(p)} = x$ we have

$$x - bx = a$$

and since $b \neq 1$,

$$x_n^{(p)} = x = \frac{a}{1-b}.$$

The general solution of the associated homogeneous equation

$$x_{n+1} - bx_n = 0$$

is

$$x_n^{(h)} = Cb^n.$$

Thus, the general solution of the nonhomogeneous equation when $b \neq 1$ is

$$x_n = Cb^n + \frac{a}{1-b}. \tag{2.14}$$

Example 2.4. The Tower of Hanoi is a puzzle consisting of a board with three pegs and n circular rings of decreasing size located on one of the pegs (see Figure 1). The problem is to transfer the rings to another peg by moving one ring at a time and never placing a ring on top of a smaller ring. The third peg can be used as a temporary resting place for rings during the transfer process. We ask the question: How many moves are required to accomplish the transfer, leaving the relative position of the rings unchanged?

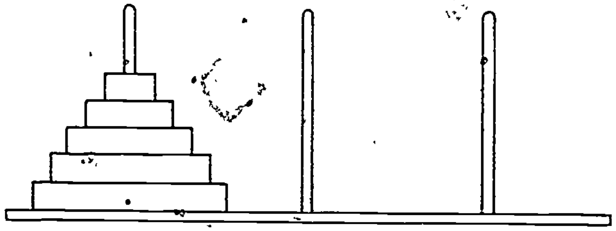


Figure 1. The Tower of Hanoi.

Let x_n denote the number of moves it takes to move n rings from one peg to another. Then x_{n+1} is related to x_n by the recurrence relation

$$x_{n+1} = 2x_n + 1$$

because we can move n rings to the second peg in x_n moves, then transfer ring n+1 to the third peg in one move, and finally move the n rings from the second peg to the third peg in another x_n moves. Hence, it takes $x_n + 1 + x_n = 2x_n + 1$ moves to transfer n+1 rings.

In this case b = 2 and a = 1 so that the general solution is

$$x_n = C2^n + \frac{1}{1-2}$$
$$= C2^n - 1.$$

The initial condition $x_1 = 1$ yields $C = 1$. Thus, the Tower of Hanoi with n rings can be solved in $x_n = 2^n - 1$ moves.

Example 2.5 (Annuities). In Example 2.1, an initial amount of money P was placed in an account that earned interest compounded at regular intervals. In an annuity, equal amounts are deposited in the account at each interest period so that the amount grows with additional deposits as well as accrued interest. This type of account is called an annuity and is usually handled by insurance companies for retirement funds, college expense funds, and so on.

The growth of money in an annuity can be described by a difference equation. Let x_n be the amount in the annuity after n interest periods, and suppose the same amount P is deposited at the beginning of each interest period. If the interest rate is r, then

$$x_{n+1} = (\text{previous amount}) + (\text{interest}) + (\text{deposit})$$

$$= x_n + rx_n + P$$

$$= (1+r)x_n + P.$$

Hence, we have the difference equation

$$(2.15) \qquad x_{n+1} - (1+r)x_n = P$$

with initial condition $x_0 = P$.

From above, we see that the general solution is

$$x_n = C(1+r)^n + \frac{P}{1-(1+r)}$$

$$= C(1+r)^n - \frac{P}{r}.$$

Since $x_0 = P$, we have $P = C - P/r$ so that

$$C = P + \frac{P}{r} = P\left(\frac{1+r}{r}\right).$$

Then we have

$$(2.16) \qquad x_n = P\left(\frac{1+r}{r}\right)(1+r)^n - \frac{P}{r}$$

$$= \frac{(1+r)^{n+1} - 1}{r}P.$$

It is common in annuities to set a goal and then determine what deposit is needed to achieve the goal. For example, suppose $r = 0.08$ compounded annually and we wish to have \$10,000 at the end of 20 years. What should our annual deposit P be? We want $x_{20} = 10{,}000$ so we must solve

$$\frac{(1.08)^{21} - 1}{0.08}P = 10{,}000$$

Using a calculator, we find that

$$P = \$192.32.$$

It is instructive to graph the solutions of difference equations to get a feeling for the long-range behavior of x_n as n gets large. We shall do this for the difference equation x_{n+1} [illegible] $bx_n = a$. We know from Equation (2.14) that the general solution when $b \neq 1$ is

$$x_n = Cb^n + \frac{a}{1-b}.$$

If $|b| < 1$, then b^n approaches zero as n gets large. Consequently, $x_n \to a/(1-b)$ as $n \to \infty$. This is illustrated in Figure 2 where lines connecting successive values of x_n have been drawn for visual emphasis. We assume that $x_0 > a/(1-b)$ and $b < 0$ in the graph. If $b > 0$, then the values of x_n simply decrease steadily instead of oscillating as shown.

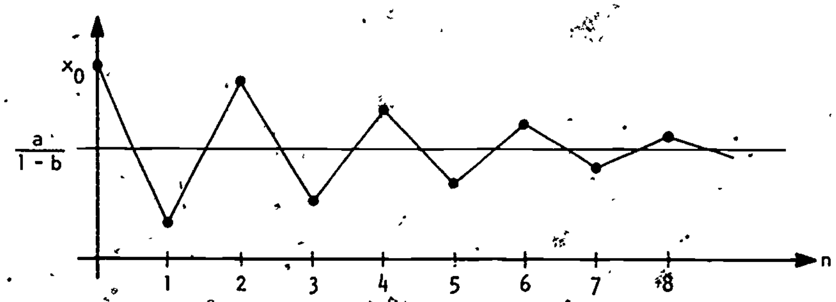


Figure 2. $|b| < 1$ and $b < 0$.

If b = -1, then x_n simply oscillates between x_0 and $-x_0 + a$ as shown in Figure 3.

If b > 1, then $b^n \to \infty$ as $n \to \infty$ and the graph would steadily rise. If b < -1, then b^n grows large in absolute value, but alternates in sign. This case is illustrated in Figure 4.

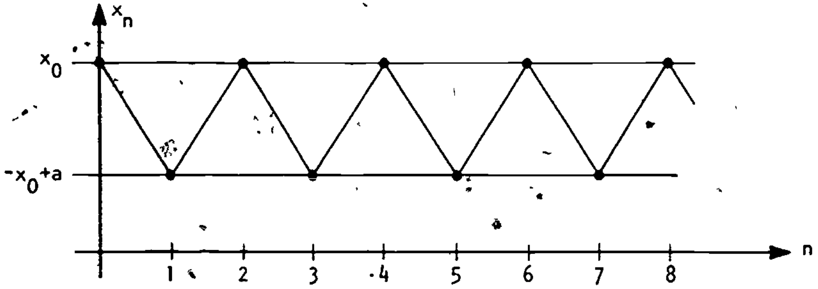


Figure 3. b = -1.

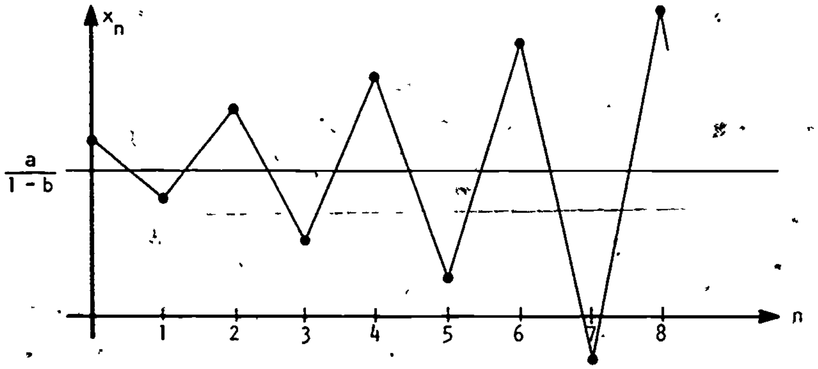


Figure 4. b < -1.

Example 2.6 (The Cobweb Theorem of Economics). In the marketplace, the supply and demand of a product are closely related to the price. A reasonable relation in many cases is illustrated by the curves as shown in Figure 5. A rise in prices lowers consumer demand, but increases supply since producers wish to take advantage of the higher price. However, a time lag occurs as prices and supply adjust to changes. Farm products such as grain or hogs are good examples of lagged supply adjustments. A fall in price one year causes a farmer to cut back production the next year, and the decreased supply then causes a rise in price during the next year, and so on. Prices thus rise and fall cyclically.

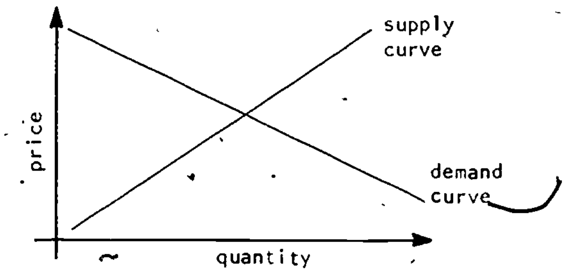


Figure 5.

An elementary model using difference equations can be used to analyze the market stability of these lagged adjustments. Let p_n and s_n denote the price and supply, respectively, of a product in the n^{th} year. We assume

(2.17) $$p_n = a - bs_n,$$

where a > 0, b > 0, since a large supply causes a low price in a given year. Similarly, we assume that price and supply in alternate years are proportional, so that the lagged adjustment is given by

(2.18) $$p_n = ks_{n+1}$$

where k is a positive constant of proportionality. If we concentrate on price, we can combine these two relationships to get a difference equation for the price:

$$\begin{aligned} p_{n+1} &= a - bs_{n+1} \\ &= a - b\left(\frac{p_n}{k}\right). \end{aligned}$$

Thus, we obtain

(2.19) $$p_{n+1} + \frac{b}{k}p_n = a.$$

As shown above, but with coefficient -b/k instead of b, the solution of this difference equation is given by

$$(2.20) \quad p_n = C\left(-\frac{b}{k}\right)^n + \frac{a}{1 + (b/k)}$$

$$= C\left(-\frac{b}{k}\right)^n + \frac{ak}{k + b} .$$

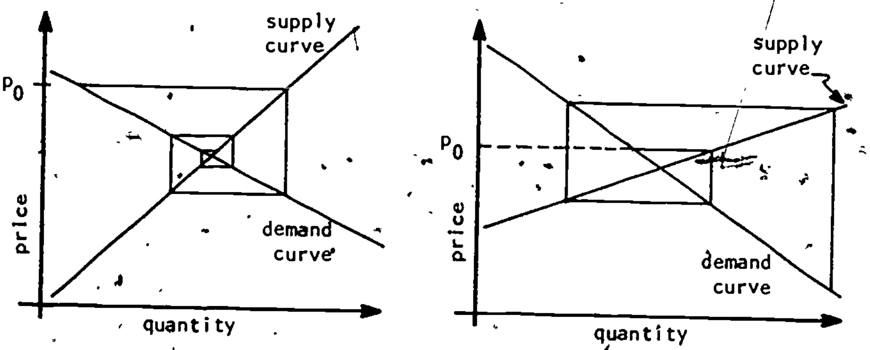
The long-range behavior of the price thus depends on the size of b/k.

Case 1. If $b/k < 1$, then $p_n \to \frac{ak}{k+b}$ as $n \to \infty$ and the market price tends to stabilize. Price variation is shown by the graph in Figure 2 on page 14.

Case 2. If $b/k = 1$, then p_n oscillates between p_0 and $-p_0 + a$, so the market is unstable. See Figure 3 on page 15.

Case 3. If $b/k > 1$, then the oscillations of p_n become larger and larger. The market is unstable, but the model fails when the price becomes negative. See the graph in Figure 4 on page 15.

The lagged adjustments can be displayed dramatically by plotting the changes on the supply and demand curves. We illustrate cases 1 and 3 in Figure 6.



Case 1 (Stable). $\frac{b}{k} < 1$. ~~Case 3~~ (Unstable). $\frac{b}{k} > 1$.

Figure 6.

The suggestive appearance of these pictures is the reason that the above analysis is referred to as the Cobweb Theorem in economics. Note that b and k are the slopes of the demand and supply curves. Thus, if supply adjusts more radically than demand to price changes, then the market will tend to stabilize (case 1). The reverse situation leads to instability (case 3). In case 2, the cobweb is reduced to a rectangle that is retraced over the years.

Exercises

1. For each of the following, find the general solution and then the solution satisfying the stated initial condition.
   a. $x_{n+1} - 5x_n = 3, \quad x_0 = 5$
   b. $x_{n+1} + x_n = 7, \quad x_0 = 1$
   c. $2x_{n+1} - x_n = 4, \quad x_0 = 1.$
2. Starting at the day of birth, parents deposit $500 per year at 6 percent interest compounded annually. How much is in the account when the child turns 19 years old?
3. In 1626, Peter Minuit of the New Netherlands province purchased Manhattan Island for goods worth $24. If this amount had been invested at 7 percent compounded quarterly, find the value of the investment in 1976. (The amount, approximately 850 billion dollars, is more than Manhattan is worth today.)
4. (Amortization of loans.) Suppose an amount L is borrowed and is to be paid in equal installments such that each payment is to include interest on the unpaid balance. That is,

   new balance = old balance + interest - payment.

   Suppose the annual interest rate r is compounded monthly (i.e., r/12 per month) and the monthly payment is P. Find a formula for B_n, the balance after n payments.

5. Apply the result of Exercise 4 to a house mortgage of $30,000 at interest rate $r = 0.10$ to be paid in 20 years. What must the monthly payment P be? How much interest is paid over the 20 years?

6. Suppose that the interest rate on a mortgage is 10 percent compounded monthly. If you can afford to pay $300 per month for 30 years, how much money can you borrow? (Estimate L so that $B_{360} \geq 0$.)

7. (Pizza slicing.) Show that n distinct straight lines in a plane that all pass through a common point divide the plane into 2n regions.

8. Suppose that the current price of oats is $1.25 per bushel and that the price p_n of oats n years from now satisfies $p_{n+1} + 0.6p_n = 1.6$. Sketch the graph of p_n.

2.4 The Method of Undetermined Coefficients

The "method of undetermined coefficients" is a method of finding a particular solution of a nonhomogeneous difference equation

$$(2.21) \qquad x_{n+1} - bx_n = g_n$$

by imitating the form of g_n. If g_n has a simple form, the method is often successful. We shall illustrate the method with a few examples.

Example 2.7. (1) Find a particular solution of the equation

$$(2.22) \qquad x_{n+1} + 2x_n = 3n + 4.$$

Because $g_n = 3n + 4$, we attempt a solution $x_n^{(p)} = An + B$ where the coefficients A and B are to be determined. This is done by putting it into Equation (2.22) as follows:

$$[A(n+1) + B] + 2[An + B] = 3n + 4.$$

Collecting terms, we get

$$3An + A + 3B = 3n + 4.$$

This should be valid for each value of n, so that

$$\begin{cases} n = 0: & A + 3B = 4 \\ n = 1: & 4A + 3B = 7. \end{cases}$$

Solving for A and B, we obtain A = 1 and B = 1. Hence, $x_n^{(p)} = n+1$ is a particular solution of (2.22).

(2) To find a particular solution of the difference equation

$$x_{n+1} - 3x_n = 2^n,$$

we try $x_n^{(p)} = A2^n$. Then

$$A2^{n+1} - 3(2^n) = 2^n$$

so that A = 2. Hence,

$$x_n^{(p)} = 2(2^n) = 2^{n+1}.$$

Difficulties can arise if the given $\{g_n\}$ has the same form as the solution of the associated homogeneous equation. However, multiplication by n can often resolve the problem as the following example illustrates.

Example 2.8. (1) A particular solution of the difference equation

$$x_{n+1} - 2x_n = 2^n$$

cannot have the form $A2^n$ since this is a solution of the associated homogeneous equation. We try $x_n^{(p)} = An2^n$. Substitution yields

$$A(n+1)2^{n+1} - 2An2^n = 2^n$$

so that A = 1. Thus, $x_n^{(p)} = n2^n$ is a particular solution.

(2) For the equation $x_{n+1} - x_n = n$ the attempt $x_n^{(p)} = An + B$ fails. Multiplying by n, we try $x_n^{(p)} = An^2 + Bn$. This leads to the condition

$$2An + A + B = n, \text{ so that } A = 1/2,\ B = -1/2.$$

Exercises

1. Find the general solution and then the solution that satisfies the given initial condition.

   a. $x_{n+1} - 5x_n = 3n, \quad x_0 = 1$

   b. $x_{n+1} + x_n = 4n^2, \quad x_0 = 1$

   c. $2x_{n+1} - x_n = 2^n, \quad x_0 = 2$.

   d. $x_{n+1} - x_n = n2^n, \quad x_0 = 1$.

2. Suppose n straight lines are drawn in a plane so that no two lines are parallel and no more than two lines intersect at any point. Let x_n be the number of different regions determined by the lines. Find and solve a difference equation to derive a formula for x_n. Note that $x_0 = 1$.

3. An empty lake is stocked with fish by putting in 100 fish the first year, 200 fish the second year, and so on. Through reproduction the number of fish increases by 50 percent each year. How many fish will there be after n years?

3. SECOND ORDER LINEAR DIFFERENCE EQUATIONS

We now consider equations of the form

(3.1) $\quad x_{n+2} + ax_{n+1} + bx_n = g_n.$

The associated homogeneous equation is

(3.2) $\quad x_{n+2} + ax_{n+1} + bx_n = 0.$

The relationship between solutions of the Equation (3.1) and the associated Equation (3.2) is the same as for first order equations.

That is, if $\{x_n^{(h)}\}$ is the general solution of the homogeneous Equation (3.2) and if $\{x_n^{(p)}\}$ is a particular solution of the Equation (3.1), then the general solution of (3.1) is given by

$$x_n = x_n^{(h)} + x_n^{(p)}.$$

Thus, adding a particular solution of the nonhomogeneous equation to the general solution of the associated homogeneous equation yields the general solution of the nonhomogeneous equation. The verification is the same as in the first order case.

3.1 Homogeneous Equations

Before developing techniques to solve homogeneous equations of the form (3.2), we make some preliminary observations. First, if $\{u_n\}$ and $\{v_n\}$ are each solutions of Equation (3.2) and if C_1 and C_2 are constants, then

(3.3) $\quad x_n = C_1u_n + C_2v_n$

is again a solution of Equation (3.2). This is directly verified by substituting (3.3) into Equation (3.2) as follows:

$$\begin{aligned}(C_1u_{n+2}+C_2v_{n+2}) + a(C_1u_{n+1}+C_2v_{n+1}) + b(C_1u_n+C_2v_n)\\ = C_1(u_{n+2}+au_{n+1}+bu_n) + C_2(v_{n+2}+av_{n+1}+bv_n)\\ = C_1 \cdot 0 + C_2 \cdot 0\\ = 0.\end{aligned}$$

The zeros result from the hypothesis that $\{u_n\}$ and $\{v_n\}$ each satisfy Equation (3.2).

Thus, if two different solutions can be found, they can be combined as in (3.3) above to yield other solutions. We next observe that two different solutions are actually needed to generate the general solution $\{x_n\}$. It is possible to express any x_n in terms of x_0 and x_1 by successive calculations. For n = 0, Equation (3.2) gives

$$x_2 = -ax_1 - bx_0,$$

and n = 1 yields

$$\begin{aligned} x_3 &= -ax_2 - bx_1 \\ &= -a(-ax_1 - bx_0) - bx_1 \\ &= (a^2 - b)x_1 + abx_0. \end{aligned}$$

For n = 2 we get after simplification that

$$x_4 = (2ab - a^3)x_1 + (b^2 - a^2 b)x_0.$$

Proceeding in this manner, we see that the value of x_n is uniquely determined by specifying values of the initial terms x_0 and x_1.

This iteration process does not lead to a useful formula in general, and without a computer it is tedious and not practical. However, if we set $C_1 = x_0$ and $C_2 = x_1$, then this can be used to show that the general form of solution of the homogeneous equation (3.2) is

$$x_n = C_1 u_n + c_2 v_n$$

where u_n and v_n are determined by the coefficients a and b as indicated above. It can also be shown that $\{u_n\}$ and $\{v_n\}$ are themselves solutions of the homogeneous equation.

In summary, the general solution of the homogeneous equation

$$x_{n+1} + ax_{n+1} + bx_n = 0 \tag{3.4}$$

is of the form

$$x_n = C_1 u_n + C_2 v_n, \tag{3.5}$$

where C_1, C_2 are arbitary constants and $\{u_n\}$, $\{v_n\}$ are distinct solutions of the equation.

Exercises

1. Verify by direct substitution that $u_n = 2^n$ and $v_n = n2^n$ are both solutions of

$$x_{n+2} - 4x_{n+1} + 4x_n = 0$$

so that the general solution is

$$x_n = C_1 2^n + C_2 n2^n.$$

2. Find the solution of Exercise 1 that is determined by the initial conditions $x_0 = 1$ and $x_1 = 2$

3. Show how the conditions $x_2 = 16$ and $x_5 = 32$ for Exercise 1 lead to a system of equations in C_1 and C_2. Solve for C_1 and C_2, and thereby find the solution satisfying these conditions.

3.2 The Auxiliary Equation

We seek solutions of the homogeneous difference equation

$$x_{n+2} + ax_{n+1} + bx_n = 0. \tag{3.6}$$

Since first order equations have solutions of the form $x_n = \lambda^n$, a geometric progression, let us see if similar solutions exist for second order equations where λ will be a constant involving the coefficients a and b. Substituting $x_n = \lambda^n$ into the difference equation, we get

$$\lambda^{n+2} + a\lambda^{n+1} + b\lambda^n = 0. \tag{3.7}$$

Rejecting the trivial zero solution, we assume $\lambda \neq 0$. Then cancelling λ^n in the preceding equation, we get

$$\lambda^2 + a\lambda + b = 0.$$

This quadratic in λ is called the *auxiliary equation* of the difference equation (3.6).

Applying the quadratic formula, we obtain the roots

$$\lambda_1 = \frac{-a + \sqrt{a^2 - 4b}}{2}, \quad \lambda_2 = \frac{-a - \sqrt{a^2 - 4b}}{2}. \tag{3.8}$$

The general solution of the difference equation depends on the nature of these roots. Three cases arise:

(i) λ_1 and λ_2 are real and unequal $(a^2 - 4b > 0)$

(ii) λ_1 and λ_2 are real and equal $(a^2 - 4b = 0)$

(iii) λ_1 and λ_2 are complex conjugates $(a^2 - 4b < 0)$.

Case (i). If λ_1 and λ_2 are real and $\lambda_1 \neq \lambda_2$, then the general solution of the homogeneous difference equation (3.6) is

$$x_n = C_1\lambda_1^n + C_2\lambda_2^n. \tag{3.9}$$

Since λ_1 and λ_2 satisfy the auxiliary equation, we see that $u_n = \lambda_1^n$ and $v_n = \lambda_2^n$ are two distinct solutions of the difference equation. Thus, according to the discussion in the preceding section, the general solution is $x_n = C_1u_n + C_2v_n$. This gives us the stated solution.

Example 3.1. We are now in a position to derive a formula for the Fibonacci numbers. The Fibonacci numbers $\{x_n\}$ satisfy the difference equation

$$x_{n+2} - x_{n+1} - x_n = 0$$

with initial conditions $x_0 = 1$, $x_1 = 1$. The auxiliary equation $\lambda^2 - \lambda - 1 = 0$ has solutions

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}, \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Hence, the difference equation has general solution

$$x_n = C_1\left(\frac{1 + \sqrt{5}}{2}\right)^n + C_2\left(\frac{1 - \sqrt{5}}{2}\right)^n.$$

We now apply the initial conditions to find the values of C_1 and C_2. The conditions $x_0 = 1$ and $x_1 = 1$ lead to a system of two equations:

$$\begin{cases} C_1 + C_2 = x_0 = 1 \\ \left(\frac{1 + \sqrt{5}}{2}\right)C_1 + \left(\frac{1 - \sqrt{5}}{2}\right)C_2 = x_1 = 1 \end{cases} \tag{3.10}$$

Solving this system, we get

$$C_1 = \frac{1 + \sqrt{5}}{2\sqrt{5}}, \quad C_2 = -\frac{1 - \sqrt{5}}{2\sqrt{5}}.$$

Hence, the formula for the n^{th} Fibonacci number is

$$x_n = \frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1 - \sqrt{5}}{2}\right)^{n+1}.$$

This may be a somewhat surprising formula in view of the fact that x_n is a positive integer for every n.

Case (ii). If $\lambda_1 = \lambda_2 = -\frac{a}{2}$, then the general solution of the difference equation (3.2) is

$$x_n = (C_1 + nC_2)\left(-\frac{a}{2}\right)^n. \tag{3.11}$$

In this case, the auxiliary equation is $(\lambda + a/2)^2 = 0$, which has only one root. We obtain one solution,

$$u_n = \left(-\frac{a}{2}\right)^n, \tag{3.12}$$

to the difference equation, but we need a second solution in order to obtain the general solution. This is done by multiplying our one solution by n; that is, we let

$$v_n = nu_n. \tag{3.13}$$

Direct substitution verifies this is in fact a solution. Thus, the general solution is

$$x_n = C_1u_n + C_2nu_n = (C_1 + nC_2)u_n,$$

as stated in Equation (3.11).

Example 3.2. Suppose gamblers A and B play a game of matching pennies where A,B start the game with N_A, N_B pennies, respectively. The game ends when one player has lost all of his pennies. We assume the coins are fair so that each player has probability 1/2 of winning on each play. What is the probability that A will win all the pennies?

Let P_n be the probability that if A has n pennies, then A will win the game. Let $N = N_A + N_B$. Clearly,

$P_0 = 0$ and $P_N = 1$. Consider a value of n such that A has n+1 pennies and $0 < n+1 < N$. After one play, A will have either n+2 pennies or n pennies, depending on whether A wins or loses on that play. Therefore,

$$P_{n+1} = \tfrac{1}{2}P_{n+2} + \tfrac{1}{2}P_n.$$

Hence, we have the difference equation

$$P_{n+2} - 2P_{n+1} + P_n = 0.$$

The auxiliary equation

$$\lambda^2 - 2\lambda + 1 = 0$$

has roots $\lambda_1 = \lambda_2 = 1$, so that the general solution is

$$P_n = C_1 + C_2 n.$$

Since $P_0 = 0$, we get $C_1 = 0$, so that $P_n = C_2 n$. Also, $1 = P_N = C_2 N$, so that $C_2 = 1/N$. Hence, the desired probabilities are

$$P_n = \frac{n}{N_A + N_B}, \quad 0 < n < N_A + N_B.$$

The probability that A will win starting with N_A pennies is therefore

$$P_{N_A} = \frac{N_A}{N_A + N_B}.$$

We conclude that it is unwise to play an even game against an opponent with greater resources. For example, if A starts with 10 pennies and B starts with 90 pennies, then the probability that A will win is only 1/10.

Case (iii). If λ_1 and λ_2 are complex, then the general solution of the homogeneous difference equation (3.6) is

$$(3.14) \qquad x_n = r^n(C_1 \cos n\theta + C_2 \sin n\theta)$$

where r and θ are given by

$$r = \sqrt{b} \quad \text{and} \quad \tan\theta = -\frac{\sqrt{4b - a^2}}{a}.$$

The derivation of this solution requires some familiarity with complex numbers. If $a^2 - 4b < 0$, then the auxiliary equation $\lambda^2 + a\lambda + b = 0$ has complex roots

$$\lambda_1 = -\frac{a}{2} + i\frac{\sqrt{4b - a^2}}{2}, \quad \lambda_2 = -\frac{a}{2} - i\frac{\sqrt{4b - a^2}}{2}.$$

These can be written in polar form

$$\lambda_1 = r(\cos\theta + i\sin\theta)$$

$$\lambda_2 = r(\cos\theta - i\sin\theta)$$

where r is the modulus $r = |\lambda_1| = |\lambda_2| = \sqrt{b}$ and θ is the argument of λ_1 as given above. Then DeMoivre's Theorem gives us

$$\lambda_1^n = r^n(\cos n\theta + i\sin n\theta)$$

$$\lambda_2^n = r^n(\cos n\theta - i\sin n\theta).$$

These are complex solutions to the difference equation, but we wish to have real solutions. This is done by taking real and imaginary parts. Setting

$$u_n = \frac{1}{2}\left(\lambda_1^n + \lambda_2^n\right) = r^n \cos n\theta$$

$$v_n = \frac{1}{2i}\left(\lambda_1^n - \lambda_2^n\right) = r^n \sin n\theta,$$

we see that u_n and v_n are solutions because λ_1^n and λ_2^n are solutions. The general real solution is thus given by $x_n = C_1 u_n + C_2 v_n$ as stated in Equation (3.14).

Example 3.3. The difference equation

$$x_{n+2} - 2x_{n+1} + 2x_n = 0$$

has auxiliary equation

$$\lambda^2 - 2\lambda + 2 = 0$$

with roots

$$\lambda_1 = 1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$$

$$\lambda_2 = 1 - i = \sqrt{2}\left(\cos\frac{\pi}{4} - i\sin\frac{\pi}{4}\right).$$

Therefore, the general solution of the difference equation is given by

$$x_n = 2^{n/2}\left(C_1 \cos\frac{n\pi}{4} + C_2 \sin\frac{n\pi}{4}\right)$$

If we also have an initial condition, say $x_0 = 5$, $x_1 = 2$, then $C_1 = 5$ and $C_1 + C_2 = 2$ so that the solution is

$$x_n = 2^{n/2}\left(5\cos\frac{n\pi}{4} - 3\sin\frac{n\pi}{4}\right).$$

3.3 Nonhomogeneous Equations

We now turn to the nonhomogeneous equation

$$(3.15) \qquad x_{n+2} + ax_{n+1} + bx_n = f_n.$$

As with first order linear equations, if we add any particular solution of (3.15) to the general solution of the associated homogeneous equation, the sum will be the general solution of the nonhomogeneous equation. Thus, we wish to determine a particular solution of (3.15). The form of a particular solution can often be inferred from the nature of the given f_n and this leads to the method of undetermined coefficients. It is essentially the same approach as used in the first order case. We illustrate the method with the following example.

Example 3.4. Consider the difference equation

$$(3.16) \qquad x_{n+2} - 3x_{n+1} + 2x_n = k^n$$

where k is some constant. The auxiliary equation

$$\lambda^2 - 3\lambda + 2 = 0$$

has roots $\lambda_1 = 1$ and $\lambda_2 = 2$, so that the corresponding homogeneous equation has general solution $\{C_1 + C_2 2^n\}$.

As an attempt to find a particular solution, let us try $x_n^{(p)} = Ak^n$ where the coefficient A is to be determined. Substitution into the difference equation yields

$$Ak^n(k-1)(k-2) = k^n$$

for all n, so that A must satisfy

$$A(k-1)(k-2) = 1.$$

If $k \neq 1$, $k \neq 2$, then we see that

$$A = \frac{1}{(k-1)(k-2)}.$$

For example, if k = 3, then $A = \frac{1}{2}$, so that $x_n^{(p)} = \frac{1}{3}3^n$ is a particular solution and

$$x_n = C_1 + C_2 2^n + \frac{1}{2}3^n$$

is the general solution of the nonhomogeneous equation.

However, if k = 1 or k = 2, that is, if k equals either of the roots of the auxiliary equation, then no value of A satisfies the required condition. This is not at all surprising because these values of k yield solutions of the homogeneous equation and therefore could not very well produce solutions of the nonhomogeneous equation. Consequently, the form of particular solutions must be modified for k = 1 and k = 2. The appropriate modification is to multiply the related solution of the homogeneous equation by n. Let us look at the two cases separately.

If k = 1, then the difference equation (3.16) becomes

$$(3.17) \qquad x_{n+2} - 3x_{n+1} + 2x_n = 1.$$

Since $\lambda_1 = 1$ is a root of the auxiliary equation, we know the constant sequence $x_n = 1$ is a solution of the homogeneous equation and therefore $x_n^{(p)} = A$ is *not* a particular solution of the nonhomogeneous equation. But if we multiply by n and try $x_n^{(p)} = An$ we shall succeed. Substituting this into the difference equation, we get

$$A(n+2) - 3A(n+1) + 2An = 1$$

so that A = -1. Hence, $x_n^{(p)} = -n$ gives us a particular solution.

If k = 2, the difference equation (3.16) becomes

(3.18) $$x_{n+2} - 3x_{n+1} + 2x_n = 2^n$$

Now $x_n = 2^n$ is a solution of the homogeneous equation, so we try $x_n^{(p)} = An2^n$ as the form of a particular solution of the nonhomogeneous equation (3.18). Then we get

$$A(n+2)2^{n+2} - 3A(n+1)2^{n+1} + 2An2^n = 2^n$$

so that

$$4A(n+2) - 6A(n+1) + 2An = 1$$

and hence A = 1/2. Thus, $x_n^{(p)} = n2^{n-1}$ gives us a particular solution of (3.18).

Exercises

1. Find the general solution and then the solution that satisfies $x_0 = 1$, $x_1 = 1$.
   a. $x_{n+2} - x_n = 0$
   b. $x_{n+2} + x_{n+1} - 6x_n = 0$
   c. $x_{n+2} + 2x_{n+1} + x_n = 0$
   d. $x_{n+2} + x_n = 0$.
2. Find the general solution and then the solution that satisfies $x_0 = 1$, $x_1 = 1$.
   a. $x_{n+2} - 5x_{n+1} + 6x_n = 2$
   b. $x_{n+2} - 4x_{n+1} + 3x_n = 2$

   c. $x_{n+2} - 4x_{n+1} + 4x_n = 3n + 2^n$
   d. $x_{n+2} + x_n = \sin\left(\frac{n\pi}{2}\right)$.
3. Consider a telegraphy system in which the symbols that can be transmitted are dots, of 1-second duration, and dashes, of 2-seconds duration. Let x_n represent the number of distinct messages of duration n seconds. Find a difference equation for x_n and so determine a formula for x_n. Note that $x_1 = 1$ (one dot) and $x_2 = 2$ (two dots or one dash).
4. Suppose the *increase* of a fish population each year is twice the increase of the previous year. If initially there are 1,000 fish, and if there are 1,100 the following year, find the population in the nth year.
5. In the preceding problem, suppose 100 fish are removed each year. Find the population in the nth year.

---

4. ANSWERS TO EXERCISES

Section 1, page 3

1a. $x_{n+1} - x_n = 3$, $x_0 = 3$.

b. $x_{n+1} = 3x_n$, $x_0 = 3$.

c. $x_{n+1} - x_n = 2^n$.

2. 3, 7, 15, 31, 63, 127, 255, 511, 1023, 2047.

3. 576.

4. 7, -4, 3, -1, 2, 1, 3, 4, 7, 11.

Section 2.1, pp. 7-8

1a. $x_n = C5^n$.

b. $x_n = C(2/3)^n$.

c. $x_n = C$.

2a. $x_n = 5^n$.

b. $x_n = 3(-3/5)^{n-5}$.

3. Solve $(1.02)^n = 2$ for $n = 35$ quarters; 8.75 years.

4. $x_{n+1} = (1 + 0.06 - 0.1)x_n = 0.96x_n$, $x_0 = 10^6$.

5. 405.

6. $r_n = r_0(0.99)^n$, $r_4 \sim 0.96r_0$, 1724 years.

Section 2.3, pp. 18-19

1a. $x_n = \frac{1}{4}(23 \cdot 5^n - 3)$.

b. $x_n = \frac{1}{2}\left(5(-1)^{n+1} + 7\right)$.

c. $x_n = -3(1/2)^n + 4$.

2. 18,393.

3. 848,000,000,000.

4. $B_n = (L - \frac{12p}{r})(1 + \frac{r}{12})^n + \frac{12p}{r}$.

5. $P = 290$, Interest $= 240 \cdot 290 - 30{,}000 = 39{,}600$.

6. 37,815.

7. $x_{n+1} - x_n = 2$, $x_1 = 2$ has solution $x_n = 2n$.

Section 2.4, page 21

1a. $x_n = C5^n - \frac{3}{4}n - \frac{3}{16}$. If $x_0 = 1$, then $C = \frac{19}{16}$.

b. $x_n = C(-1)^n + 2n^2 - 2n$. If $x_0 = 1$, then $C = 1$.

c. $x_n = C(1/2)^n + 2^{n-1}$. If $x_0 = 2$, then $C = 3/2$.

d. $x_n = C + (n-2)2^n$. If $x_0 = 1$, then $C = 3$.

2. $x_{n+1} = x_n + n + 1$, $x_0 = 1$, has solution $x_n = \frac{1}{2}(n^2 + n + 2)$.

3. $x_{n+1} - \frac{3}{2}x_n = 100(n+1)$, $x_0 = 0$, has solution

$x_n = 600(3/2)^n - 200n - 600$.

Section 3.1, page 23-24

2. $x_n = 2^n - n2^{n-2}$.

3. $\begin{cases} C_1 + 2C_2 = 4 \\ C_1 + 5C_2 = 1 \end{cases}$

$C_1 = 6$, $C_2 = -1$.

Section 3.3, pp. 31-32

1a. $x_n = C_1 + C_2(-1)^n$, $x_n = 1$.

b. $x_n = C_1 2^n + C_2(-3)^n$, $x_n = \frac{1}{5}(2^{n+2} + (-3)^n)$.

c. $x_n = (C_1 + C_2 n)(-1)^n$, $x_n = (1-2n)(-1)^n$.

d. $x_n = C_1 \cos\left(\frac{n\pi}{2}\right) + C_2 \sin\left(\frac{n\pi}{2}\right)$, $x_n = \cos\left(\frac{n\pi}{2}\right) + \sin\left(\frac{n\pi}{2}\right)$.

2a. $x_n = C_1 3^n + C_2 2^n + 1$; $C_1 = -2$, $C_2 = 2$.

b. $x_n = C_1 + C_2 3^n - n$; $C_1 = 3/2$, $C_2 = -1/2$.

c. $x_n = (C_1 + nC_2)2^n + n^2 2^{n-3} + 3n + 6$; $C_1 = -5$, $C_2 = -1/8$.

d. $x_n = C_1 \cos\left(\frac{n\pi}{2}\right) + C_2 \sin\left(\frac{n\pi}{2}\right)$; $C_1 = 1$; $C_2 = -1/2$.

3. $x_{n+2} = x_{n+1} + x_n$, the Fibonacci numbers.

4. $x_n = 900 + 100 \cdot 2^n$.

5. 1000 + 100n.

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 1

Request for Help

Student: If you have trouble with a specific part of this unit, please fill out this form and take it to your instructor for assistance. The information you give will help the author to revise the unit.

Your Name______________________________ Unit No.________

Page______ ○ Upper ○ Middle ○ Lower	OR	Section______ Paragraph______	OR	Model Exam Problem No._____ Text Problem No._____

Description of Difficulty: (Please be specific)

Instructor: Please indicate your resolution of the difficulty in this box.

○ Corrected errors in materials. List corrections here:

○ Gave student better explanation, example, or procedure than in unit. Give brief outline of your addition here:

○ Assisted student in acquiring general learning and problem-solving skills (not using examples from this unit.)

Instructor's Signature______________________

Please use reverse if necessary.

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 2

Unit Questionnaire

Name________________________________ Unit No.____________ Date____________

Institution____________________________ Course No.______________________

Check the choice for each question that comes closest to your personal opinion.

1. How useful was the amount of detail in the unit?

____Not enough detail to understand the unit
____Unit would have been clearer with more detail
____Appropriate amount of detail
____Unit was occasionally too detailed, but this was not distracting
____Too much detail; I was often distracted

2. How helpful were the problem answers?

____Sample solutions were too brief; I could not do the intermediate steps
____Sufficient information was given to solve the problems
____Sample solutions were too detailed; I didn't need them

3. Except for fulfilling the prerequisites, how much did you use other sources (for example, instructor, friends, or other books) in order to understand the unit?

____A Lot ____Somewhat ____A Little ____Not at all

4. How long was this unit in comparison to the amount of time you generally spend on a lesson (lecture and homework assignment) in a typical math or science course?

____Much Longer ____Somewhat Longer ____About the Same ____Somewhat Shorter ____Much Shorter

5. Were any of the following parts of the unit confusing or distracting? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Paragraph headings
____Examples
____Special Assistance Supplement (if present)
____Other, please explain____

6. Were any of the following parts of the unit particularly helpful? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Examples
____Problems
____Paragraph headings
____Table of Contents
____Special Assistance Supplement (if present)
____Other, please explain____

Please describe anything in the unit that you did not particularly like.

Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)

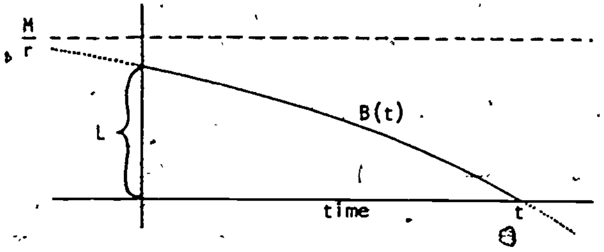
umap

UNIT 381

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT

SELECTED APPLICATIONS OF MATHEMATICS TO FINANCE AND INVESTMENT

by Frederick W. Luttmann



APPLICATIONS OF ELEMENTARY ALGEBRA
TO FINANCE

edc/umap 55chapel st / newton, mass 02160

SELECTED APPLICATIONS OF MATHEMATICS IN FINANCE AND INVESTMENT

by

Frederick W. Luttmann
Department of Mathematics
Sonoma State University
Rohnert Park, CA 94928

TABLE OF CONTENTS

Intermodular Description Sheet: *UMAP Unit 381*

Title: SELECTED APPLICATIONS OF MATHEMATICS TO FINANCE AND INVESTMENT

Author: Dr. Frederick W. Luttman
Department of Mathematics
Sonoma State University
Rohnert Park, CA 94928

Review Stage/Date: III 3/22/79

Classification: APPL ELEM ALG/FINANCE

Prerequisite Skills:
Elementary algebra, including exponents, logarithms, factoring, and geometric progressions. A few sections are specifically directed at students familiar with differential and integral calculus, but these may be skipped if desired, or read with minimal "faith" required at certain steps! (There is a brief and inevitable reference to limits in the definition of e, but the unit is comprehensible to pre-calculus students. Newton's Method is referred to briefly in the section of estimating APR's; section 6 on continuous approximation requires differential calculus; and the latter part of section 7 on present value of future payments requires integral calculus; all can be easily skipped by the student with insufficient background. At one stage in the last section, on the optimal time to sell, differential calculus is used to find an extremum---this step can be handwaved for students without preparation.)

Output Skills:
1. Understand business and financial concepts and formulas.
2. Be able to apply these concepts and formulas in practical circumstances.
3. Appreciate the power and usefulness of mathematical concepts and techniques.

Other Related Units:

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT (UMAP)

The goal of UMAP is to develop, through a community of users and developers, a system of instructional modules in undergraduate mathematics and its applications which may be used to supplement existing courses and from which complete courses may eventually be built.

The Project is guided by a National Steering Committee of mathematicians, scientists and educators. UMAP is funded by a grant from the National Science Foundation to Education Development Center, Inc., a publicly supported, nonprofit corporation engaged in educational research in the U.S. and abroad.

PROJECT STAFF

Ross L. Finney	Director
Solomon Garfunkel	Associate Director/Consortium Coordinator
Felicia DeMay	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Dianne Lally	Project Secretary
Paula M. Santillo	Administrative Assistant
Carol Forray	Production Assistant
Zachary Zevitas	Order Processor

NATIONAL STEERING COMMITTEE

W.T. Martin	MIT (Chairman)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
Ernest J. Henley	University of Houston
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
R. Duncan Luce	Harvard University
George Miller	Nassau Community College
Frederick Mosteller	Harvard University
Walter E. Sears	University of Michigan Press
George Springer	Indiana University
Arnold A. Strassenburg	SUNY at Stony Brook
Alfred B. Willcox	Mathematical Association of America

The Project would like to thank the members of the UMAP Subject Matter Panel in Analysis and Computation, Carroll O. Wilde, Chairman, for their reviews, and all others who assisted in the production of this unit.

This material was prepared with the support of National Science Foundation Grant No. SED76-19615 A02. Recommendations expressed are those of the author and do not necessarily reflect the views of the NSF, nor of the National Steering Committee.

SELECTED APPLICATIONS OF MATHEMATICS IN FINANCE AND INVESTMENT

1. REVIEW OF INTEREST CALCULATIONS; THE DEFINITION OF "e"; CONTINUOUS GROWTH

It is impossible to do any interesting or significant problems in the theory of finance without understanding continuous compounding of interest and, more generally, continuous increase in value with constant percentage growth rate.

First we review some useful terminology. We should make clear the difference between *interest* and *interest rate*. Interest is an actual sum of money, paid by a borrower to a lender for the privilege of having held a loan. Generally the longer the borrower has the loan, the higher an amount of interest he must pay. The interest *rate* is an expression of an amount of money *per time period*, usually per year, and it specifies how the *interest* itself is to be calculated at the end of any particular period of time. For example, a lender may say that on a loan of $3,000 the interest rate is $25 per month, which could also be expressed as $300 per year. This number in effect determines a rule for calculating the interest at any time, e.g., after three months the interest is $75, after two years the interest is $600, etc.

Interest rates are normally expressed as percentages of the loan, rather than as actual dollar amounts as above. Thus the interest rate described above could have been expressed as "5/6 of 1% per month" or "10% per year."

Percentages, of course, are completely interchangeable with fractions or decimals, so we could express "10% per year" as "0.10 per year" (the suppressed decimal in 10 is moved two places left) or "10/100 per year" (the 10 is placed over 100), which reduces to "1/10 per year." Also "5/6 of 1%" is the same as "5/6%" (though this is a little hard to read) and can be written as "0.833%" or "0.00833" or "5/600" (which reduces to "1/120").

When interest is added to a loan to find the total owing, we can divide the new sum by the original to find the *growth factor*. For example, suppose the interest after one year is added to that $3,000 loan discussed above. The interest rate was 10% per year, so in one year the interest is $300 and the total owing is $3,300. Dividing $3,300 by $3,000 we have the growth factor 1.1 or 110%. If you subtract 1 or 100% from these, as appropriate, you have the interest back (1/10 or 10%). This works in the other direction also—add 1 or 100% to the interest, as appropriate, and you have the growth factor.

Exercise 1. The annual percentage rate on a loan of $1,000 is 9%. What is this interest rate as a decimal? as a fraction? What is the monthly percentage rate? How much is the interest for four months? five years? What are the growth factors for these two periods of time?

Suppose money is deposited in an account to accumulate interest. To be specific, let $1,000 be deposited for one year at 12% interest. (This may be usurious, but the numbers are convenient for illustration!) One possibility is that the interest earned will not be credited until the end of the year. Thus at that time the $1,000 will have earned $120 interest and the total value will be $1,120. Now suppose the depositor makes the following argument: "If my $1,000 earns $120 in one year, then it earns $60 in six months. I *want* the $60 after six months—or else I want it added to what you owe me, and interest calculated on it." The depositor would get slightly more at the end of the year this way, since the $60 credited after six months would itself be earning interest for the second half of the year, namely $3.60. The total value of the deposit would then be $1,123.60.

Now suppose the depositor is greedy and wants the interest reckoned after each month. Then in each month his money earns 1% (i.e., 1/12 of the annual interest rate of 12%). So after a month $10 is added to the account and $1,010 earns interest the second month; this earns $10.10 in a month, and so during the third month $1,020.10 is earning interest, etc. At the end of the year $1,126.83 will be due, a few dollars more than under the previous plan.

If the depositor is *still* greedier and wants the interest earned each *day* to be added, or each minute, or each second, the amount in the account at the end of the year will clearly become greater and greater. But how great can it get if the depositor is infinitely greedy? Is there an upper bound on how much interest can be earned no matter how frequently the depositor demands the interest be compounded?

There is a bound. Here's why: At the end of the year the account will include the original $1,000 plus the interest on it, plus some interest on this interest, plus interest on the interest on the interest, etc. Even if this process is continued to infinity, it turns out a finite sum is obtained.

The 12% interest on the $1,000 will itself be earning interest for at least part of the year—actually, if compounding takes place many times, bits of it will earn from nearly the beginning and other bits not until near the end. But in no way could the interest on the interest exceed 12% of the interest (12% of $120, or $14.40), because it would only earn the full 12% if it were *all* invested for the entire year. Now this second-level interest, the "interest-interest," so to speak, also earns interest for part of the year, but for the same reason as above the interest earned by the interest-interest cannot exceed 12% of itself (12% of $14.40, or $1.73).

This reasoning may be continued indefinitely. If we can sum this series of diminishing interests on interests to *infinity*, we will surely have an upper bound on the total the borrower would be liable for at the end of the year. This bound is independent of the number of times compounding is required, because it is a valid bound for every such number. For we have overestimated at each stage, and overestimated the number of stages, which of course would inevitably be finite no matter how often the depositor actually demanded compounding. The upper bound is

$1,000 + 12% of $1,000 + 12% of 12% of $1,000 + ...

This is a simple geometric series with first term $1,000 and common ratio 0.12; its sum, by a well-known formula from high-school algebra, is the first term divided by 1 minus the common ratio: $1,000/(1 - 0.12) = $1,000/0.88 = $1,136.36.

To see this directly, in case you've forgotten the formu[illegible]t

$$S = 1,000 + 0.12(1,000) + 0.12^2(1,000) + \dots + 0.12^n(1,000) + \dots$$

Multiply by 0.12 on both sides:

$$0.12S = 0.12(1,000) + 0.12^2(1,000) + \dots$$

Each term is turned into its successor. Now subtract, and we obtain 0.88S = 1,000, as before.

(Strictly speaking, we have *assumed* here that the series converges, and merely determined what it *would* converge to if it converges at all. It is not hard to prove rigorously that it does indeed converge. It is done in many elementary textbooks. We use almost the identical trick in the section below on annuities.)

Let us now proceed algebraically, so that our results have more general validity. Suppose an amount P is invested for n years at an annual interest rate of r.

Suppose that compounding occurs t times per year. Then each time interest is compounded the rate is r/t. After the first period the interest is (r/t)P, and this is added to P for the second period:

$$P + \frac{r}{t}P = P(1 + \frac{r}{t}).$$

After the second period, *this* amount is multiplied by (1 + r/t) (the 1 times P(1 + r/t) because the principal is still included, plus r/t times P(1 + r/t), which is again the interest). At the end of each period the amount on account is multiplied by another factor of (1 + r/t). In n years there are nt such periods—hence at the end of n years the amount that P has grown to is

$$(1) \qquad A = P(1 + \frac{r}{t})^{nt}.$$

This is a very important and useful formula in its own right, but we are interested in it for help in answering the question posed above: What is the effect of indefinitely increasing the frequency of compounding? We want, in the jargon of calculus, to take the limit as $t \to \infty$.

Exercise 2. Let $5,000 be deposited at 8% annual interest. Find out how large the account will be in five years if compounding takes place (a) once; (b) annually; (c) quarterly; (d) daily; and (e) every minute!

Exercise 3. Suppose a $5,000 loan is paid back double in five years. Find the annual percentage rate of interest if compounding takes place (a) once; (b) annually; and (c) daily.

It is shown in calculus classes that the expression $(1 + z)^{1/z}$ approaches some limit as z gets very small. The argument rests on the twin facts that the quantity continues to increase as z shrinks, but not without bound; 3, for example, is easily shown to be a bound. This argument shows that the limit must exist, but it does not reveal what its value is, except that it is obviously not greater than any bound. This limit, whatever it may be, is denoted "e," and it turns out by further considerations we can determine its value to be approximately 2.71828....

It would be nice to nail down exactly which number this is, but alas it is irrational (this can be shown without much difficulty), so it is not possible to express it as a quotient of integers, nor to give it as a repeating decimal. As a matter of fact it is a *transcendental* number, which means it is not even the root of any polynomial with integer coefficients (this is somewhat more difficult to prove; it was first done by Hermite in 1873). So it is also not possible to nail the number down by specifying such a polynomial, as we can do for example with $\sqrt{2}$, which satisfies $x^2 - 2 = 0$, and is basically *defined* by the property of satisfying this equation.

So we know that 1 plus a small number, raised to a power that is the reciprocal of that small number, approaches something which we have chosen to denote by e. It isn't *really* essential to understand where e comes from; the main things are that it's a number, its value is about 2.718, and $(1 + z)^{1/z}$ is close to it when z is small.

This isn't quite enough to answer the limit question about interest, however. In our expression $P(1 + r/t)^{nt}$, it is true that r/t is getting small as $t \to \infty$, but we are not raising (1 + r/t) to the power t/r, which would be the reciprocal of r/t. If we were, the limit would be e. This difficulty is easily remedied. Let us rewrite $P(1 + r/t)^{nt}$ as $P[(1 + r/t)^{t/r}]^{nr}$, using the well-known rules of exponents. The part in brackets goes to e as t increases; so the limit we seek is Pe^{nr}.

Let us recapitulate what this quantity means: This is the *limit* of the amount to which an investment of P could grow in n years at an annual interest rate of r, as the frequency of compounding increases indefinitely. It is higher than any amount that could be obtained by any finite number of compoundings, however large. It is also the smallest number that has this property—it is just exactly large enough. We might say picturesquely that this expression represents the effect of *compounding continuously*—that is to say, compounding at every instant. It is as though interest is added exactly as it is earned—there is no waiting period at all before adding the interest so it begins earning interest itself. In a sense it is the natural and inevitable conclusion of the greedy lender's line of argument that he should have any interest credited the very moment it is earned so that it in turn can begin to earn interest.

It is important to realize that all of this does not make one formal whit of sense! It is a bit like trying to talk about the bottom of a bottomless pit, or the smallest positive fraction. If you're going to compound interest you have to wait *some* amount of time or there is nothing to compound. But any length of time you actually wait is too long, as there would always be a shorter time. There is simply no way to compound "as often as possible." Nonetheless, the concept of continuous compounding has intuitive appeal, and is in any case formally defined as above in a way that is thoroughly logical and defensible.

The behavior of the total investment's value over time under continuous compounding is commonly referred to as *continuous growth*. Note that the quantity changes continuously now, whereas when we compounded t times per year it grew in discrete jumps and was constant in between.

To summarize then: When an investment P is allowed to grow for t years at an annual interest rate r compounded continuously (or where an initial value P experiences continuous growth for t years at an annual percentage rate of r), the final value is

(2) $$A = Pe^{rt}.$$

Though there are obvious conceptual difficulties here, and the concept of continuous growth takes some getting used to, computationally it is much easier to deal with than discrete compounding, especially for frequent compoundings. Compare, for example, the calculation of daily interest for one year at an annual percentage rate of 12% versus the calculation of continuous interest:

Daily	Continuous
$\$1,000 \times (1 + \frac{0.12}{365})^{365}$	$\$1,000 \times e^{0.12}$
$= \$1,127.42$	$= \$1,127.50$

Incidentally, note also that as expected the continuous compounding gives a slightly higher value.

Many hand calculators have keys for calculating these quantities easily. To calculate e^x where x is some number, punch in x on the keyboard and then hit the e^x key. This was the method used in the right-hand calculation above. The calculation on the left is facilitated by the y^x key: after dividing 0.12 by 365 and adding 1, hit y^x, then punch in 365 followed by the = key.

Exercise 4. If $1,200 is loaned for 20 years at an annual percentage rate of 5% compounded continuously, how large does the account grow?

Exercise 5. Refer to Exercises 2 and 3 and answer both questions if compounding takes place continuously.

---

2. APPRECIATION

Though the formulas for continuous compounding are useful in actual interest problems, they are even more useful for understanding growth in value generally. Suppose, for example, that a piece of property appreciates in value from $30,000 to $45,000 over a five-year period. It is not reasonable to regard the increase of $15,000 in value as having occurred in five equal jumps of $3,000 separated by intervals of one year. One would expect that each year it would increase a fixed *percentage* in value. This is because any property would presumably be increasing at any moment at a dollar rate that would be proportional to its value, and this in turn stems from the obvious fact that all individual dollars of this value would be growing at the same fixed rate.

Call the annual percentage increase r, so that in five years the property is worth $\$30,000(1 + r)^5$, which must equal \$45,000 by our hypothesis. Thus r can be solved for: $(1 + r)^5 = 1.5$, $1 + r = \sqrt[5]{1.5} = 1.084472$ (using the y^x key again), and so $r = 8.4472\%$. What we are saying is that the succession of values at the end of each year should form a *geometric* progression rather than an *arithmetic* progression. The values increase in fixed ratio rather than by a fixed difference.

Even this figure we have obtained, however, is misleading, for it suggests that the property took a discrete jump upward in value at the end of each year and remained constant otherwise, whereas presumably its value would have changed smoothly and continuously throughout the five years. We use the formula Pe^{nr} with $P = \$30,000$, and $n = 5$. Set this equal to \$45,000 and solve for r:

$$e^{5r} = 1.5; \qquad 5r = \ln 1.5; \qquad r = 8.1093\%$$

(The "ln" operation is, so to speak, the "opposite" of the "e^x" operation, and it can also be accomplished on most hand calculators by pressing a button.)

Note that in one year at continuous compounding with this r, the growth factor is $e^{0.081093} = 1.084472$, so the actual increase in a year is 8.4472% as shown in our earlier calculation.

Exercise 6. If a \$100,000 investment experiences continous growth at an annual percentage rate of 12%, what will it be worth after two years?

Exercise 7: If a house is purchased for \$40,000 and sold after three years for \$60,000, what was the annual growth rate r? By what percentage did it actually increase in each year? (This can be computed by two methods, which should agree.)

3. THE "RULE OF 72"

Investors often use the doubling-time of an investment as a good measure of how profitable the investment is. The more rapidly you can double your money, of course, the more desirable an investment is.

According to our formula, e^{nr} is the growth factor by which the value of an investment will grow in n years at rate r. When this factor is 2, the investment has doubled: $e^{nr} = 2$, so $nr = \ln 2 = 0.6931$. If r is expressed as a percentage, then $nr = 69.31$, i.e., the doubling time in years times the interest rate is about 69. Oddly enough, *this* is the "Rule of 72"! Actually, the figure of 72 is more reasonable in place of 69 if r denotes the equivalent annual increase after allowing for continuous compounding, instead of the actual instantaneous growth rate—for example (see above), if $r = 8.1093\%$, then in one year at this growth rate there will be an increase in value of 8.4472%. The slight difference between these two figures will make the product roughly 72 instead of 69. So the "Rule of 72" says

> "The product of the doubling time by the actual annual growth rate is 72."

For example, money will double in 12 years at 6%, in 8 years at 9%, etc.

The Rule of 69 is always valid, but, strictly speaking, the Rule of 72 is not. It holds approximately when the interest rate is about 8%, but breaks down for values of r much greater or much less than this. In practice, it "always" holds.

Exercise 8. About how long will it take an investment to double at an effective annual percentage rate of 10%?

Exercise 9. At what annual percentage rate is an investment growing that doubles in four years?

---

4. ANNUITIES

Next, let us consider a slightly more elaborate problem. Suppose that regular monthly additions are made to an interest-bearing account. We will develop a formula for the increasing sum in the account as time passes. For convenience, let us assume that interest is compounded monthly. If the payments into the account are made regularly at some other interval than monthly, and the interest is compounded at the same interval (or more often depending on the situation), all our formulas will continue to hold with suitable reinterpretation of the basic unit of time.

Let G be the amount added to the account each month and r be the *monthly* interest rate (which can of course be obtained from the annual percentage rate by dividing by 12). We will calculate S_n, the amount in the account during month n. During month 1, there is G on deposit, earning interest rG. At the end of the month, this interest is paid and a second sum of G is deposited, bringing the total to G(1+r) + G on deposit during month 2. At the end of this second month, this amount is multiplied by (1+r) and another G added, so that during month 3 the amount on deposit is $G(1+r)^2 + G(1+r) + G$. Continuing likewise, we infer that during month n there is on deposit the amount

$$S_n = G(1+r)^{n-1} + G(1+r)^{n-2} + \dots + G(1+r) + G.$$

You should be able to visualize this expression as the sum of the following (in reverse order): a new deposit G, a month-old deposit of G with its interest (simple), a two-month-old deposit of G with its interest (twice-compounded), and so on up to the original deposit with n-1 months' interest.

Now we can use an old trick to express this quantity more conveniently. (This is the trick we alluded to in the first section for summing a series.) Let both sides be multiplied by 1 + r. We obtain

$$(1+r)S_n = G(1+r)^n + G(1+r)^{n-1} + \dots + (1+r)G$$

Now we subtract the original equation from this:

$$rS_n = G(1+r)^n - G.$$

Note that most of the terms cancelled because there were so many in common, and we are left with a closed-form expression. Thus we achieve our final result:

$$S_n = \frac{G}{r}[(1+r)^n - 1]. \tag{3}$$

This is the formula we have been seeking. Remember, r is the *monthly* interest.

This formula is of considerable value in its own right, but it can be turned to even greater use in understanding mortgages or, for that matter, any loan which is paid off in regular installments and on which interest is computed only on the remaining balance.

Exercise 10. If deposits of $100 are made monthly to an account paying 5% annual interest, compounded monthly, how much is in the account after two and one-half years?

---

5. MORTGAGES

Let us consider a typical mortgage situation. Suppose someone takes out a $20,000 loan at 9% annual interest and a monthly payment obligation of $200. (These figures, by the way, are totally hypothetical and are chosen for convenience in illustrating the theory.) After one month, the interest is 3/4 of 1% or $150. (As usual, we divide 9% by 12 to get the monthly interest rate.) Of the $200 payment, then, $150 goes toward interest and the remaining $50 reduces the principal. In the second month the debt is $19,950. At the end of the second month, the interest is $149.62. Out of the $200 payment there remains now $50.38 to go toward the principal, so in the third month the debt is $19,899.62. Projecting ahead several months, here's what happens:

End of Month	Interest	$ Toward Principal	New Balance
3	$149.25	$50.75	$19,848.87
4	148.87	51.13	19,797.74
5	148.48	51.52	19,746.22
6	148.10	51.90	19,694.32
7	147.71	52.29	19,642.03
8	147.32	52.68	19,589.35

The calculations can be quite tedious if performed one-by-one like this, and it is hard to see a pattern. But . . . mathematics to the rescue! The matter is simplified substantially if we resort to the following trick: Let us imagine that each month's payment M goes first to pay off the interest on the principal as usual, but that the rest will be considered a counter-loan instead of a reduction on the original principal. Let us suppose further that this counter-loan earns interest at the same rate as the principal. The principal loan balance thus remains unchanged under this artifice, but in actuality the borrower's liability at any time is the difference between the main loan balance and the counter-loan, which is equivalent to the amount calculated more "naturally" above. By the distributive law, the interest calculations would be equivalent: $r(L - C) = rL - rC$, i.e., the left-hand side is the interest on the remaining balance as normally calculated, whereas the right-hand side represents the net interest owed by the borrower (the difference between the interest due on the original, unreduced principal and the interest earned on the counter-loan).

The advantage of viewing the situation in this manner, is that, although nothing has been changed in substance, we can now apply our annuity formula, since that is, in effect, what the counter-loan is.

To clarify matters, let's take a look at a specific loan and examine the first few months of bookkeeping, computed according to both the conventional and the new viewpoints. Note that the "net" column of Method 2 agrees with that of Method 1.

Loan: $20,000 @ 9% per year; $200 monthly payment

METHOD 1 (in $)		METHOD 2 (in $)				
		Principal		Counter-Loan		Net
Month 1:	20,000	Month 1:	20,000		0	20,000
interest	+150	interest	+150		0	+150
payment	-200	payment	-150	"deposit"	50	-200
Month 2:	19,950	Month 2:	20,000		50	19,950
interest	+149.62	interest	+150	"interest"	0.38	+149.62
payment	-200	payment	-150	"deposit"	50	-200
Month 3:	19,899.62	Month 3:	20,000		100.38	19,899.62

We can now make the following algebraic generalization: Let L be the original loan, M the monthly payment, and r the *monthly* interest rate. (We keep emphasizing *monthly* rate when we use it, because the federal Truth-in-Lending Law not only requires [in contracts, not math books] that all interest rates must be expressed in the form of an Annual Percentage Rate (APR), but also vaguely suggests that it is immoral or crooked to do otherwise! We'll discuss this matter more fully below.)

The monthly interest due on the principal L is rL, leaving (M-rL) to go into the counter-loan. This is the "G" of our earlier work, so after the nth payment, the value of the counter-loan is $[(M-rL)/r][(1+r)^n-1]$, and the actual balance due (B) is L minus the counter-loan:

$$(4)\qquad B = L + \frac{rL-M}{r}[(1+r)^n-1]$$

$$\text{or}\qquad B = L(1+r)^n - \frac{M}{r}[(1+r)^n-1].$$

(In Formula (4), the expression rL-M will normally be negative, and it might seem more natural to write ". . . L - (M-rL)/r . . ." However, the form given is more suitable for use with a hand calculator.)

This formula can be used as it stands for two purposes. One is to find the actual balance due on a loan after n months of regular payment—*without* calculating all the intermediate balances. For example, after 10 years or 120 months of payment on the loan used as an illustration above, the remaining balance would be $10,324.37. (Here again, the y^x key on a hand calculator comes in handy for computing $(1+r)^n$.)

Exercise 11: If monthly payments of $150 are made on a $15,000 loan at an annual percentage rate of 10%, how much of the principal remains after 1 year? 2 years? 5 years? 10 years? 200 months?

The other purpose is to find a "balloon" payment—except that this is really the same thing. (And a good thing, too, that "balance" and "balloon" both begin with B!) If you agree to make payments on a loan for a certain period that is insufficient to amortize it (that is, pay it off completely), then a rather large sum may be due at the end of the term agreed upon. For example, if the above loan had a 10-year maturity, the balloon payment would be $10,324.37.

Exercise 12. If monthly payments of $70 are made on a loan of $7,000 at an annual percentage rate of 9%, and the entire balance is due in three years, what is the balloon payment?

There are numerous other uses of our formula for the declining mortgage balance. Let's look at one more. Suppose that after 5 years of making $200 monthly payments on that $20,000 loan @ 9%, an extra $2,000 is paid against the principal—perhaps, for example, out of an inheritance. What effect does this have on the length of the loan, the future behavior of the declining balance, and the total interest paid?

First, we calculate the balance after 60 months: $16,228.81. Then, we subtract the $2,000, so the principal becomes $14,228.81. Now we simply start over with this as our new L. From this point on, the loan will take 102 months (8 years and 6 months) to amortize. (The rule for determining this, as well as the original length of the loan, is discussed below.) Since the original loan would have taken 15 years and 6 months to pay off, the total time has been shortened by 24 months (2 years), which means 24 payments of $200—a considerable savings!

(Do not be misled, however: $2,000 has been tied up for many years in order to achieve this saving. In fact, using our counter-loan techniques, we see clearly what is happening: The savings of $4,800 is exactly equal to the interest and principal that would be generated over the remaining life of the loan by a $2,000 bank account earning 9% interest. Or to put it another way, if instead of using that $2,000 as payment on principal, it were invested at 9% per year (compounded monthly), then it would be exactly sufficient to pay off the last 24 monthly loan payments as they came due. If you can find a place to put the $2,000 where it earns more than 9%, you're better off doing that and pocketing the difference. If you can't, you're better off making the principal payment.)

By modifying Formula (4) in certain ways, it can be put to other important uses. Basically, what we can do is solve for any one of the variables in terms of the others. Suppose we want to know how long it will take to amortize a loan completely. This means we want the loan balance to be zero, i.e., $L - [(M-rL)/r][(1+r)^n - 1] = 0$, or $L = [(M-rL)/r][(1+r)^n - 1]$. Notice the second form of the equation merely states that the counter-loan equals the original loan. We solve for n:

$$\frac{rL}{M-rL} = (1+r)^n - 1,$$

$$(1+r)^n = \frac{rL}{M-rL} + 1 = \frac{M}{M-rL}, \text{ and so}$$

$$(5) \qquad n = \frac{\log \frac{M}{M-rL}}{\log (1+r)}.$$

It is immaterial which base the logarithms are as long as they are the same. The ratio of logs is always independent of base. Normally one would use either common logs (base 10) or natural logs (base e). There is one advantage to using natural logs—when r is small, as it usually will be in practice, ln(1+r) is *almost* equal to r. This is not true in any other base. So

$$(6) \qquad n \cong \frac{1}{r} \ln \frac{M}{M-rL}.$$

The quantity M/(M-rL) is the ratio of the monthly payment to the portion of it that represents gain on principal after paying interest the first month. This might be called the "Payback Ratio." We will refer to it as the PBR. For instance, in our example, the PBR is 200/50 = 4. To compute the amortization period of this loan, we evaluate (log 4)/(log 403/400) and get 185.53. This means that in 186 months (or 15-1/2 years) the loan is paid off.

The last payment would not need to be a full one—that is the significance of the fractional part of this number. To find the exact amount of the final payment, compute the remaining balance after 185 months, using Formula (4): $105.88. Add the interest on this, which is $0.79, and you have the final payment, $106.67.

Exercise 13. For the loan of Exercise 11, find out how many months will be required to pay it off, and what the amount of the final payment will be. Do the same for the loan of Exercise 12 (assuming it is allowed to run out to maturity).

Suppose you are planning a loan (L) and you know the interest rate (r) and how long you want the loan repayment to take (n). We can calculate the required monthly payment (M) as follows: We still have

$$L = \frac{M-rL}{r}[(1+r)^n - 1].$$

We solve now for M:

$$M = \frac{rL}{(1+r)^n - 1} + rL = rL\left[\frac{(1+r)^n}{(1+r)^n - 1}\right], \text{ or}$$

$$(7) \qquad M = \frac{rL}{1 - 1/(1+r)^n}.$$

Exercise 14. If you need to borrow $50,000, if the interest rate is 11%, and if the lender will amortize over 25 years, find the required monthly payment.

On the other hand, if the monthly payment (M) is fixed and we want to know how much of a loan (L) we can swing given an interest rate (r), amortized in n months, we solve for L instead:

$$(8) \qquad L = \frac{M}{r}\left[1 - \frac{1}{(1+r)^n}\right].$$

Exercise 15. If you have $250 per month to pay on a loan, if the interest rate is 9-1/2%, and if the lender will amortize over 30 years, how much can you borrow?

There are circumstances when a loan is given with interest, but the specific interest or interest rate is not explicitly provided, and sometimes the actual amount of the loan is not even made entirely clear. The lender is considered paid off when a certain number (n) of monthly payments (M) have been made.

In saying this, we are taking a very strict point of view, looking closely at what people are actually doing rather than what they are saying. Rarely does anyone intentionally not specify an interest rate, and usually there is one floating around even if for possibly obscure reasons. The lender may have one in mind and, based on it, may do some calculations that satisfy him as to what he expects from the borrower. But, unless the lender has been careful to do the calculations exactly as we described them above, his figure may bear little relation to the actual effective interest rate that the borrower is paying, and thus, from a strictly mathematical point of view, the real interest rate has never been made explicitly clear.

What we are going to look at, then, is the problem of calculating exactly what the effective rate (r) of interest is when we are given: (1) how much the borrower actually gets (L); (2) how big his loan payments are (M); and (3) how long he has to make them (n). We will make our computation without listening *at all* to what the lender did to figure out which values of M and n will satisfy him.

This is the kind of calculation that must be made in order to comply with the federal Truth-in-Lending Law, since all interest charges, however the lender may arrive at them and however he may think of them, must be expressed as though calculated in a standard manner, namely, the way we have done it. This is called "the actuarial method." Moreover, when the actual effective rate of interest is finally determined, it must be stated in the form of an Annual Percentage Rate (APR). This is so that people will always see this important quantity expressed in the same units and thus will not be confused by having to compare numbers that really don't mean the same thing or measure in the standard way.

The formula for the equivalent monthly interest rate is found by solving our familiar equation (Formulas (5), (7), or (8) for r:

$$-M(1+r)^n + rL(1+r)^n + M = 0.$$

It is easier to solve for the growth factor 1+r, so we write

$$-M(1+r)^n + (1+r-1)L(1+r)^n + M = 0$$

or

$$L(1+r)^{n+1} - (L+M)(1+r)^n + M = 0.$$

When X denotes the growth factor 1+r and a is the ratio M/L, the equation becomes

$$X^{n+1} - (1+a)X^n + a = 0.$$

Unfortunately, though this is merely a polynomial of degree n+1, such an equation in general cannot be solved explicitly. Some method of estimating roots would have to be used, such as Newton's Method.

Newton's Method is a clever device for turning a guess at the root of an equation into a better guess. By repeatedly using the method, it is possible, in just a few steps, to come extremely close to the root—so close that, for all intents and purposes, we "have" it. The formula is a bit messy in this case, but with a hand calculator the arithmetic is reduced to pushing a few buttons.

Let

$$f(x) = x^{n+1} - (1+a)x^n + a = a - x^n(1+a-x).$$

Then

$$f'(x) = [(n+1)x - n(1+a)]x^{n-1},$$

and the Newton's Method formula for producing a better guess x' out of an original guess x is:

$$x' = x - \frac{f(x)}{f'(x)} = x - \frac{a - x^n(1+a-x)}{[(n+1)x - n(1+a)]x^{n-1}}$$

$$= \frac{[nx - (n-1)(a+1)]x^n - a}{[(n+1)x - n(1+a)]x^{n-1}} .$$

Let our first guess be 1+a. Then

$$x' = 1 + a - \frac{a}{(1+a)^n} ,$$

which suggests that

$$r = a \frac{(1+a)^n - 1}{(1+a)^n} .$$

We could apply the method again using this value for x and calculate a new x'. At this point it is easier to do numerically than algebraically.

In many cases it is just as easy to guess values of r and see how close they come. By repeated guessing we may be able to come very close to the true value of r. Let us look at an example: Suppose you buy a car for $2,000 down and carry $3,000. Suppose you make 30 monthly payments of $200, a total of $6,000. How much was the interest? Here, M = 200, L = 3,000, and our first guess might be that r = 5.7% per month (or 68.5% per year), based on the first approximation obtained by Newton's Method above. We can easily check how close we came. Let's use Formula (5) to figure how many monthly payments of $200 would be required to amortize a $3,000 loan at 5.7% per month. Then rL = 171.15, M-rL = 28.85, the PBR is 6.93, and the amortization would take 34 months. Since, in fact, the loan is paid off in 30 months, this trial interest rate must be too high. If we try r = 5% per month, the loan is amortized in 28.4 months. Thus this figure is too low, and the correct figure apparently is somewhere between 5% and 5.7%. On the basis of linear interpolation, we might figure that 30 is about 30% of the way from 28.4 to 34, so perhaps next we should try guessing a value of r that is about 30% of the way from 5 to 5.7, i.e., about 5.2%. With a hand calculator it would not take long to find r correct to the hundredths place by repeated trial and error of this sort.

Exercise 16. A used-car dealer tells a customer the following:

"We'll sell you this car for $600. You pay $100 down and we'll carry the rest at 14% interest for one year. So the loan will be $500 at an Annual Percentage Rate of 14%. [Note the Truth-in-Lending jargon creeping in.]

"Now, 14% of $500 is $70 [watch him carefully] so you'll owe us a total of $570 at the end of the year [here's where it starts to get fishy]. You pay us back in 12 equal monthly installments of 570/12 = $47.50." (You're paying interest on the whole amount for the entire year even though you're not keeping the whole amount the entire year. This calculation would be perfectly correct if you were only required to make one payment of $570 at the year's end.)

Calculate the effective APR this customer would be paying.

Exercise 17. The used-car dealer next door argues as follows: "We'll sell you this car for $600. You pay $100 down and we'll carry the rest at 14% interest for one year. Now, 14% of $500 is $70, so we'll take our $70 now [here's where he diverges from his neighbor], and you can pay us the $500 back in 12 equal monthly installments of 500/12 = $41.67." (This is called "front-end interest," because it is taken out before the loan starts.) The customer actually has to pay this dealer $170 now ($100 "down" and $70 "interest") so he's really getting a loan of only $430. No wonder the payments are lower!

Calculate the effective APR this customer would be paying.

Exercise 18. The customer above says he has only $100 to pay now, so he's got to borrow the full $500 even if it costs him more. The dealer says OK, he'll settle for $100 now plus 12 equal monthly payments of $41.67 *times* 500/430 (or $48.45): since the loan has increased by the ratio 500 to 430, so should the monthly payments.

Calculate the effective APR this customer would be paying.

Exercise 19. A lender makes a "nominal" loan of $10,000 at a "nominal" APR of 10%. But he charges a "loan fee" of $1,000, so the borrower actually receives only $9,000. Yet his repayment schedule is as though based on the nominal value of the loan.

(a) Calculate how long he would have to pay $150 per month to pay off a $10,000 loan. (This actually is how long he will have to pay to meet the lender's terms.)

(b) Calculate his effective APR if, in order to satisfy a $9,000 loan, he paid $150 per month for the number of months calculated above.

Another use of this technique for finding an effective annual percentage rate, which is closely related to Exercise 19 above, is to compute the effective yield of a discounted note. Sometimes an investor who holds a note as security on a loan finds that he needs the money and must sell the note to another investor. To "sweeten the pot" and induce another investor to buy his note, he may sell it at a discount, that is, he may accept less than the actual balance due at that time. Or, when a lender originally accepts a note from a borrower as security for a loan, he may require that the borrower pay a certain fee. This results in the borrower actually receiving less of a loan than the face value of the note.

What may complicate both of these situations is that the note may require full repayment prior to amortization, so there would be a substantial balloon payment.

Suppose, for example, that the loan in Exercise 19 must be repaid after three years. We can use Formula (4) to calculate the nominal balance after 36 months of paying $150 per month on a nominal loan of $10,000 at a nominal annual percentage rate of 10%. In this case, L = 10,000, r = 1/120, M = 150, and n = 36, so the balance is $7,214.59, the actual balloon payment due. But remember that the loan was only $9,000, so the effective annual percentage rate must have been somewhat higher than 10%. Let us use the trial and error method to find it. We take L = 9,000, M = 150, n = 36, with various trial values of r, and calculate the balance by (4), trying to get as near as possible to $7,214.59:

When

r = 12%, balance due is $6,415.39—too low
= 13%, $6,703.50—still too low
= 14%, $7,001.01—still a bit low
= 14.7%, $7,214.99—*very* close!

So, 14.7% is the effective annual percentage rate that is actually paid by the borrower and actually received by the lender.

Exercise 20. A borrower wants an actual $18,000, and is willing to pay a nominal 10% annual percentage rate, and a 20% discount. What would be the nominal or face value of the loan? Find the effective annual percentage rate if the loan is paid back in full at the end of: one year; two years; and five years. Assume a monthly payment of 1% of the nominal value of the loan.

Exercise 21. The face value of a note is $5,500, and the nominal annual percentage rate is 10%. Assume the borrower pays a discount of 10%, but that a loan broker takes 2% and leaves 8% for the lender. What actual amount does the borrower get? What actual amount does the lender put up? Assuming a monthly repayment of $55, carried on for two years until a balloon payment is due, find the balloon payment, and the effective annual percentage rates for both borrower and lender (they will be different).

Exercise 22. A three-year note for $11,500 at 8% annual percentage rate is taken without discount by an investor. He receives monthly payments of $115 for two years, then sells the note at a 10% discount (off its current value). How much does he get? What has been his effective annual percentage rate of return? What is the ballon payment when the note is paid off after another year? What is the note buyer's effective annual percentage rate of return?

6. CONTINUOUS APPROXIMATION

Anyone who has looked at a table of declining balances to find a pattern has probably been perplexed, but, nonetheless, has undoubtedly noticed at least that the balance declines slowly at first, and then more and more rapidly. For example, here is what happens to one of the loans we have been studying:

Initial Balance: $20,000.00

Paid in first year:	$ 625.37	leaving	$19,374.63 due
second year:	684.04		18,690.59
third year:	748.21		17,942.38
fourth year:	818.40		17,123.98
fifth year:	895.17		16,228.81
sixth year:	979.13		15,249.68
seventh year:	1,070.99		14,178.69
eighth year:	1,171.45		13,007.24
ninth year:	1,281.33		11,725.91
tenth year:	1,401.54		10,324.37

For readers who know calculus, we can make these observations more quantitative by the following device: Let B(t) be the declining balance. It actually decreases in steps of course, namely M-rB per month. But let us suppose it declines continuously at the monthly rate of M-rB. This is approximately accurate. We have, then, the equation

$$\frac{dB}{dt} = -(M-rB) \quad \text{or} \quad \frac{dB}{dt} - rB = -M.$$

If we multiply both sides by e^{-rt}, we obtain

$$e^{-rt}\frac{dB}{dt} - e^{-rt}rB = -Me^{-rt}.$$

This step may seem unmotivated, but it has the advantage of turning the left-hand side into the exact derivative of the product of the functions B(t) and e^{-rt}. This may be checked by the well-known product rule of calculus.

So, whatever B(t) may be, at least we know that $(Be^{-rt})' = -Me^{-rt}$. Thus,

$$Be^{-rt} = \frac{M}{r}e^{-rt} + C,$$

and so

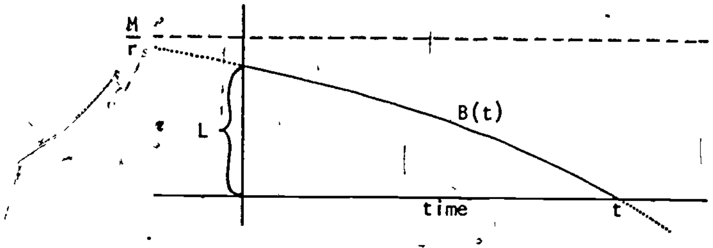
$$B = \frac{M}{r} + Ce^{rt}.$$

The C is a constant introduced in taking the antiderivative. It can be evaluated by letting t = 0, and using the observation that B(0) is L, the initial loan. The equation becomes L = M/r + C. Hence, C = L - M/r, and, at last, we obtain

$$B(t) = \frac{M}{r} + (L - \frac{M}{r})e^{rt}, \text{ or}$$

$$B(t) = \frac{M}{r}(1 - \frac{M-rL}{M}e^{rt}).$$

We see, then, that B is experiencing exponential growth down away from the value M/r. Note that the payback ratio (inverted) appears in this formula, as the coefficient of the exponential. We can graph B(t):



The value marked t, where B(t) crosses the time axis and vanishes, is the time it takes to amortize L. This will occur when B(t) = 0, i.e., when e^{rt} is equal to the payback ratio, or t = (1/r) ln PBR. Interestingly, this was the approximation we obtained in an earlier section, Formula (6).

7. PRESENT VALUE OF FUTURE PAYMENTS

Suppose you sell your car today for $1,000, but the buyer asks you to accept payment in a year. This amounts to your giving him a loan of $1,000 for one year. He should pay interest to you at the prevailing rate. Conversely, if he does not, then you aren't really getting $1,000 for your car: you're getting whatever amount it is that would grow to $1,000 in one year at the prevailing interest rate. Think of this another way: Suppose the buyer signs a note agreeing to pay $1,000 in a year—what would the note be worth now if you tried to sell it? You would have to sell it at a "discount" in order to induce someone else to buy it, for after all the other party would expect to make some profit during the year he's going to tie up his money.

To find out exactly what $1,000-one-year-from-now is worth today, let's reason as follows: If r is the prevailing interest rate, and an amount A is invested today, it would grow to Ae^{r} in one year. For this to equal $1,000, A must be $1,000e^{-r}$. More generally, if a payment P is expected in t years, it is worth Pe^{-rt} today. The payment is discounted by the factor e^{-rt} to account for the forfeited interest in the meantime.

Exercise 23. Suppose a forest could be cut down now and a profit realized of $3,000 per acre; or, it could be allowed to grow for five years and then harvested at a profit of $5,000 per acre. Which is better? Try discount rates of 10% and 12%.

If there are several payments involved at different times, say P_k at time t_k, then the present value of these is

$$\sum_{k=1}^{n} P_k e^{-rt_k}.$$

Suppose the payments P_k are all equal, say to P, and the payments are equally spaced, say at intervals of T, so that $t_k = kT$. Assume they start now and continue indefinitely. The formula above becomes

$$P\sum_{k=0}^{\infty}(e^{-rt})^{k},$$

which is a geometric series and sums to $P/(1-e^{-rt})$.

Exercise 24. Suppose you win a sweepstakes and are offered your choice of the following prize options:

(a) $100,000 now;
(b) $50,000 now plus $250 per month for life;
(c) $500 per month for life.

Calculate the present value of each of these, and determine which is best. Assume you will live a "long time," and make the calculations as though you will live forever. Try both r = 5% and r = 6% as the prevailing interest rate. (Ignore all income tax consequences.)

There is yet another method for treating this kind of problem if the payments are made more or less continuously—as, for example, income from a business, revenue from a toll facility, etc. This method requires concepts from the integral calculus. Let the function I(t) represent the rate of receipt of income at time t over an interval from time a to time b in the future. Let [a,b] be partitioned into a large number n of small intervals of length Δt_k and let $I(t_k^*)$ be the income rate at a typical time t_k^* in the kth subinterval. Then $I(t_k^*)\Delta t_k$ is approximately the income earned in the kth subinterval, its present value is $I(t_k^*)\Delta t_k e^{-rt_k^*}$, and the sum

$$\sum_{k=1}^{n} I(t_k^*)\Delta t_k e^{-rt_k^*}$$

is an approximation to the present value of this anticipated income. Since, as the partition becomes finer and finer, it is a better and better approximation, the integral

$\int_a^b I(t)e^{-rt}dt$ must be the exact present value of the anticipated income from I(t). If income is anticipated from the immediate present into the indefinite future, the improper integral $\int_0^\infty I(t)e^{-rt}dt$ gives the desired value.

(Note that this quantity is a function of r. The function so obtained from any given income function I(t) is generally known as the Laplace Transform of I(t), and is widely studied in engineering and applied mathematics.)

Suppose, for example, it is projected that a shoe store will make a net profit of $1,000 per month, indefinitely. Thus, I(t) = 12,000, if t is in years. The present value of this shoe store is $\int_0^\infty 12{,}000e^{-rt}dt = 12{,}000/r$. At 10% discount, then, it would be traded fairly at $120,000. Conversely, if the asking price was $200,000, this would imply a discount rate of 6%.

Exercise 25. A proposed hydroelectric power plant will cost $50,000,000 to build. Is it worth building if the rate at which it will generate revenue after t years is $1,000,000 $\sqrt{t}$ per year? Note the revenue rate will continue to increase indefinitely, starting at 0, but its rate of increase will gradually slow down. Assume r = 6.75%. (Solution of this exercise requires knowledge of the gamma function.)

If there are lump sum payments superimposed upon a continuous flow of income, then the best technique for representing the present value of these future payments is an advanced mathematical construction known as the Riemann-Stieltjes integral.

Whether a sum or an integral is more appropriate to measure the present value of future income, discounting future payments in this way is a standard concept in contemporary economic theory. For example, in examining the wisdom of waiting another year for a cow to get fatter before it is butchered, or waiting a decade for a forest to grow larger before it is sawed down, it is unreasonable to compare the respective profits now and later as if dollars now and later are equivalent. Of course the profit will be nominally increased by waiting. But the matter becomes more realistic (and more interesting) if the delayed profit is discounted before being compared to the present profit, as illustrated above.

It is important to give careful thought to the value of r in the discounting formula. The exercises have shown how critical decisions may sometimes be reversed depending on the value of r. It should realistically represent the value of doing without the money for a while. This is not a mathematical matter, but it does have obvious mathematical consequences. It could be taken simply as the value of inflation—even the most conservative investor would at least expect his buying power to be restored after the waiting period. It could be taken as the current interest rate paid by savings and loans, or the current cost of borrowing prime money. Or it could be taken to be the interest you expect to receive from your investment program—which could be quite high, if you're a shrewd investor! You might, for example, take it to be the return rate calculated in the section below on the optimal time to hold an investment.

The discounting formula can be just as useful to the person paying as to the person being paid. However, the two parties to a transaction may have good reasons for using different discount rates, and therefore they might differ as to the true "cost" of postponing a payment. Actually, it's probably a good thing they differ, just as it's good that people place different "utility" values on various combinations of commodities—it's primarily this difference that makes trade possible, for each side can think it's getting a good deal from the same transaction.

For example, take the shoe store discussed above. If a buyer discounted future income at 10% and a seller only at 6%, the buyer would be willing to pay $200,000, while the seller would settle for $120,000. Thus they each would think that a price of, say, $150,000 for the business was fantastic, since it represents a compromise discount rate of 8%.

8. LEVERAGE

Most investors in real estate take advantage of the willingness of lending institutions to loan money generously against real-estate investments. The general stability of the real-estate market makes lending institutions relatively confident that their money will be safe, so they are willing to loan high fractions of the purchase price at relatively low interest rates. Thus, most investors actually own a rather low equity in their property. When there is an appreciation in value of the property, however, it all belongs to the investor, and this fact causes a surprising magnification in the rate of growth. Let us illustrate: Suppose an investor buys a $50,000 property with $10,000 down (his own money) and the balance of $40,000 financed (borrowed). Now, let's say in a year the property appreciates 10%, so it is worth $55,000. The investor still owes (roughly) $40,000, but now his equity is $15,000. In one year his $10,000 turned into $15,000. This is a 50% increase, five times the rate of increase of the property. The factor five is directly related to the amount of his bank loan—he bought a property worth fives times as much as his own investment. This phenomenon—the magnification of growth rates through borrowing—is known as leverage. The general formula is as follows: If you have put up 1/n of the money yourself and borrowed the rest, and if the holding appreciates by a factor of r, then your investment has appreciated by a factor of nr.

We are ignoring the obvious fact that an investor who borrows $40,000 has to make pretty hefty payments to the lender on this loan, most of which will be interest (if, as usual, the loan is amortized over 30 years). But, if we assume that the property under discussion generates income (e.g., an apartment house, an office building), then, roughly speaking, we can presume that the income will be enough to balance all the expenses including debt service.

The leverage principle actually applies to all investments, not just real estate, but it is of greatest significance in real estate due to the generally large values of n, as compared to the stock market, for example.

9. THE OPTIMAL TIME TO SELL

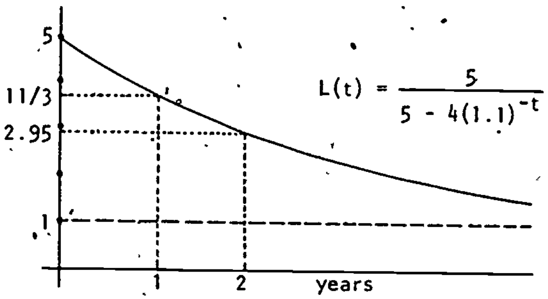
In the preceding example, the 50% increase in the investor's funds is, so to speak, "too good to last." His equity after one year is a much higher percentage than it was initially—$15,000 out of $55,000, as opposed to $10,000 out of $50,000—and during the second year his leverage is correspondingly lower. If the property goes up 10% more to $60,500 during the second year, his equity increases to $20,500. Thus, during the next year his equity goes only from $15,000 to $20,500, an increase of 37%—still very nice, of course, but not what it was. As his equity continues to go up over the course of time (which is what he desires, after all), his leverage will continue to decline, and percentagewise his investment will not be as good.

It might be suggested that after the second year he should sell his property and buy something more valuable. After all, with banks evidently willing to finance 80%, and with $20,500 for a down payment, he could buy a $102,500 property. Clearly, he is better off with the profits from the appreciation of a $102,500 property instead of a $60,500 one.

But if selling after two years is a good idea because of the fact that his leverage, which has fallen so low, can be restored, perhaps selling after one year would be an even better idea. After one year, he could realize $15,000 on selling his property (above the mortgage), which as a down payment on a second investment would permit him to buy into a $75,000 property. After another year this would appreciate to $82,500 and he would have a gain of $22,500, a good $2,000 better than before. If he wishes, he could now sell again and buy into a $112,500 property as a third investment.

But again, if selling after one year is such a good idea, why not sell every six months when the average leverage is even higher? Or every month, or every day, or every minute?!

A graph can illustrate what is happening. The figure below shows the investor's leverage (ratio of the property's value to his equity) as a function of time:



Since this curve is decreasing as the property appreciates, clearly the sooner the investment is sold, the higher his *average* leverage.

If one pushes this reasoning to its logical extreme, it suggests buying and selling property every instant. It also suggests very clearly that there is an important factor we have not as yet taken into account: the actual costs of buying and selling. Every time a property is bought or sold there are substantial costs, such as title insurance, agent's fees, etc. These will place a lower limit on the length of time it is realistic to hold an investment.

Let us make a theoretical calculation of the optimal length of time to hold an investment. This will be good practice in mathematical modeling, too—and in understanding its limitations! We will have to make some simplifying assumptions in order to make the calculations tractable. Unfortunately, these assumptions also cast some doubt on the validity of the conclusions. They would be limited, to say the least. But on some bright day, when you are feeling adventuresome, you can try complicating the model to make it more realistic! In the meantime, at least we'll have some ballpark guesses as to the most prudent course for an investor to follow.

In buying a property there are certain costs of acquisition known collectively as "closing costs"—bank fees, title insurance, etc. To preserve generality, let's say they are c times the purchase price. In selling, there are also costs, largely the agent's fee. Let's say the selling costs are f times the selling price.

There are also certain continuing costs of maintaining ownership of a property, largely the cost of debt service (i.e., interest on the loan), but also taxes, insurance, maintenance, etc. These can be substantial, but it is reasonable to assume that there is income from the property to mitigate these, and not unreasonable to hypothesize, for the sake of convenience, that this income is exactly sufficient to balance the costs of ownership. Thus, our profit on the investment will derive exclusively from the appreciation, and be diminished only by the buying and selling costs. There are also tax consequences of both ownership and sale of a property, but we will ignore these, too, for simplicity in this analysis. (Incidentally, we are also neglecting any gain associated with repaying the mortgage during the period of ownership—but this is normally very small, especially over the initial period of a 30-year loan.)

Let us suppose we have an amount I to invest and we find a bank willing to finance all but 1/n of the purchase price P of a property. The closing costs will be cP, and we have available I-cP for the actual down payment. Thus, we can buy a property worth P = n(I-cP), that is, n times what we have available after allowing for closing costs.

We solve for P and find P = nI/(1+cn). Now we know what we can afford to buy with the available money.

Next, let us assume that the property continuously appreciates at the rate of r per year. In t years it is worth Pe^{rt}. If it is sold at this point, we pay fPe^{rt} in commissions and other costs of divestiture, and we pay off the bank mortgage, which is still pretty close to P(1-1/n), its original value. We have left

$$Pe^{rt} - fPe^{rt} - P(1-1/n) = \frac{I}{1+cn}[(1-f)ne^{rt}-(n-1)].$$

To find our gain per dollar, that is, our investment gain ratio, we divide by the input I, obtaining

$$(9) \qquad \frac{(1-f)ne^{rt} - (n-1)}{1+cn}.$$

It took t years to produce this gain. At a uniform rate of increase of x per year, the growth factor in t years would have been e^{tx}, as shown earlier. What is the value of x to which our gain is equivalent?

$$(10) \qquad e^{xt} = \frac{(1-f)ne^{rt} - (n-1)}{1+cn},$$

so,

$$x = \frac{1}{t}\ln\frac{(1-f)ne^{rt} - (n-1)}{1+cn}.$$

This x is a function of t and it is our goal to maximize it—that is, we sell at such a time that our average rate of gain is as large as possible. So we compute dx/dt and set it to 0:

$$\frac{1}{t}\,\frac{(1-f)rne^{rt}}{(1-f)ne^{rt} - (n-1)} - \frac{1}{t^2}\ln\frac{(1-f)ne^{rt} - (n-1)}{1+cn} = 0$$

or

$$\frac{(1-f)rtne^{rt}}{(1-f)ne^{rt} - (n-1)} = \ln\frac{(1-f)ne^{rt} - (n-1)}{1+cn}$$

This, needless to say, is a mess. To facilitate solving this equation, let us introduce the symbol y for $(1-f)ne^{rt}$, so that $rt = \ln y/n(1-f)$. Let us also set $w = y - (n-1)$.

With this notation, the condition becomes

$$\frac{y\ln\frac{y}{n(1-f)}}{w} = \ln\frac{w}{1+cn}$$

or

$$y[\ln y - \ln n(1-f)] = w[\ln w - \ln(1+cn)].$$

There is no hope of solving this equation analytically. However, the following technique is useful for approximating solutions.

Let c, f, and n have actual values; for example, c = 0.03, f = 0.06, and n = 5 are reasonable. Thus, ln n(1-f) = 1.5476, ln(1+cn) = 0.13976, and w = y - 4. Then make a table for various values of both y(ln y - 1.5476) and w(ln w - 0.13976), and see where the former at some y is equal to the latter at a w that is 4 less.

With these numbers it turns out that $y \cong 6.85$. Thus, $rt \cong 0.3766$. Or if you let n = 4, then $y \cong 5.70$ and $rt \cong 0.4160$. Now you merely plug in the actual average rate of appreciation and solve for t. Some examples are summarized in the table below:

	n = 5 (80% financing)	n = 4 (75% financing)
r = 0.10	t = 3 years, 9 mos.	t = 4 years, 2 mos.
0.12	3 years, 2 mos.	3 years, 6 mos.
0.14	2 years, 8 mos.	3 years, 0 mos.
0.16	2 years, 4 mos.	2 years, 7 mos.

It is interesting to note that the parameters fix only the product rt. Since the growth factor is e^{rt}, it appears that the optimal time to sell is when the growth factor reaches a certain value, irrespective of how long it takes or what the appropriate appreciation rate is. The value of e^{rt} at the "right moment" is y/n(1-f), and by Formula (9), the gain per dollar invested is then

$$\frac{y - (n-1)}{1+cn}.$$

From Formula (10) the effective annual percentage rate is one less than the t^{th} root of this, where t is the optimal length of time to keep the investment.

If you sell before these values have been reached, you make less profit and, more to the point, your average earning rate is less; if you sell after these values have been reached, you make more profit, but it takes longer and, it turns out, your average rate of earning is again less.

For the data above, if n = 5, the growth factor of the property at the optimal time to sell is 1.46, and each dollar invested will grow to \$2.48. The effective annual percentage rates are, respectively, 27.4%, 33.2%, 40.6%, and 47.6%. If n = 4, the growth factor is 1.52 and each dollar will grow to \$2.41. The effective annual percentage rates are 23.5%, 28.6%, 34.1%, and 40.6%, respectively.

10. SOLUTIONS TO EXERCISES

1. 9% = .09 = 9/100. The monthly rate is 3/4 of 1% or .75% or .0075 or 9/1200 or 3/400. The interest for 4 months is \$7.50 x 4 = \$30; for 5 years it is \$450. The growth factors are, respectively, 1.030 and 1.450. (Note that \$1,000, the actual amount of the loan, is only required for the computations of the actual interests in dollars.)

2. Using $P(1+r/t)^{nt}$ with P = 5,000, r = 0.08, n = 5, and
   (a) t = 1/5 (once in 5 years is 1/5 times per year), we get \$7,000
   (b) t = 1, we get \$7,346.64
   (c) t = 4, we get \$7,429.73
   (d) t = 1825, we get \$7.449.21
   (e) t = 2,688,000, beyond my calculator!

3. We know $5000(1+r/t)^{5t} = 10{,}000$, or $(1+r/t)^{5t} = 2$. Using various values of t, we solve for r:
   (a) t = 1/5, yields r = 20%
   (b) t = 1, yields r = 14.87%
   (c) t = 1825, yields r = 13.87%

4. Using Pe^{rt} with P = 1,200, r = 0.05, and t = 20, we have \$3,261.94.

5. Using P = 5,000, r = 0.08, and t = 5, we get \$7,459.12 for Exercise 2. For Exercise 3, we want $5000e^{5r} = 10{,}000$, or $r = 1/5 \ln 2 = 13.86\%$.

6. Using P = 100,000, r = 0.12, t = 2, we get \$127,124.90 .

7. $40{,}000e^{3r} = 60{,}000$, so $e^{3r} = 1.5$, $3r = \ln 1.5$, $r = 13.52\%$. In each year it appreciated by a factor of $e^{0.1352} = 1.1447$, so the increase is 14.47%. Note that 1.1447 is also the cube (3 years) root of 1.5 (the 3-year growth factor).

8. 7.2 years (72 ÷ 10)

9. 18% (72 ÷ 4)

10. With G = 100, r = 0.05/12, and n = 30 (the number of months in 2 1/2 years), we get \$3,188.40.

11. L = 15,000, M = 150, r = 5/600 = 1/120 (this is 5/6 of 1%), so
    when n = 12, \$14,685.87 remains;
    when n = 24, \$14,338.84 remains;
    when n = 60, \$13,064.11 remains;
    when n = 120, \$9,879.00 remains; and
    when n = 200, \$2,226.23 remains.

12. L = 7,000, M = 70, r = 3/400, n = 36, so \$6,279.83 remains. Note this is most of the original loan. It would take over 15 years to pay off this loan completely at the given rate.

13. (a) 216 months (or 18 years). In 215 months the balance (by Formula (4) is \$14,864.73, leaving \$135.27 to pay; allowing for interest over the final month, the last payment is \$136.40.
    (b) 186 months. In 185 months the balance is \$6,962.94, leaving \$37.06 to pay; calculating interest, the final payment is \$37.34.

14. L = 50,000, r = 11/1200, n = 300, so M = \$490.06.

15. M = 250, r = 9.5/1200, n = 360, so L = \$29,731.58 (approximately \$29,750; the monthly payment for that additional \$20 or so would be only 15¢).

16. L = 500, M = 47.50, n = 12, so r = 24.9%. This was obtained by trying various guesses in Formula (4), attempting always to get as close to \$0 as possible. For example, the guess r = 20% leads to

a final balance of -\$15.57, so the interest is higher; the guess $r = 21\%$ produces -\$12.40; we moved closer by about \$3 and need to move \$12 more; so next, guess $r = 25\%$, this produces 31¢, so it's slightly too high; $r = 24.9\%$ produces 3¢, which is plenty close enough!

17. $L = 430$, $M = 41.67$, $n = 12$, and using the trial and error method, $r = 28.8\%$. This dealer is thus charging somewhat more interest on his loans.

18. $L = 500$, $M = 48.45$, $n = 12$, so using Formula (4), $r = 28.8\%$. Since all the figures are proportional, it is reasonable that the interest rate should be the same as in Exercise 17.

19. (a) Using Formula (5), with $M = 150$, $r = 5/600 = 1/120$, and $L = 10{,}000$, we find $n = 98$ months (actually 97.72).

    (b) By trying various values of r in the same formula, with $M = 150$, $L = 9{,}000$, we calculate various corresponding values of n and try to get close to 98. For example, when

    $r = 12$, we find $n = 92$;
    $r = 13$, $n = 97.43$;
    $r = 13.1$, $n = 98.02$; and
    $r = 13.05$, $n = 97.72$.

    Thus, the borrower is paying at the effective annual percentage rate of 13.05% for the use of the \$9,000 he actually received.

20. We want to know what amount \$18,000 is 80% of, so we solve $0.80x = 18{,}000$ and find $x = 18{,}000/0.8 = \$22{,}500$. If this were the amount of the loan, and the monthly payment were \$225 (which is 1% of the face value), then after one year, the balance would be \$22,028.80 (using Formula (4) with $L = 22{,}500$, $M = 225$, $n = 12$, and $r = 1/120$). Note this is higher than the actual loan (\$18,000), which means the monthly payment of \$225 is insufficient even to pay the interest, and thus the principal is actually growing larger rather than smaller. Our formulas all apply, nonetheless. As before, we search for the effective annual percentage rate (r) by trial and error. With $L = 18{,}000$, $M = 225$, $n = 12$, and using Formula (4):

for $r = 15$, balance is \$18,000.00, not high enough, the principal stays level;
$r = 16$, \$18,193.80, way too low;
$r = 20$, \$18,987.25, still a lot low;
$r = 30$, \$21,103.98, still over \$800 low;
$r = 34$, \$22,006.62, very close;
$r = 34.1$, \$22,029.61, virtually exact.

For 2 years, the balloon would be \$21,508.26, and the effective annual percentage rate would be 22.8%.

For 5 years, the balloon would be \$19,596.16, and the APR would be 16.2%. Note then, that the discount has substantially more effect when the loan is repaid quickly. This is reasonable, since it is a one-time charge, and its effect on the interest rate is lessened when it is averaged over longer periods.

21. The borrower gets \$4,950, which is \$5,500 less 10%. The lender puts up \$5,060, which is \$5,500 less 8%. The broker gets the difference, which is 2% or \$110. From Formula (4), with $L = 5{,}500$, $M = 55$, $r = 1/120$, and $n = 24$: the balloon is \$5,257.57.

    For the borrower, we seek an r such that $L = 4{,}950$, $M = 55$, $n = 24$, and the balloon is \$5,257.57. By trial and error, when $r = 16\%$, the balloon is \$5,259.73.

    For the lender, we seek an r such that $L = 5{,}060$, $M = 55$, $n = 24$, and the balloon is \$5,257.57. By trial and error, when $r = 14.75\%$, the balloon is \$5,259.46.

22. By (4), with $L = 11{,}500$, $M = 115$, $r = 8/1200 = 1/150$, and $n = 24$, the current value is \$10,505.91. The sale price agreed upon is 90% of this, or \$9,455.32. By trial and error, we seek r such that with $L = 11{,}500$, $M = 115$, $n = 24$, the balance will be \$9,455.32. When $r = 3.4\%$, the balance is \$9.456.32. Not too good for this investor. But the other one should do correspondingly better.

    The balloon after three years is \$9,946.16, using (4), with $L = 11{,}500$, $M = 115$, $n = 36$, and $r = 1/150$. So the second investor paid \$9,455.32, received 12 monthly payments of \$115, and then got \$9,946.16 back. The principal was growing, so the \$115 per month was insufficient to pay interest. Here, $L = 9{,}455.32$, $M = 115$, $n = 12$; by guessing, $r = 19.4\%$ yields a balloon of \$9,952.30.

23. $5{,}000e^{-5r}$ is \$3,032.65 when r = 10% and it is \$2,744.06 when r = 12%. Thus, in the former case is is better to wait, in the latter to cut now.

24. Using r = 5%, the present value of

(a) is \$100,000;

(b) is \$50,000 plus $\frac{250e^{-5/1200}}{1 - e^{-5/1200}}$ or \$109,800;

(c) is $\frac{500e^{-5/1200}}{1 - e^{-5/1200}} = \frac{500}{(e^{1/240} - 1)} = \$119,760.$

Clearly, (c) is better than (b), and (b) is better than (a).

Using r = 6%, the same calculations give, respectively, \$100,000, \$99,880.29, and \$99,760.57. So here, (c) is worse than (b), and (b) is worse than (a).

25. Using r = 6.75%, the present value of total revenue generated is

$$\int_0^{\infty} 1{,}000{,}000\sqrt{t}\, e^{-rt}\, dt = \frac{1{,}000{,}000}{r\sqrt{r}}\,\Gamma\left(\frac{3}{2}\right)$$

$$= \frac{\sqrt{\pi}\ 500{,}000}{r^{3/2}} = \$50{,}534{,}616.$$

This is about 1% (one-half million dollars) more than it would cost. The project is worth doing under the assumption of this r, but clearly, it's "close" and r should be carefully investigated.

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 1

Request for Help

Student: If you have trouble with a specific part of this unit, please fill out this form and take it to your instructor for assistance. The information you give will help the author to revise the unit.

Your Name____ Unit No.________

Page_______ ○ Upper ○ Middle ○ Lower	OR	Section________ Paragraph________	OR	Model Exam Problem No._____ Text Problem No._____

Description of Difficulty: (Please be specific)

Instructor: Please indicate your resolution of the difficulty in this box.

○ Corrected errors in materials. List corrections here:

○ Gave student better explanation, example, or procedure than in unit. Give brief outline of your addition here:

○ Assisted student in acquiring general learning and problem-solving skills (not using examples from this unit.)

Instructor's Signature____________________________

Please use reverse if necessary.

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 2

Unit Questionnaire

Name________________________________Unit No.______________Date______________

Institution____________________________Course No._____________________________

Check the choice for each question that comes closest to your personal opinion.

1. How useful was the amount of detail in the unit?

____Not enough detail to understand the unit
____Unit would have been clearer with more detail
____Appropriate amount of detail
____Unit was occasionally too detailed, but this was not distracting
____Too much detail; I was often distracted

2. How helpful were the problem answers?

____Sample solutions were too brief; I could not do the intermediate steps
____Sufficient information was given to solve the problems
____Sample solutions were too detailed; I didn't need them

3. Except for fulfilling the prerequisites, how much did you use other sources (for example, instructor, friends, or other books) in order to understand the unit?

____A Lot ____Somewhat ____A Little ____Not at all

4. How long was this unit in comparison to the amount of time you generally spend on a lesson (lecture and homework assignment) in a typical math or science course?

Much	Somewhat	About	Somewhat	Much
____Longer	____Longer	____the Same	____Shorter	____Shorter

5. Were any of the following parts of the unit confusing or distracting? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Paragraph headings
____Examples
____Special Assistance Supplement (if present)
____Other, please explain____

6. Were any of the following parts of the unit particularly helpful? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Examples
____Problems
____Paragraph headings
____Table of Contents
____Special Assistance Supplement (if present)
____Other, please explain____

Please describe anything in the unit that you did not particularly like.

Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)

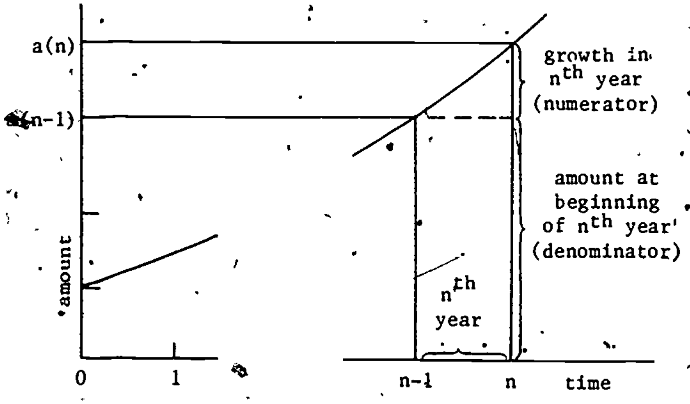
umap

UNIT 382

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT

THE FORCE OF INTEREST

by Michael E. Mays



APPLICATIONS OF CALCULUS TO FINANCE

edc/umap/55chapel st/newton, mass 02160

THE FORCE OF INTEREST

by

Michael E. Mays
Department of Mathematics
West Virginia University
Morgantown, West Virginia 26506

TABLE OF CONTENTS

Intermodular Description Sheet: UMAP Unit 382

Title: THE FORCE OF INTEREST

Author: Michael E. Mays
Department of Mathematics
West Virginia University
Morgantown, WV 26506

Review Stage/Date: III 2/14/80

Classification: APPL CALC/FINANCE

Prerequisite Skills:
1. Be able to differentiate exponential functions.
2. Be able to perform calculations with logarithms.

Output Skills:
1. To appreciate a natural connection between exponential functions and compound interest.
2. To apply the definition of the derivative in a non-science situation.
3. To understand terms used in finance such as simple interest, compound interest, yield rate, nominal rate, and force of interest.

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT (UMAP)

The goal of UMAP is to develop, through a community of users and developers, a system of instructional modules in undergraduate mathematics and its applications which may be used to supplement existing courses and from which complete courses may eventually be built.

The Project is guided by a National Steering Committee of mathematicians, scientists, and educators. UMAP is funded by a grant from the National Science Foundation to Education Development Center, Inc., a publicly supported, nonprofit corporation engaged in educational research in the U.S. and abroad.

PROJECT STAFF

Ross L. Finney	Director
Solomon Garfunkel	Associate Director/Consortium Coordinator
Felicia DeMay	Associate Director
Barbara Kelczewski	Coordinator for Materials Production
Paula M. Santillo	Administrative Assistant
Donna Diduca	Project Secretary
Janet Webber	Technical Typist
Zachary Zevitas	Staff Assistant

NATIONAL STEERING COMMITTEE

W.T. Martin	M.I.T. (Chair)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
Ernest J. Henley	University of Houston
William Hogan	Harvard University
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
R. Duncan Luce	Harvard University
George Miller	Nassau Community College
Frederick Mosteller	Harvard University
Walter E. Sears	University of Michigan Press
George Springer	Indiana University
Arnold A. Strassenburg	SUNY at Stony Brook
Alfred B. Willcox	Mathematical Association of America

The Project would like to thank Norman V. Schmidt of North Iowa Area Community College, Benjamin Greenberg of the College of Staten Island, and Paul C. Rogers of the University of Southern Maine for their reviews, and all others who assisted in the production of this unit.

This material was prepared with the partial support of National Science Foundation Grant No. SED76-19615 A02. Recommendations expressed are those of the author and do not necessarily reflect the views of the NSF, or the National Steering Committee.

1. HOW MONEY GROWS

1.1 Accumulation Functions

Money is worth money. Banks and other financial institutions are willing to pay you for the use of your money, which they in turn lend to others. Corporations use the money you invest in their stock for capital improvements such as new factories or machinery, and pay you dividends for the use of your investment.

A quantity of money invested grows as what you are paid for its use is added to your initial investment. How fast it grows is a measure of how much it is needed by the borrower. We will be concerned with various ways of measuring how fast this growth occurs.

An accumulation function a(t) is a function which gives the amount to which an initial investment of $1 has accumulated at time t (t is usually measured in years.) A change of scale lets you use an accumulation function to determine the value at time t of any initial investment by multiplying by an appropriate constant. Thus if you know that a(3) = 1.15, which means that an initial investment of $1 grows to $1.15 in 3 years, then in 3 years $16 grows to 16($1.15) = $18.40, and $100 grows to $115. The graphs of some reasonable accumulation functions are shown in Figure 1. While the functions are different in some respects, you should notice that these functions are all non-decreasing functions of t whose values are 1 when t = 0. The first three of these will turn up again later.

1.2 Simple Interest

Figure 1a) shows a particularly attractive accumulation function, one whose graph is a straight

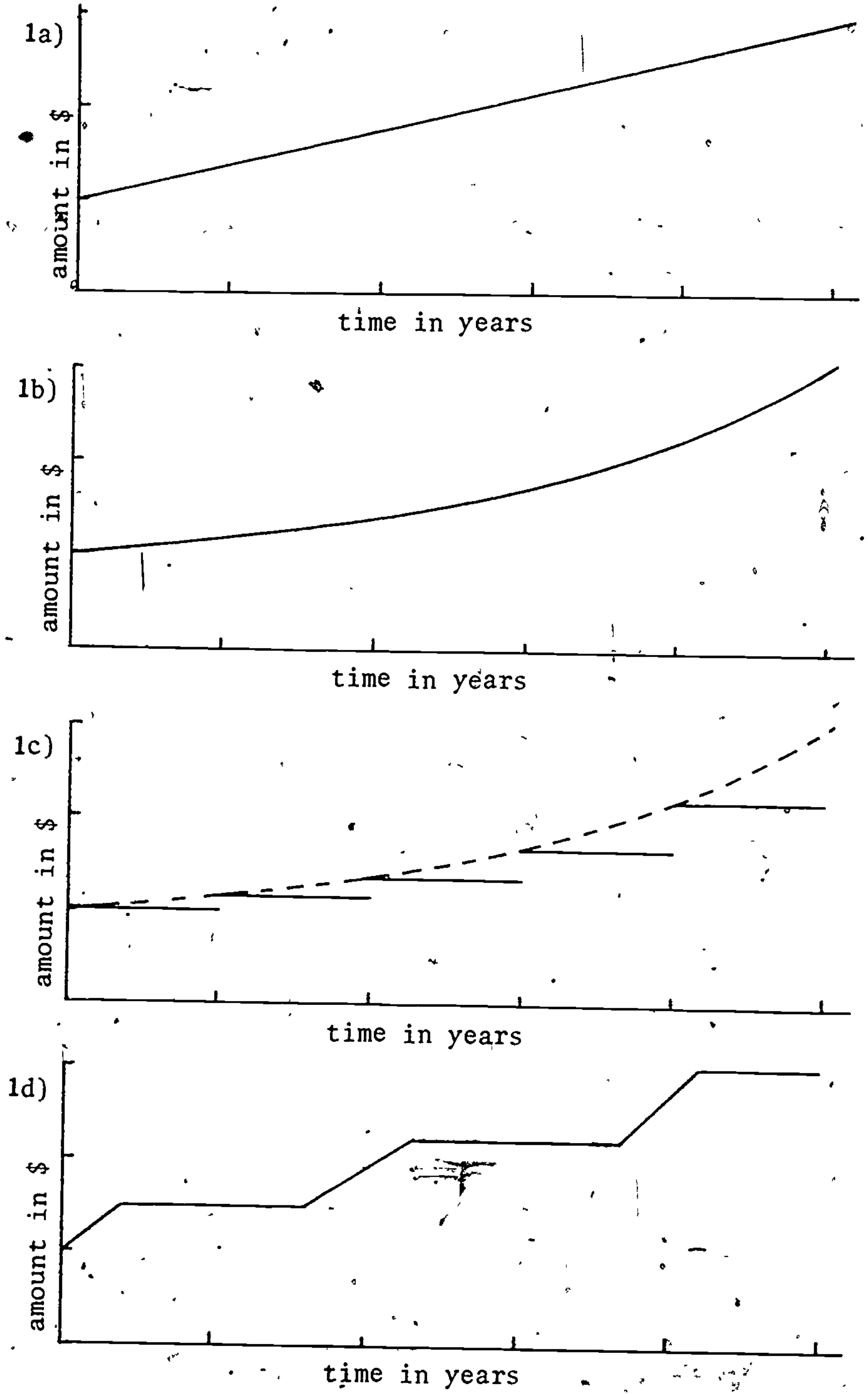


Figure 1. Several possible accumulation functions.

line. You might choose it as an accumulation function by reasoning that, since in the first year an investment of $1 grows by an amount (say) i, the growth in later years should be by an amount of i per year also. You may have seen problems in interest based on the formula

$$I = prt,$$

where I is the amount of interest paid, p is the principal (the amount originally invested), r is the interest rate, and t is the time the money was invested. I computed by this method is called the <u>amount of simple interest</u> earned.

<u>Exercise 1.</u> An investment of $100 earned $27 in interest over a period of 3 years. At what rate of simple interest was the investment made?

The graph of the accumulation function for simple interest is the straight line which passes through the points (0,1) and (1,1+i). For any time t, we can calculate the amount to which an original investment of $1 has accumulated from the formula

$$a(t) = 1 + it.$$

Once we choose the straight line in Figure 1a) as the graph of the accumulation function for simple interest, we can calculate values for a(t) even when t is not an integer. The expression 1 + it is defined for t any real number. Thus, for i = .06, a(1/2) = 1.03 is how much 1 is worth in six months, and a(74/365) = 1 + 74(.06)/365. 1.012 is what 1 invested on January 1 has accumulated to by the Ides of March.

<u>Exercise 2.</u> It takes 8 months for an investment to earn $40 at a rate of simple interest of 6% per year. How long will it take the investment to earn $100?

<u>1.3 Compound Interest</u>

Interest is said to be <u>compounded</u> when it is reinvested to begin earning interest on itself. Simple interest has the fault that even though there is more money in the fund at later times than at time t = 0, interest is paid only on the initial investment of 1. Compound interest is used to calculate the value of the fund at time t on the basis of the value of the fund at time t − 1. Thus if in the first year 1 grows to 1 + i, then in the second year not only does the 1 grow again by a factor of 1 + i, but the i does too. A hundred dollars at 6% interest earns $6 in the first year. If that $6 is compounded, then in the second year the fund grows to $106(1.06) = $112.36. Using the accumulation function for simple interest would yield a(2) = 1 + 2(.06) = 1.12, so $100 would grow to only $112. The extra 36¢ comes from compounding.

The accumulation function for compound interest is given by

$$a(t) = (1 + i)^t.$$

Some banks only credit interest to an account periodically, and an accumulation function reflecting such a policy is graphed in 1c). The advantages of working with a continuous (and differentiable) function are so great, though, that the function in Figure 1b) is more often used in mathematical treatments of compound interest. A bank using the accumulation function graphed in 1b) could advertise 'interest paid from day of deposit to day of withdrawal.'

Exercise 3. How long does it take an investment of $100 to double at a rate of simple interest of 5%? How long does it take to double at 5% if interest is compounded evey year?

---

2. MEASURING INTEREST

2.1 Effective Rates of Interest

A rate of interest measures the growth of money in a fund. One way to measure is given by the effective rate of interest, which gives the rate of growth over a particular year per unit invested at the beginning of the year. If $30 grows to $40 in a year, for example, the interest earned was $40 - $30 = $10, and the rate of growth for that year was 10/30 = .33, or 33%. During the n^{th} year, an initial investment of 1 grows from $a(n - 1)$ to $a(n)$, so denoting the effective rate of interest during the n^{th} year by i_n, we have

$$(4) \qquad i_n = \frac{a(n) - a(n - 1)}{a(n - 1)} .$$

For example, if $93 at the beginning of the fifth year (i.e. at time t = 4) grows to $98 at the end of that year, the effective rate of interest earned during the fifth year, i_5, is given by

$$(5) \qquad i_5 = \frac{a(5) - a(4)}{a(4)} = \frac{98-93}{93} = .0538, \text{ or } 5.38\%.$$

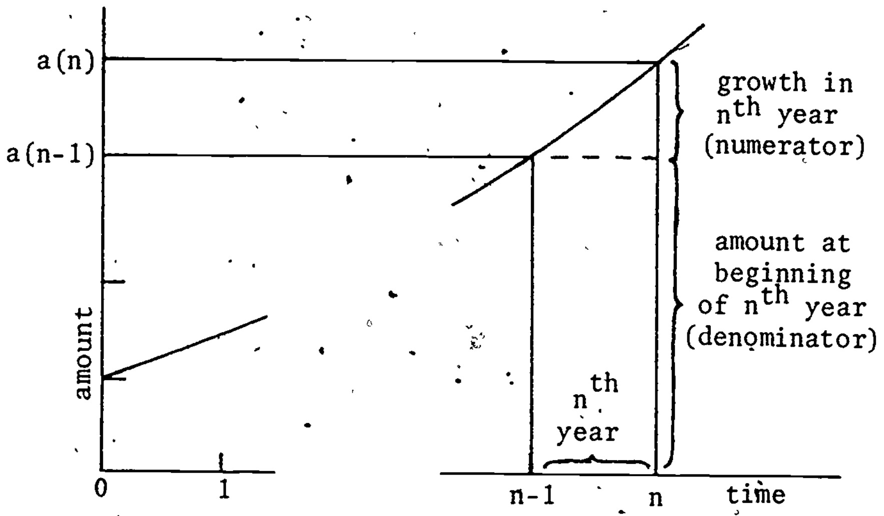


Figure 2. Sorting out i_n.

Exercise 4. During the first year a fund grows from $1000 to $1060. By the end of the second year it has grown to $1121. In which year is the effective rate of interest earned greater?

If we know a formula for $a(n)$ then we can find i_n as a function of n. For instance, if $a(n)$ is the accumulation function for simple interest, then

$$(6) \qquad i_n = \frac{a(n) - a(n - 1)}{a(n - 1)} = \frac{1 + in - (1 + i(n - 1))}{1 + i(n - 1)} = \frac{i}{1 + i(n - 1)}$$

Since this last expression gets smaller as n gets larger, simple interest yields an ever decreasing effective rate of interest. This fact jibes with the criticism of simple interest mentioned in Section 1.3.

On the other hand, for the accumulation function $a(t) = (1 + i)^t$ we have

$$i_n = \frac{a(n) - a(n - 1)}{a(n - 1)}$$

$$= \frac{(1 + i)^n - (1 + i)^{n-1}}{(1 + i)^{n-1}}$$

$$(7) \qquad = \frac{(1 + i)(1 + i)^{n-1} - (1 + i)^{n-1}}{(1 + i)^{n-1}}$$

$$= \frac{(1 + i)^{n-1}(1 + i - 1)}{(1 + i)^{n-1}}$$

$$= 1 + i - 1 = i.$$

Hence compound interest gives a constant effective rate of interest, always equal to the rate of interest in the first year. This is one reason for viewing compound interest as the fairest or most natural way to compute interest.

Exercise 5. Show that at an effective rate of interest of 4%, money left on deposit will double in 17.7 years. How long does it take money to double at an effective rate of 6% ? 8% ?

2.2 Nominal Rates of Interest

How often interest is compounded, or added to the account to begin earning interest on its own, can influence the rate of growth of an investment. The accumulation function for compound interest was built on the assumption that interest was compounded at the end of every year, but it is possible to build accumulation functions based on other assumptions. If interest on an investment of 1 were to be compounded every six months at a rate of 2.5%, then after six months .025 is deposited and after the second six months this amount grows to

$$(1.025).025 + 1.025 = (1.025)^2 = 1.050625.$$

If 1 were compounded annually at 5%, it would grow to only 1.05. Those extra four decimal places on the end of 1.050625 are there because the interest deposited at the end of six months earned interest itself the second half year. We keep track of how often the interest is compounded by saying that the money earns interest at a nominal rate of 5% compounded semiannually. The calculation above shows that a nominal rate of 5% compounded semiannually is equivalent to an effective rate of interest of 5.0625%. Alternately, a nominal rate of 4.939% compounded semiannually is equivalent to an effective rate of 5% because $(1 + .04939/2)^2 = 1.05$.

It seems reasonable that the more often interest is compounded, the faster is the rate of growth of money in a fund. The following table shows this to be true, but we will see that the rate of growth does not increase without bound.

TABLE 1

Effect of More Frequent Compounding on a Nominal Rate of 5%

Number of times per year Interest is compounded:	Corresponding effective rate: $(1 + .05/n)^n - 1$
1 (annually)	5%
2 (semiannually)	5.0625%
4 (quarterly)	5.0945%
12 (monthly)	5.1162%
52 (weekly)	5.1246%
365 (daily)	5.1267%
21900 (every minute)	5.1271%

The number in the last column is $(1 + .05/n)^n - 1$, expressed as a percentage, where n is the number of times per year that interest is compounded (called interest conversion periods) given in the first column. The largest value to hope for in the last column would be

$$(8) \quad \lim_{n\to\infty} (1 + .05/n)^n$$

This would correspond to a fund in which interest is compounded continuously, so that each instant the interest earned begins earning interest on its own.

Exercise 6. What effective rate of interest corresponds to a nominal rate of 8% compounded quarterly?

Exercise 7. Suppose that $100 was invested at a nominal rate of 4% compounded quarterly for a period of 18 months. How much interest was earned?

Another way to describe the growth of money is to compute nominal interest rates that are necessary to yield an effective rate of interest of 5% ate iear. If interest is to be compounded k times a year, at a nominal rate i, then we want

$$(1 + i/k)^k = 1.05.$$

This gives

$$1 + i/k = (1.05)^{1/k},$$

so

$$(9) \qquad i = k((1.05)^{1/k} - 1).$$

Table 2 shows values of i (as percentages) for selected values of k.

TABLE 2

Effect of More Frequent Compounding to Yield an Effective Rate of 5%, Computed from Eq. (9)

Number of times per year interest is compounded (k)	Nominal rate i required to yield 5% effective
1	5%
2	4.9390%
4	4.9089%
12	4.8889%
52	4.8813%
365	4.8793%
21900	4.8790%

You might suspect from the values shown in Table 2 that the more often interest is compounded, the smaller the nominal rate must be to achieve a given effective rate. The smallest value to hope for here is

$$(10) \quad \lim_{n\to\infty} n((1.05)^{1/n} - 1).$$

Exercise 8. What expression would have to be evaluated to compute the value of i in Table 2 corresponding to interest computed every second ?

2.3 Notation

A notation to handle nominal rates of interest must include the number of times interest is reinvested per year. The standard way of writing a nominal interest rate of i compounded n times a year is $i^{(n)}$ (read as i upper n). This rate specifies an effective rate of interest of i/n compounded every n^{th} fraction of a year. Hence $i^{(2)}$ is a nominal rate of i compounded semiannually and $i^{(12)}$ is a nominal rate of i compounded monthly. The last column of Table 1 then gives, for instance, the effective rate of interest corresponding to $.05^{(4)}$ is 5.0945%. Money in a fund earning 5.0945% interest compounded once a year grows just as fast as money invested at a nominal rate of 5% compounded quarterly.

Exercise 9. Use Table 1 to find $.05^{(12)}$.

Exercise 10. Friendly Harry's Loan Shoppe offers unsecured loans of up to $500 with interest payments of 5% per month (these payments are called 'vigorish' in the trade). Write the nominal annual rate Harry earns. What effective rate does he earn on these investments?

---

3. THE FORCE OF INTEREST

3.1 How to Use the Derivative

The derivative of a function at a point has a natural interpretation as the rate of change of the function at that point, so it would be nice to harness the derivative of the accumulation function to describe the rate of growth of a monetary fund due to interest accruing. The derivative of the accumulation function alone, however, is inappropriate because it is influenced by the value of a(t), whereas interest rates quoted before were concerned not with absolute growth but with growth relative to the amount invested. A fund with $200 in it earns interest twice as fast (it earns twice as much in interest in a given time) as a fund with $100 in it, yet the rate of interest is the same for each.

In Equation (4) the absolute growth of the fund is given by a(n) − a(n − 1), but i_n, the effective rate of interest, is that difference divided by a(n − 1). We saw in Section 2.1 that if the accumulation function is the one for simple interest then there is a decreasing effective rate of interest. But when a(t) = 1 + it, da/dt = i, a constant. We need to take into account the value of a(t) as well to get the true picture of the growth of money.

3.2 A Rate of Change per Unit Invested

What matters in the computation of an effective rate of interest i_n is not only the growth during the n^{th} year but also the amount that the n^{th} year started with. To measure the instantaneous rate of change of a(t) we will use the derivative of a(t), but the instantaneous rate of change <u>per unit invested</u> is the measure of the rate of change associated with interest. With this in mind, we define <u>the force of interest function δ(t)</u> associated with a particular accumulation function a(t) by

$$\delta(t) = a'(t)/a(t).$$

The function δ(t) gives the <u>relative change</u> in a(t).

For a given t, δ(t) will usually be given as a decimal which can be expressed as a percent to measure how fast the fund is growing at that time. A useful observation is that δ(t) = d(ln(a(t)))/dt. To see why this is true, notice first that a(t) is 1 when t is 0 and nondecreasing thereafter, so a(t) > 0 and there is no danger that ln a(t) will be undefined. An application of the Chain Rule gives

$$\delta(t) = d(\ln(a(t)))/dt = \frac{1}{a(t)}\frac{d}{dt}a(t). \tag{12}$$

For example, if

$$a(t) = (1 + i)^t$$

then

$$\begin{aligned} \delta(t) &= d(\ln(1 + i)^t)/dt \\ &= d(t \ln(1 + i))/dt \qquad (13) \\ &= \ln(1 + i)\, dt/dt \\ &= \ln(1 + i). \end{aligned}$$

The constant $\ln(1 + i)$ is the force of interest at any time t if we assume a(t) is the accumulation function for compound interest.

<u>Exercise 11.</u> Show $a(t) = 1 + t^2$ is an increasing function for $t \geq 0$ (and therefore an allowable accumulation function). Calculate $\delta(t)$, and find $\delta(4)$.

<u>Exercise 12.</u> Find $\delta(t)$ for a(t) the accumulation function for simple interest. How does $\delta(t)$ behave as t increases?

3.3 Relation to Other Measurements of Interest

We have given the most emphasis so far to compound interest, and shall continue to do so. The function $a(t) = (1 + i)^t$ can also be represented in terms of nominal interest rates. If $i^{(2)}$ is the nominal rate compounded semiannually which yields the effective rate i, then

$$(1 + i^{(2)}/2)^2 = 1 + i,$$

and so

$$(1 + i^{(2)}/2)^{2t} = (1 + i)^t.$$

Thus

$$(1 + i^{(2)}/2)^{2t}$$

gives the same accumulation function as

$$(1 + i)^t = a(t).$$

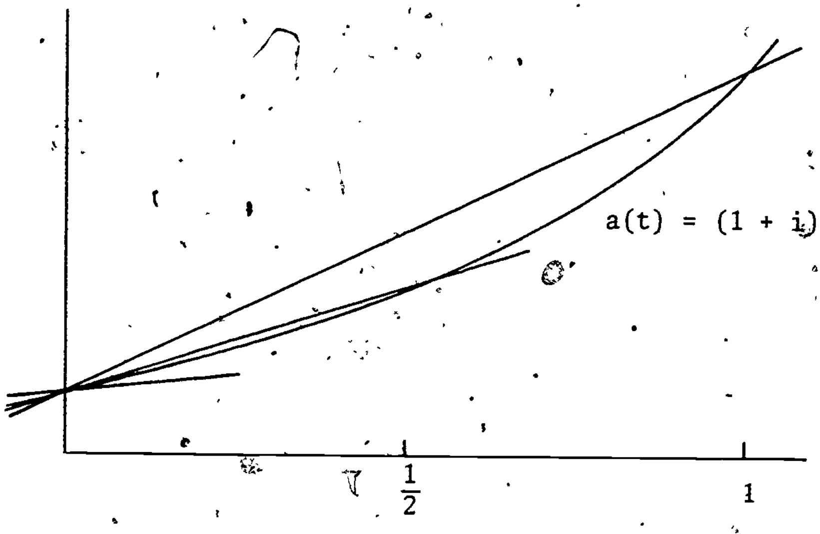


Figure 3. Secants and a tangent: Average rates of growth and an instantaneous rate.

The average rate of growth of the function $a(t) = (1 + i)^t$ over the interval [0,1] is given by the difference quotient.

$$\frac{(1 + i)^1 - (1 + i)^0}{1 - 0} = i. \qquad (14)$$

This means that the average rate of growth over the first year is i. Equivalently, i is the slope of the secant line between (0,1) and (1,1 + i). The average rate of growth over the interval [0,1/2] is

$$\frac{(1 + i)^{1/2} - (1 + i)^0}{1/2 - 0} = 2((1 + i)^{1/2} - 1) = i^{(2)}. \qquad (15)$$

To verify that the last equality holds, solve the equation $(1 + i^{(2)}/2)^2 = 1 + i$ for $i^{(2)}$.

The derivative is defined as the limit of a difference quotient as the width of the interval approaches zero. For the portion of a year from time 0 to time 1/n, the difference quotient is

$$(16) \quad \frac{(1 + i)^{1/n} - (1 + i)^0}{1/n - 0} = n((1 + i)^{1/n} - 1) = i^{(n)},$$

The width of the interval here is 1/n, and we can make the width approach 0 by letting n get larger. Then the instantaneous rate of change of the accumulation function $a(t) = (1 + i)^t$ at $t = 0$ is

$$a'(0) = \lim_{n\to\infty} \frac{(1 + i)^{1/n} - (1 + i)^0}{1/n - 0}$$

$$= \lim_{n\to\infty} i^{(n)}.$$

The limit in this equation is the same as the limit that occurred in Section 2.2 for i = 5%, but now we can evaluate it as $d(a(t))/dt$ at $t = 0$ for $a(t) = (1 + i)^t$. It is

$$d/dt\ e^{t \ln 1.05} = e^{t \ln 1.05}(\ln 1.05),$$

which is just ln 1.05 when t = 0. You can verify that ln 1.05 = 0.048790 to 6 decimal places to see that the last entry in Table 2 hits pretty close. The decreasing numbers in the second column of Table 2 correspond to the (exageratedly) decreasing slopes as the width of the subinterval gets smaller in Figure 3.

Exercise 13. Find $\lim_{n\to\infty} (0.06)^{(n)}$.

3.4 Continuous Compounding and Yield Rates

Banks are limited by law as to the largest interest rate they can pay on savings accounts. Such laws are intended to keep banks from offering fiscally unsound rates in the spirit of competition. A nominal rate so specified, however, can be compounded by the bank as often as it likes. The effect of compounding more often is to raise the effective rate of interest, as illustrated in Table 1 with a nominal rate of 5%. The highest effective rate a bank can pay is found by taking the nominal rate quoted to be δ(t). Thus if the nominal rate quoted is 5%, the effective rate i can be found by solving the equation

$$\ln(1 + i) = .05.$$

The equation is solved by exponentiating both sides to get

$$e^{\ln(1 + i)} = e^{.05},$$

so that

$$1 + i = e^{.05}$$

and

$$i = e^{.05} - 1 = .051271.$$

The value of i found in this manner is often called the yield rate, so that a typical bank advertisement might say, 'Your savings earn interest at an annual rate of 5%, compounded continuously to give you a whopping 5.13% annual yield.' In practice, since there is little difference ~~in~~ the rate whether a bank compounds daily or continuously (see the last entry in Table 1), more often the yield rate quoted is based on a daily compounding of interest to spare the bank from having to explain calculus to its customers.

Exercise 14. What yield rate is associated with a nominal rate of 8%?

Exercise 15. What nominal rate is associated with a yield rate of 8%?

MODEL EXAM

1. Why can't accumulation functions be used to model the behavior of common stocks?
2. Which of the following functions are suitable to be accumulation functions?

   a) $t^2 + 1$ b) $t^2 + t$ c) $1 - t^2$ d) 1
3. Suppose you know $\delta(t)$ has the value 0.045 for any t. Describe $a(t)$.
4. What effective rate of interest corresponds to a nominal rate of 12% compounded monthly?
5. Rank the following effective interest rates in ascending order: 0.05, $0.05^{(2)}$, $0.05^{(4)}$, $\ln(1.05)$.
6. When does $\delta(t) = a'(t)$?
7. A bank pays 8% compounded continuously. What rate compounded yearly must another bank pay for deposits to grow as fast as at the first bank?

ANSWERS TO EXERCISES

1. We are given I = 27, p = 100, and t = 3. Substitution into I = prt gives r = .09, or 9%.

2. Eight months is 2/3 of a year. Using t = 2/3, I = 40, and r = .06 in I = prt gives p = 1000. Now we want t given p = 1000, r = .06, and I = 100, so plugging this information into I = prt gives t = 5/3 years, or 20 months.

3. $100 grows to $200 in the same amount of time it takes $1 to grow to $2. Solve a(t) = 1 + it = 1 + .05 t = 2 for t to find t = 20 years, for simple interest. For compound interest we must solve $a(t) = (1 + i)^t = (1.05)^t = 2$. Taking the natural logarithm of both sides gives $\ln(1.05)^t = \ln 2$, so t = ln 2/ln 1.05 = 14.2 years.

4. i_1 = (a(1) - a(0))/a(0) = (1060 - 1000)/1000 = .06, and i_2 = (a(2) - a(1))/a(1) = (1121 - 1060/1060 = .0575. The effective rate is greater the first year.

5. Solve $1.04^t = 2$ by taking the natural logarithm of both sides to get t = ln 2/ln 1.04 = 17.7. When i = .06, t = ln 2/ln 1.06 = 11.9 and when i = .08, t = ln 2/ln 1.08 = 9.0. A handy rule of thumb called <u>the Rule of 72</u> is that the time it takes a given amount of money to double at rate i is aproximately 72 divided by the interest rate expressed as percent. Money invested at 12% will double in about 72/12 = 6 years.

6. 8% compounded quarterly is 2% every quarter. During one year (four quarters) 1 grows to $(1.02)^4$ = 1.0824, so the effective rate is .0824, or 8.24%.

7. 100 grew to $(1.01)^6$ 100 = 106.51. The interest earned is $6.51.

8. There are 21900x60 = 1,314,000 seconds in a year. We would have to compute

$$1314000((1.05)^{1/1314000} - 1).$$

9. $.05^{(12)}$ = .051162, or 5.1162%.

10. The nominal annual rate he earns is 12(5%) = 60%. If Harry's customers don't pay on time, he can charge them interest on the unpaid interest to reap an effective annual rate of $1.05^{12} - 1 = 0.796$, or almost 80%. That is why he drives an El Dorado.

11. $a(0) = 1 + 0^2 = 1$, and $a'(t) = 2t \geq 0$ for $t \geq 0$ says that a(t) is an increasing function for $t \geq 0$. $\delta(t) = 2t/1 + t^2$ and $\delta(4) = 8/17$.

12. If a(t) = 1 + it, then a'(t) = i and $\delta(t) = i/1 + it$. This expression decreases as t increases, just as i_n decreases for simple interest.

13. ln 1.06 = 0.0583

14. Solve ln(1 + i) = .08 to get i = .0833, or 8.33%.

15. Solve ln(1.08) = i to get i = 0.0770, or 7.7%.

SOLUTIONS TO MODEL EXAM

1. Unfortunately, common stocks can not be relied on not to decrease in value.

2. (a) and (d) are. (b) is not suitable because $a(0) \neq 1$. (c) is not suitable because it is decreasing for $t > 0$.

3. δ(t) was constant, and equal to ln(1 + i), when a(t) was an accumulation function for compound interest. With δ(t) = .045, $a(t) = e^{.045t}$.

4. $1.01^{12} = 1.1268$, so $i = 12.68\%$.

5. $\ln(1.05) < .05 < .05^{(2)} < .05^{(4)}$.

6. $a'(t)/a(t) = a'(t)$ if $a(t) = 1$. This happens when $t = 0$.

7. The effective rate paid by the first bank is $e^{.08} - 1 = 0.0833$, or 8.33%. The second bank must pay 8.33% compounded annually.

STUDENT FORM 1

Request for Help

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

Student: If you have trouble with a specific part of this unit, please fill out this form and take it to your instructor for assistance. The information you give will help the author to revise the unit.

Your Name____________________ Unit No.________

Page_____ ○ Upper ○ Middle ○ Lower	OR	Section_____ Paragraph_____	OR	Model Exam. Problem No._____ Text Problem No._____

Description of Difficulty: (Please be specific)

Instructor: Please indicate your resolution of the difficulty in this box.

○ Corrected errors in materials. List corrections here:

○ Gave student better explanation, example, or procedure than in unit. Give brief outline of your addition here:

○ Assisted student in acquiring general learning and problem-solving skills (not using examples from this unit.)

Instructor's Signature____________________

Please use reverse if necessary.

Return to:
EDC/UMAP
55 Chapel St.
Newton, MA 02160

STUDENT FORM 2

Unit Questionnaire

Name________________________________ Unit No.____________ Date____________

Institution____________________________ Course No.____________________

Check the choice for each question that comes closest to your personal opinion.

1. How useful was the amount of detail in the unit?

____Not enough detail to understand the unit
____Unit would have been clearer with more detail
____Appropriate amount of detail
____Unit was occasionally too detailed, but this was not distracting
____Too much detail; I was often distracted

2. How helpful were the problem answers?

____Sample solutions were too brief; I could not do the intermediate steps
____Sufficient information was given to solve the problems
____Sample solutions were too detailed; I didn't need them

3. Except for fulfilling the prerequisites, how much did you use other sources (for example, instructor, friends, or other books) in order to understand the unit?

____A Lot ____Somewhat ____A Little ____Not at all

4. How long was this unit in comparison to the amount of time you generally spend on a lesson (lecture and homework assignment) in a typical math or science course?

____Much Longer ____Somewhat Longer ____About the Same ____Somewhat Shorter ____Much Shorter

5. Were any of the following parts of the unit confusing or distracting? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Paragraph headings
____Examples
____Special Assistance Supplement (if present)
____Other, please explain________________________________

6. Were any of the following parts of the unit particularly helpful? (Check as many as apply.)

____Prerequisites
____Statement of skills and concepts (objectives)
____Examples
____Problems
____Paragraph headings
____Table of Contents
____Special Assistance Supplement (if present)
____Other, please explain________________________________

Please describe anything in the unit that you did not particularly like.

Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)