DOCUMENT RESUME

ED 214 787 SE 036 478

AUTHOR Schoenfield, Alan H.; And Others
TITLE UMAP Modules-Units 203-211, 215-216, 231-232.
INSTITUTION Education Development Center, Inc., Newton, Mass.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 80
GRANT SED76-19615; SED76-19615-A02
NOTE 382p.; Contains occasional light type.

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Biology; Calculus; Chemistry; *College Mathematics; Economics; Engineering; Geography; Harvesting; Higher Education; *Learning Modules; *Mathematical Applications; Mathematical Models; Medicine; Physics; *Problem Solving; Social Sciences; Supplementary Reading Materials; Topology
IDENTIFIERS *Integration (Mathematics); *Linear Algebra

ABSTRACT

One module is presented in units 203, 204, and 205, as a guide for students, and presents a general strategy for solving integrals effectively. With this material is a solutions manual to exercises. This document set also includes a unit featuring applications of calculus to geography: 206-Mercator's World Map and the Calculus. Unit 207-Management of A Buffalo Herd, features a Leslie-type model covering applications of linear algebra to harvesting. Two units include applications of linear algebra to economics: 208-Economic Equilibrium-Simple Linear Models, and 209-General Equilibrium-A Leontief Economic Model. Unit 210-Vicous Fluid Flow and the Integral Calculus, contains applications of calculus to engineering. Module 211-The Human Cough, views calculus applications to physics, biological, and medical sciences. Social science applications of calculus are viewed in 215-Zipf's Law and His Efforts to Use Infinite Series in Linguistics. Unit 216-Curves and their Parametrization, and 231-The Alexander Horned Sphere, focus on introductory topology. Finally, 232-Kinetics of Single Reactant Reactions, views calculus applications to chemistry. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214787

umap Units 203, 204, 205

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

Alan H. Schoenfeld

INTEGRATION: Getting It All Together

A guide for students, presenting a general strategy for solving integrals effectively

June 1977

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

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it

Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
P. BABB

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SE 036 478

Intermodular Description Sheet: Unit 203, 204, 205

Title: INTEGRATION: GETTING IT ALL TOGETHER

Author/Correspondent: Alan H. Schoenfeld
Group in Science and Mathematics Education
c/o Physics Department
University of California
Berkeley, CA 94720

Review Stage/Date: IV, 6/30/77

Classification: Calculus, Indefinite Integration, Review

Use Experience: Second quarter calculus class at University of California, Berkeley. See paper by A. Schoenfeld, "Presenting a Strategy for Indefinite Integration," Sesame, University of California, for details.

Length: 7 hours

Suggested Support Material: None

Prerequisite Skills: Be able to solve problems in indefinite integration knowing that a specific technique (i.e., substitution, partial fractions, integration by parts, trigonometric substitution) is applicable. Recall and use the essential formulas given in the table on the inside of the back cover.

Output Skills: Given an indefinite integral problem solvable by one of the above techniques, find a technique which is appropriate and solve it. Specifically,

Simplify integrals by algebraic substitution
Classify integration problems into the appropriate technique
Modify integration problems so that they can be classified and solved by one of these techniques.

Other Related Units: None

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 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 one of many projects of Education Development Center, Inc., a publicly supported, nonprofit corporation engaged in educational research and development in the U.S. and abroad.

PROJECT STAFF

Ross L. Finney	Codirector
William U. Walton	Codirector
Solomon Garfunkel	Consortium Coordinator
Joyce Taggart	Editorial Assistant
Thomas Good	Editorial Assistant

NATIONAL STEERING COMMITTEE

W. T. Martin	MIT (Chairman)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
James D. Forman	Rochester Institute of Technology
Ernest J. Henley	University of Houston
William F. Lucas	Cornell University
Walter E. Sears	University of Michigan Press
Peter Signell	Michigan State University
George Springer	Indiana University
Robert H. Tamarin	Boston University
Alfred B. Willcox	Mathematical Association of America
Nancy J. Kopell	Northeastern University

This material was prepared with the support of National Science Foundation Grant No. SED 76-19615. Recommendations expressed are those of the authors and do not necessarily reflect the views of the NSF, nor of the National Steering Committee.

CONTENTS

INTRODUCTION

In order to solve problems in indefinite integration effectively, students need both a mastery of the special techniques of integration and a general procedure for choosing and applying these techniques to problems. Most textbook space and classroom time in this subject area is devoted to teaching and practicing the special techniques of integration. It is generally assumed that with much practice and the help of insightful comments from their teachers, students will develop a "feel" for the material that enables them to solve problems effectively. In my experience, however, many students have had difficulty learning to approach problems in integration systematically and effectively -- even after lengthy classroom discussions of problem solutions.

To overcome this problem, this booklet provides students directly with a general procedure for approaching and solving problems in integration. Based on observations of "experts" working on integrals, the procedure has three steps: SIMPLIFY, CLASSIFY, and MODIFY.

In step 1, SIMPLIFY, we try to reduce a problem to one which can be solved by a formula or can be done easily. If this fails to solve the problem we proceed to step 2, CLASSIFY. Here we use the form of the integrand to decide which special technique (integration by parts, by partial fractions, etc.) to use on the problem. If we are unable to CLASSIFY the integrand, we go to step 3, MODIFY. There we try to manipulate the integrand into a more familiar or manageable form. We always check for simple alternatives before beginning complicated calculations, and start the process over with step 1 whenever we have succeeded in transforming the integral to something easier. The general procedure is outlined in the table on page 3, and summarized in full detail on the last page of this booklet.

INTEGRATION:

A GENERAL PROCEDURE

Proceed from one step to the next when the techniques of that step fail to solve the problem. Always look for easy alternatives before beginning complicated calculations. If you succeed in transforming the problem to something easier, begin again at Step 1.

Step 1: SIMPLIFY!	
Easy Algebraic Manipulations	Obvious Substitutions

↓↓

Step 2: CLASSIFY!			
Rational Functions	Products	Trigonometric Functions	Special Functions

↓↓

Step 3: MODIFY!		
Problem Similarities	Special Manipulations	Needs Analysis

HOW TO USE THESE MATERIALS

Work the pre-test in Appendix I. These materials are written for people who have mastered the basic techniques of integration. If you miss more than one of the pre-test problems, or if you find them difficult, you should review your textbook's sections on basic antiderivatives and substitutions before you start Chapter 1. Before you work on Chapter 2, you should be familiar with the techniques of partial fractions, integration by parts, and trigonometric substitutions.

This booklet is organized like the General Procedure, given in the chart on page 3. The three chapters in the booklet and the sections they are divided into correspond to the three steps in the general procedure and their subdivisions. You should work through this booklet following the procedure closely, until using it becomes automatic. If it does, you will be able to solve problems in integration like an expert.

Each section begins with a description of some technique of integration, which is summarized in table form. The table is followed by sample problems, which serve as review problems and examples. You should try to solve each sample problem yourself. Then compare your answer with the solution given. Just reading through the solutions will not be enough! You should focus on the process of solution, which is as important as the answer.

Each chapter ends with exercises designed to reinforce the procedures you have just learned. Work the exercises as if they were a test. Detailed solutions are in a separate solutions manual.

Note: It's easy to "lose" terms in an intergal if we're not careful. I've chosen to write all the terms in an integral at each stage of the process, and I suggest you do the same. This takes some extra time, but it helps prevent costly mistakes.

ERIC Full Text Provided by ERIC

SIMPLIFY!

There is one general rule that you should keep in mind whenever you are solving problems:

> ALWAYS CHECK FOR EASY ALTERNATIVES BEFORE BEGINNING ANY COMPLICATED OR TIME-CONSUMING OPERATIONS.

As the sample problems below illustrate, it is worth taking a few moments to look for a quick or easy solution to a problem before jumping into a complicated procedure. This is especially true in integration, where a timely observation can save tremendous amounts of work. The two types of SIMPLIFYing operations we will discuss are summarized below.

Step 1: SIMPLIFY	
Easy Algebraic Manipulations	Obvious Substitutions

EASY ALGEBRAIC MANIPULATIONS

Some algebraic manipulations are easy enough to use that it's worth considering them automatically before going on to anything else. For example, we almost always break the integral of a sum into a sum of integrals and then integrate term by term. Before doing this, however, we should look for other alternatives. Sometimes an algebraic or trigonometric identity will simplify the term facing us, before we try to integrate it. Another operation which is more complicated but also worth considering is simplifying rational functions by long division.

We call a rational function (the quotient of two polynomials) a "proper fraction" if the degree of the numerator is less than the degree of the denominator. Proper fractions are usually easier to manipulate than others. Also, we can only apply the technique of partial fractions to proper fractions. Thus we should consider division as a preliminary simplification. In sum, we have:

EASY ALGEBRAIC MANIPULATIONS
(1) Break integrals into sums
(2) Exploit Identities
(3) Reduce rational functions to Proper Fractions by division

SAMPLE PROBLEMS

Each of the following sample problems can be SIMPLIFIED by an easy algebraic manipulation. Try to solve each problem before you read the solution, and then compare your method with mine.

1. $\int \frac{1 + \sin x}{\cos^2 x}\, dx$ 2. $\int (\sin x + \cos x)^2\, dx$

3. $\int \frac{x^3}{x^2 + 1}\, dx$

SOLUTIONS

1. $\int \frac{1 + \sin x}{\cos^2 x}\, dx.$

This integrand contains a sum, so we should consider breaking the problem into a sum of integrals. This gives us

$$\int \frac{1}{\cos^2 x}\, dx + \int \frac{\sin x}{\cos^2 x}\, dx = \int \sec^2 x\, dx + \int \frac{\sin x}{\cos^2 x}\, dx.$$

The first integral can now be done directly. In the second, we notice that the denominator contains the term cos x. Since the numerator is sin x, which (except for a minus sign) is the derivative of cos x, this suggests that we make the substitutions

$$u = \cos x, \quad du = -\sin x\, dx.$$

Then the integrals become

$$\int \sec^2 x\, dx - \int \frac{-\sin x\, dx}{\cos^2 x} = \tan x - \int \frac{du}{u^2}$$

$$= \tan x - \int u^{-2} du = \tan x - (-u^{-1}) + C = \tan x + \frac{1}{u} + C$$

$$= \tan x + \frac{1}{\cos x} + C = \underline{\tan x + \sec x + C.}$$

2. $\int (\sin x + \cos x)^2\, dx$

The first thing we should notice is that the integral can't be done directly, so some sort of manipulation is called for. If we square the term (sin x + cos x), we obtain

$$\int (\sin^2 x + 2 \sin x \cos x + \cos^2 x)\, dx.$$

STOP! While the integral can be broken into three terms and each done separately, there are simplifications. Do you see them? Recalling the trig identities ($\sin^2 x + \cos^2 x = 1$) and ($\sin 2x = 2 \sin x \cos x$), we can write the above as

$$\int (\sin^2 x + \cos^2 x)\, dx + \int 2 \sin x \cos x\, dx = \int 1\, dx + \int \sin 2x\, dx$$

$$= \underline{x - \frac{1}{2} \cos 2x + C.}$$

NOTE: The term $\int 2\sin x \cos x\, dx$ can also be solved by the substitution u=sin x or u = cos x. These give two equivalent solutions to the problem,

$$\underline{x + \sin^2 x + C} \quad \text{and} \quad \underline{x - \cos^2 x + C.}$$

3. $\int \frac{x^3}{x^2 + 1}\, dx$

The integrand in this problem is an "improper fraction", so we should perform a division. The division gives us a quotient of (x) and a remainder of (-x), so we obtain

$$\int \frac{x^3}{x^2 + 1}\, dx = \int \left(x - \frac{x}{x^2 + 1}\right) dx = \int x\, dx - \int \frac{x}{x^2 + 1}\, dx.$$

In the second integrand, we notice that the numerator is one half the derivative of the denominator. If we make the substitutions $u = (x^2+1)$, $du = (2x\, dx)$, the above becomes

$$\int x\, dx - \frac{1}{2}\int \frac{du}{u} = \frac{1}{2} x^2 - \frac{1}{2} \ln|u| + C$$

$$= \underline{\frac{1}{2} x^2 - \frac{1}{2} \ln|x^2+1| + C.}$$

Section 2

"OBVIOUS" SUBSTITUTIONS

Using substitutions is one of the most powerful tools we have for simplifying and solving integrals. I always look for substitutions before I try more complex procedures. There are two guidelines I use in looking for substitutions:

(1) Does the integrand contain a function of a function? If it does, try a substitution with u as the "inside" function. Consider the integrals

$$\int \frac{x \tan^{-1}(x^2)}{1 + x^4}\,dx \quad \text{and} \quad \int \frac{\sin x}{\cos^2 x}\,dx.$$

The term $\tan^{-1}(x^2)$ appears in the first integral, with x^2 as an inside function. I would try the substitution $u = x^2$ in that problem. The denominator of the second integral is $\cos^2 x = (\cos x)^2$, so $\cos x$ is an inside function. I would try $u = \cos x$.

(2) Does the integrand contain a complicated or "nasty" function, particularly in the denominator of a fraction? If so, try a substitution with u as the "nasty" function. Consider

$$\int (\tan^{-1}x + x)\left(\frac{x^2+2}{x^2+1}\right)dx \quad \text{and} \quad \int \frac{x}{x^2 - 9}\,dx.$$

In the first problem I would try $u = (\tan^{-1}x + x)$, and *hope* that it helps. [It does; see sample problem 2.] In the second problem the denominator isn't *particularly* "nasty", but it's worth *trying* the substitution $u = x^2 - 9$. Then $du = 2x\,dx$, and the integral is

$$\frac{1}{2}\int \frac{2x\,dx}{x^2 - 9} = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\ln|u| + C = \frac{1}{2}\ln|x^2-9| + C.$$

Note: If the problem I just discussed were $\int \frac{1}{x^2-9}\,dx$, the substitution $u = x^2-9$ would not have helped. In general, a substitution $u = f(x)$ will only help if you can find the term $du = f'(x)dx$ somewhere in the integral. If you try a substitution and it looks like you're getting involved in a complicated procedure, stop to consider other alternatives. The procedures of chapter 1 are designed to help SIMPLIFY and solve an integral rapidly. You should explore all simple alternatives before trying anything complicated. If need be, you can always return to a complicated substitution later.

OBVIOUS SUBSTITUTIONS
(1) "Inside" functions (2) "Nasty" terms and denominators

SAMPLE PROBLEMS

Each of problems 1 through 3 can be solved by a substitution. Try to solve each problem before you read the solution, and then compare your method with mine.

1. $\int \frac{e^{\tan^{-1}x}}{1 + x^2}\,dx$

2. $\int (\tan^{-1}x + x)\left(\frac{x^2+2}{x^2+1}\right)dx$

3. $\int \frac{e^x + e^{-x}}{e^x - e^{-x}}\,dx$

4. One of the following two integrals is much easier to solve than the other. Decide which it is, and solve it.

(a) $\int x^3(1+x^4)^5\,dx$ (b) $\int (1+x^4)^5\,dx$

ERIC Full Text Provided by ERIC

SOLUTIONS

1. $\int \frac{e^{\tan^{-1}x}}{1+x^2}\,dx$

In this problem we have the term $e^{\tan^{-1}x}$, so $\tan^{-1}x$ is an "inside" function. If we try

$$u = \tan^{-1}x, \quad \text{then} \quad du = \frac{1}{1+x^2}\,dx.$$

Since du does appear in the integral, we can make the substitution. The integral becomes

$$\int e^{\tan^{-1}x}\left(\frac{1}{1+x^2}\,dx\right) = \int e^u\,du = e^u + C$$

$$= e^{\tan^{-1}x} + C.$$

2. $\int (\tan^{-1}x + x)\left(\frac{x^2+2}{x^2+1}\right)dx$

In this expression the term $(\tan^{-1}x + x)$ is rather "nasty". We might *consider* the substitution

$$u = \tan^{-1}x + x,$$

and see if it helps. We obtain

$$du = \left(\frac{1}{1+x^2} + 1\right)dx = \left(\frac{1}{1+x^2} + \frac{1+x^2}{1+x^2}\right)dx$$

$$= \left(\frac{x^2+2}{1+x^2}\right)dx,$$

and we're in luck. The integral then becomes

$$\int (\tan^{-1}x + x)\left(\frac{x^2+2}{x^2+1}\,dx\right) = \int u\,du = \frac{1}{2}u^2 + C$$

$$= \frac{1}{2}(\tan^{-1}x + x)^2 + C.$$

3. $\int \frac{e^x + e^{-x}}{e^x - e^{-x}}\,dx$

I'd like to work this problem using all the methods of this chapter, to illustrate how I would think about this problem if I didn't know where it came from.

As a first step, I look for algebraic simplifications. The numerator is a sum, so I might consider breaking the integral up into

$$\int \frac{e^x}{e^x - e^{-x}}\,dx + \int \frac{e^{-x}}{e^x - e^{-x}}\,dx.$$

This doesn't seem to help, so I look for substitutions. I might be tempted to try the substitution $u = e^x$ at first, since all the terms in the integral are expressed in terms of e^x. But $du = e^x dx$, and I don't see that in the integral. For that reason I won't explore the substitution further now. If necessary, I can return to it.

Finally, I might try a substitution for the denominator,

$$u = (e^x - e^{-x}).$$

This gives $du = (e^x + e^{-x})dx$,

which does appear in the integral. From here on the problem is easy. We have

$$\int \frac{1}{e^x - e^{-x}}[(e^x + e^{-x})\,dx] = \int \frac{1}{u}\,du = \ln|u| + C$$

$$= \ln\left|\frac{e^x + e^{-x}}{e^x - e^{-x}}\right| + C.$$

4. One of the following two integrals is much easier to solve than the other. Decide which it is, and solve it.

(a) $\int x^3(1+x^4)^5\,dx$ (b) $\int (1+x^4)^5\,dx$

As always, I start working on a problem by looking for algebraic simplifications. In both parts (a) and (b) of this problem, I *can* multiply $(1+x^4)$ by itself five times, and then integrate term by term. That seems too complicated, however, so I look for other alternatives.

In both parts of the problem I see $(1+x^4)^5$, so that the term $(1+x^4)$ is an "inside" function. If I try

$$u = 1 + x^4, \qquad \text{then} \qquad du = 4x^3\,dx.$$

Since the term (x^3dx) appears in part (a), that integral will be easy to solve. It becomes

$$\int x^3(1+x^4)^5\,dx = \frac{1}{4}\int (1+x^4)^5(4x^3dx) = \frac{1}{4}\int u^5\,du = \left(\frac{1}{4}\right)\left(\frac{1}{6}u^6\right) + C$$

$$= \underline{\frac{1}{24}(1+x^4)^6 + C.}$$

* * * <u>WARNING</u> * * *

The sample problems you've worked through in this chapter may have seemed very easy, because you were on guard for simple solutions. On tests I've seen students spend ten or fifteen minutes trying to solve

$$\int \frac{x\,dx}{x^2 - 9}$$

by partial fractions or by using the substitution $x = 3\sin\theta$! The moral of this chapter is:

> *When you start working on a problem, always check for an easy algebraic manipulation or obvious substitution. Only when you're sure the problem cannot be SIMPLIFIED should you try anything else.*

<u>*EXERCISES FOR CHAPTER 1*</u>

Detailed solutions of these exercises are available in a separate solutions manual. The order of the solutions is scrambled, to keep you from accidentally seeing the answer to the next problem you are working on. The solution number of the exercise you are working on is underneath the exercise number. For example,

| 1. sol. 5 | means that solution #5 presents a discussion of exercise 1.

* * * * * * * *

In each of the following exercises, <u>one</u> problem can be done easily. Use the techniques of easy algebraic manipulations and obvious substitutions to determine which it is, and solve it.

1. (sol. 5)
(a) $\int \frac{dx}{2 + \sin x}$
(b) $\int \frac{\cos x\,dx}{2 + \sin x}$

2. (sol. 2)
(a) $\int \frac{x + 1}{x^3 + x^2 + 1}\,dx$
(b) $\int \frac{x^3 + x^2 + 1}{x + 1}\,dx$

3. (sol. 8)
(a) $\int \tan^4 x \sec x\,dx$
(b) $\int \sec^4 x \tan x\,dx$

4. (sol. 1)
(a) $\int \frac{\tan^{-1} x}{x^2 + 1}\,dx$
(b) $\int \tan^{-1} x\,dx$

5. (sol. 3)
(a) $\int \ln(e^x)\,dx$
(b) $\int \ln(x)\,dx$

6. (sol. 4)
(a) $\int \frac{1}{(\sqrt{x})(1 + \sqrt{x})^5}\,dx$
(b) $\int \frac{1}{(1 + \sqrt{x})^5}\,dx$

7. (sol. 6)
(a) $\int \frac{e^x}{e^{5x} - 1}\,dx$
(b) $\int \frac{e^{5x} - 1}{e^x}\,dx$

8. (sol. 7)
(a) $\int \frac{1}{x^2 - 4x + 3}\,dx$
(b) $\int \frac{x - 2}{x^2 - 4x + 3}\,dx$

ERIC Full Text Provided by ERIC

Chapter 2

CLASSIFY!

As we noted in the introduction, experts generally follow a three-step procedure when solving integrals. The first step, which we discussed in Chapter 1, consists in looking for simplifications or easy solutions to a problem. The second step, if necessary, consists of choosing and applying the technique most likely to solve a problem.

This choice of technique is usually based on the *FORM* of the integrand. Ask an expert why he chooses to solve $\int x \sin x \, dx$ using integration by parts, for example, and he'll say "because it's a product of dissimilar functions." The solution to a problem follows routinely once the right technique has been chosen.

In this chapter we will classify integrals in four basic categories, and discuss the techniques most often effective in dealing with them. Our classification is summarized by the second box in the General Procedure:

Step 2: CLASSIFY			
Rational Functions	Products	Trigonometric Functions	Special Functions

Your goal in working through this section should be to classify integrands by form and recall the techniques appropriate to them. *If you systematically use the simplifications of Chapter 1 and the classification scheme of this section, you should be able to solve most of the problems at the end of your text's chapter on integration.*

Section 1

RATIONAL FUNCTIONS

A rational function is the quotient of two polynomials. The procedure for integrating rational functions is straightforward, although it may sometimes be long and involved. A large part of that procedure is purely algebraic, and consists of "breaking up" complicated rational functions into sums of simpler ones. We will begin by examining the simple or "basic" rational functions, and then discuss how to break up the more complicated ones.

Part 1:

BASIC RATIONAL FUNCTIONS

Definition: A *Basic Rational Function* is a "proper fraction" of the form

$$\frac{r}{ax+b}, \quad \frac{r}{(ax+b)^n}, \quad \text{or} \quad \frac{rx+s}{ax^2+bx+c}.$$

Basic rational functions of the first two types are easy to integrate. If the denominator is $(ax+b)$ or $(ax+b)^n$, the substitution $u = (ax+b)$ will solve the problem. See sample problems 1 and 2.

Things are more complicated if the denominator is quadratic. If the denominator factors easily, we use partial fractions to break up the integrand. For example,

$$\int \frac{(x-5)\,dx}{x^2-4x+3} = \int \frac{(x-5)dx}{(x-1)(x-3)} = \int \left(\frac{2}{x-1} - \frac{1}{x-3}\right) dx$$

$$= 2\ln|x-1| - \ln|x-3| + C.$$

We will discuss the technique of partial fractions in part 2 of this section.

Suppose the denominator does not factor easily. Then complete the square and make a substitution for the u term in the denominator. There are two possibilities.

(i) If the denominator is of the form (u^2+a^2), we will obtain something of the form

$$\int \frac{bu+c}{u^2+a^2}\,du = b\int \frac{u}{u^2+a^2}\,du + c\int \frac{1}{u^2+a^2}\,du.$$

The first integral on the right will yield a logarithm, and the second gives an arctangent.

(ii) If the denominator is of the form (u^2-a^2), we obtain

$$\int \frac{bu+c}{u^2-a^2}\,du.$$

There are two ways to continue from here. One is to factor the denominator and use partial fractions to break up the expression

$$\frac{bu+c}{(u+a)(u-a)}.$$

If the factors $(u+a)$ and $(u-a)$ look reasonable, this is probably a good way to finish the problem. We do have another alternative, however.

We can write the integral as

$$b\int \frac{u}{u^2-a^2}\,du + c\int \frac{1}{u^2-a^2}\,du.$$

The first integral is a logarithm, and the second can be solved easily using the formula given below. See sample problems 3 through 5.

$$\frac{1}{u^2-a^2} = \frac{1}{2a}\left(\frac{1}{u-a} - \frac{1}{u+a}\right)$$

INTEGRATING BASIC RATIONAL FUNCTIONS

(1) If the denominator is $(ax+b)$ or $(ax+b)^n$, substitute $u = (ax+b)$. This reduces the problem to standard form.

(2) If the denominator is quadratic and factors easily, use partial fractions to finish the problem.

(3) If the denominator is quadratic and does not factor easily, complete the square. If the denominator is then

i: (u^2+a^2), integrate directly to obtain a logarithm and/or inverse tangent.

ii: (u^2-a^2), either use partial fractions or break up the integral and use the formula on p.17.

<u>Note:</u> *Make sure you have checked for SIMPLIFICATIONS <u>before</u> you use the procedure for rational functions.*

SAMPLE PROBLEMS

The solutions to these problems illustrate the techniques described above. Try to solve them before you read my solutions. If they cause you a great deal of difficulty, you should probably practice on some similar problems from your textbook.

1. $\int \frac{4}{5x+7}\,dx$

2. $\int \frac{5}{(4x+3)^6}\,dx$

3. $\int \frac{3x+7}{x^2+4x+13}\,dx$

4. $\int \frac{x+2}{x^2+4x+13}\,dx$

5. $\int \frac{x+5}{x^2+4x+2}\,dx$

1. $\int \frac{4}{5x+7}\,dx$

There is no algebraic simplification possible. Since the denominator is (5x+7), we make the substitutions

$$u = 5x+7; \quad du = 5\,dx.$$

The integral then becomes

$$\frac{4}{5}\int \frac{5\,dx}{5x+7} = \frac{4}{5}\int \frac{du}{u} = \frac{4}{5}\ln|u| + C = \underline{\frac{4}{5}\ln|5x+7| + C.}$$

2. $\int \frac{5}{(4x+3)^6}\,dx$

Again, I see no algebraic simplification. Since the denominator is $(4x+3)^6$, the substitutions

$$u = 4x+6; \quad du = 4\,dx$$

are called for. The integral then becomes

$$\frac{5}{4}\int \frac{4\,dx}{(4x+3)^6} = \frac{5}{4}\int \frac{du}{u^6} = \frac{5}{4}\int u^{-6}du = \frac{5}{4}\left(\frac{u^{-5}}{-5}\right) + C$$

$$= -\frac{1}{4}\left(\frac{1}{u^5}\right) + C = \underline{\frac{-1}{4(4x+3)^5} + C.}$$

3. $\int \frac{3x+7}{x^2+4x+13}\,dx$

As a preliminary algebraic manipulation I would consider breaking the integral into a sum, but that doesn't look like it will help yet. Checking for obvious substitutions, I would *consider* substituting for the denominator, $u = x^2+4x+13$. This gives $du = (2x+4)dx$, which does not appear in the numerator. I can't factor the denominator, so I should complete the square. Since

$$x^2 + 4x + 13 = (x^2 + 4x + 4) + 9 = (x+2)^2 + (3)^2,$$

ERIC Full Text Provided by ERIC

the denominator is of the form $(u^2 + a^2)$, where $u = (x+2)$ and $a = 3$. Making the substitutions $u=x+2$ and $du=dx$, we obtain

$$\int \frac{(3x+7)\,dx}{x^2+4x+13} = \int \frac{[3(u-2)+7]du}{u^2+3^2} = \int \frac{3u\,du}{u^2+3^2} + \int \frac{du}{u^2+3^2}$$

$$= \frac{3}{2}\ln|u^2+3^2| + \frac{1}{3}\tan^{-1}\left(\frac{u}{3}\right) + C$$

$$= \underline{\frac{3}{2}\ln|x^2+4x+13| + \frac{1}{3}\tan^{-1}\left(\frac{x+2}{3}\right) + C.}$$

4. $\int \frac{x+2}{x^2+4x+13}\,dx$

As always, I begin work on this problem by looking for easy algebraic manipulations. The integral can be broken into a sum of two integrals, but this does not look especially promising. I see no useful identities and this is already a "proper fraction", so I look for obvious substitutions next.

The "nasty" term is the denominator, so I should *consider* the substitution

$$u = x^2 + 4x + 13.$$

This would give

$$du = (2x + 4)\,dx,$$

which is double the numerator in this problem! The rest is easy. The integral is

$$\frac{1}{2}\int \frac{(2x+4)\,dx}{x^2+4x+13} = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\ln|u| + C$$

$$= \underline{\frac{1}{2}\ln|x^2+4x+13| + C.}$$

Notice: This problem could have been done by completing the square in the denominator, like we did in problem 3. The advantage of the SIMPLIFY step is that it saved us the trouble.

5. $\int \frac{x + 5}{x^2 + 4x + 2}\, dx$

A preliminary check indicates that none of the SIMPLIFYING procedures will be of assistance here. Since I cannot factor the denominator easily, I complete the square to obtain

$$x^2 + 4x + 2 = (x^2 + 4x + 4) - 2 = (x + 2)^2 - (\sqrt{2})^2.$$

Thus the denominator is of the form u^2-a^2, where $u = x+2$ and $a = \sqrt{2}$. Making the substitutions $u = (x+2)$ and $du = dx$, we obtain

$$\int \frac{(x+5)\,dx}{x^2+4x+2} = \int \frac{[(u-2)+5]du}{u^2 - 2} = \int \frac{u\,du}{u^2 - 2} + \int \frac{3\,du}{u^2 - 2}.$$

The first integral is easy, and yields a logarithm. For the second integral we can use the formula on page 17 to obtain

$$\frac{1}{u^2 - 2} = \frac{1}{2\sqrt{2}}\left(\frac{1}{u - \sqrt{2}} - \frac{1}{u + \sqrt{2}}\right),$$

and the integral becomes

$$\int \frac{u\,du}{u^2 - 2} + \frac{3}{2\sqrt{2}}\int\left(\frac{1}{u - \sqrt{2}} - \frac{1}{u + \sqrt{2}}\right) du =$$

$$\frac{1}{2}\ln|u^2-2| + \frac{3}{2\sqrt{2}}\left(\ln|u - \sqrt{2}| - \ln|u + \sqrt{2}|\right) + C =$$

$$\frac{1}{2}\ln|u^2-2| + \frac{3}{2\sqrt{2}}\ln\left|\frac{u - \sqrt{2}}{u + \sqrt{2}}\right| + C =$$

$$\frac{1}{2}\ln|x^2+4x+2| + \frac{3}{2\sqrt{2}}\ln\left|\frac{(x+2) - \sqrt{2}}{(x+2) + \sqrt{2}}\right| + C.$$

Part 2:

DECOMPOSING RATIONAL FUNCTIONS

In part 1 of this section we learned to integrate the basic rational functions. It is a fact that *any* rational function can be decomposed into a sum of basic rational functions. The techniqes we use are summarized in the following table.

DECOMPOSING RATIONAL FUNCTIONS
(1) If the function is an "improper fraction", divide to obtain the sum of a polynomial and a proper fraction.
(2) Factor the denominator as far as you can, into a product of linear and quadratic terms.
(3) Use the technique of partial fractions to decompose the proper fraction into a sum of simpler terms.

We have already discussed step (1) in the SIMPLIFY chapter. If you are trying to integrate an improper rational function, your first step should *always* be to divide, and then to look for further simplifications.

Step (2), factoring the denominator, can sometimes be difficult if the denominator is complicated. The following rules from algebra often make this task easier.

Rule 1:	If a polynomial with whole numbers for coefficients has a root which is a whole number, that root is a divisor of the constant term of the polynomial.
Rule 2:	For any polynomial $P(x)$, the term $(x-a)$ is a factor of $P(x)$ if and only if $P(a) = 0$.

To see how these rules work, let's factor the polynomial

$$P(x) = x^3 + x^2 + x + 6.$$

By *Rule 1*, any number which is a root of $P(x)$ must be a divisor of the constant term 6. Thus the only *candidates* for a whole number root of $P(x)$ are

$$+1, -1, +2, -2, +3, -3, +6, \text{ and } -6.$$

Now we use *Rule 2* to see if any of these are roots of P(x). Testing the candidates one at a time, we obtain

$P(+1) = 1^3 + 1^2 + 1 + 6 = 9$, so (+1) is NOT a root of P(x).

$P(-1) = (-1)^3 + (-1)^2 + (-1) + 6 = 5$,

so (-1) is NOT a root of P(x).

$P(+2) = 2^3 + 2^2 + 2 + 6 = 20$, so (+2) is NOT a root of P(x).

$P(-2) = (-2)^3 + (-2)^2 + (-2) + 6 =$

$-8 + 4 - 2 + 6 = 0$. Thus (-2) IS a root of P(x).

Using *Rule 2*, we now have that $x - (-2) = (x + 2)$ is a *factor* of $P(x) = x^3 + x^2 + x + 6$. We can *divide* to find the other factor:

$$
\begin{array}{r|l}
 & x^2 - x + 3 \\
x + 2 & x^3 + x^2 + x + 6 \\
 & \underline{x^3 + 2x^2} \\
 & -x^2 + x \\
 & \underline{-x^2 - 2x} \\
 & 3x + 6 \\
 & \underline{3x + 6} \\
 & 0
\end{array}
$$

Thus $P(x) = x^3 + x^2 + x + 6 = (x + 2)(x^2 - x + 3)$. The quadratic term cannot be factored further, so we stop here.

Step (3) in the procedure calls for using the technique of partial fractions. Since your textbook describes it in detail, I'll just summarize it here.

The technique of partial fractions is used to decompose a *proper fraction* into a sum of basic rational functions. Make sure you have a proper fraction before you try to use the technique.

Each term in the denominator of the fraction you are trying to break up will give one or more terms when you use partial fractions.

If (ax+b) appears in the denominator, there will be a term of the form

$$\frac{A}{ax + b} \text{ in the decomposition.}$$

If $(ax+b)^n$ appears in the denominator, there will be terms of the form

$$\frac{A_1}{(ax+b)}, \frac{A_2}{(ax+b)^2}, \ldots, \frac{A_n}{(ax+b)^n} \text{ in the decomposition.}$$

If the term (ax^2+bx+c) appears in the denominator, there will be a term of the form

$$\frac{Cx+D}{ax^2+bx+c} \text{ in the decomposition.}$$

You will rarely, if ever, encounter terms like $(ax^2+bx+c)^n$ in the denominator. We will not deal with such functions here.

To use partial fractions, follow this procedure:

Step 1: Decide what terms will appear in the decomposition, using the guidelines given above. Write an equation, with the coefficients still to be determined.

Step 2: Multiply both sides of the equation by the denominator of the fraction you are trying to break up. Write both sides of the equation as polynomials in x.

Step 3: Compare the coefficients of x on both sides of the equation. These enable you to solve for the terms A, B, C, etc. in the decomposition.

SAMPLE PROBLEMS

Decompose these two functions into sums of basic functions, using the techniques we have just discussed. Make sure to try the problems before you read my solutions. Then compare your work with mine.

1. $f(x) = \dfrac{x^3 + x^2 - 6x + 5}{x^2 + x - 6}$

2. $g(x) = \dfrac{x^4 - x^3 + 3}{x^3 - 1}$

SOLUTIONS

1. $f(x) = \dfrac{x^3 + x^2 - 6x + 5}{x^2 + x - 6}$

The first thing we should do is reduce the "improper fraction" by division. That division has a quotient of (x) and a remainder of (5), so

ERIC

$$f(x) = x + \frac{5}{x^2 + x - 6}$$

$$= x + \frac{5}{(x+3)(x-2)}.$$

Since the terms in the denominator are both linear, the partial fractions decomposition will be of the form

$$\frac{5}{(x+3)(x-2)} = \frac{A}{x+3} + \frac{B}{x-2}.$$

Multiplying both sides of this equation by $(x+3)(x-2)$, we get

$$5 = A(x-2) + B(x+3), \quad \text{or}$$

$$(0)x + 5 = (A + B)x + (-2A + 3B).$$

(Remember that if a term does not appear, its coefficient is 0.)

Comparing coefficients, we obtain the equations

$$\begin{bmatrix} A + B = 0 \\ -2A + 3B = 5 \end{bmatrix}, \text{ so that } \begin{bmatrix} A = -1 \\ B = 1 \end{bmatrix}.$$

Thus

$$\frac{5}{(x+3)(x-2)} = \frac{-1}{x+3} + \frac{1}{x-2}, \quad \text{and}$$

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

2. $$g(x) = \frac{x^4 - x^3 + 3}{x^3 - 1}$$

This function is also an improper fraction, so we divide to obtain

$$g(x) = x + \frac{3}{x^3 - 1}.$$

Our next step is to factor the denominator. Since the constant

term in the denominator is 1, the only candidates for roots are $x = +1$ and $x = -1$. Since

$$(1)^3 - 1 = 0, \quad x = +1 \text{ is a root of } x^3 - 1.$$

This tells us that $(x-1)$ is a factor of (x^3-1). We can divide to find the other factor. This gives us

$$x^3 - 1 = (x - 1)(x^2 + x + 1).$$

Thus

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

and our problem is to decompose

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

into a sum of basic functions. Using the criteria on pp. 23-24, we see that the decomposition will be of the form

$$(*) \quad \frac{3}{(x-1)(x^2+x+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+x+1}.$$

Multiplying through by $(x-1)(x^2+x+1)$, we obtain

$$3 = A(x^2+x+1) + (Bx+C)(x-1)$$
$$= Ax^2 + Ax + A + Bx^2 - Bx + Cx - C.$$

Thus

$$(0)x^2 + (0)x + 3 = (A + B)x^2 + (A - B + C)x + (A - C).$$

This gives us the three equations

$$\begin{bmatrix} A + B = 0 \\ A - B + C = 0 \\ A - C = 3 \end{bmatrix}, \text{ so that } \begin{bmatrix} A = 1 \\ B = -1 \\ C = -2 \end{bmatrix}.$$

Plugging these three values back into (*), we obtain

$$\frac{3}{(x-1)(x^2+x+1)} = \frac{1}{x-1} + \frac{(-1)x + (-2)}{x^2 + x + 1}, \quad \text{so that}$$

$$g(x) = x + \frac{1}{x-1} - \frac{x+2}{x^2+x+1}.$$

Section 2

PRODUCTS

If the integrand is a product, and *especially if the integrand is a product of dissimilar functions,* you should consider using integration by parts to solve the problem. The formula is derived from the formula for the differential of a product,

$$d(uv) = u\,dv + v\,du.$$

Integrating each term, we obtain

$$uv = \int u\,dv + \int v\,du.$$

Rearranging this gives

$$\boxed{\int u\,dv = uv - \int v\,du.}$$

To apply this formula, we separate the integrand into two parts. We call one u and the other dv. We differentiate u to obtain du, and integrate dv to obtain v. If we can then integrate the term $\int v\,du$, the problem is solved. The goal of this procedure, then, is to choose u and dv such that the term $\int v\,du$ is easier to solve than the original problem. As the sample problems illustrate, this usually happens when u is simplified by differentiation. These comments are summarized in the box below.

INTEGRATING PRODUCTS

Consider integration by parts. The formula is

$$\boxed{\int u\,dv = uv - \int v\,du},$$

and your choice of u and dv should be governed by two things:

(1) You must be able to integrate the term you call dv.

(2) You want $\int v\,du$ to be easier than the original integral. This often happens when u is simplified by differentiation.

Note: This formula also has special application to the integration of single terms that we can't integrate otherwise. Since $\int f(x)dx$ can be written as $\int[f(x)][1\,dx]$, we can think of that integrand as a product and try integration by parts with $u=f(x)$ and $dv=dx$. See sample problems 3 and 4.

SAMPLE PROBLEMS

As usual, try these problems before you read my solutions. *Pay particular attention to the reasoning I use in making my choices of u and dv in each problem.*

1. $\int x \cos x\,dx$
2. $\int x^2 \tan^{-1}x\,dx$
3. $\int \sin^{-1}x\,dx$
4. $\int (\ln x)^2\,dx$

SOLUTIONS

1. $\int x \cos x\,dx$

There are two possible choices of u and dv in this problem; $\left|\begin{matrix} u = x \\ dv = \cos x\,dx\end{matrix}\right|$ and $\left|\begin{matrix} u = \cos x \\ dv = x\,dx\end{matrix}\right|$. To see which is more promising, we should determine du and v in each. In the first case we obtain $\left|\begin{matrix} du = dx \\ v = \sin x\end{matrix}\right|$, and in the second $\left|\begin{matrix} du = -\sin x\,dx \\ v = \frac{1}{2}x^2\end{matrix}\right|$. Clearly $\int v\,du$ is easier to solve in the first case, so we make the substitutions $u=x$, $dv=\cos x\,dx$. Then

$$\int \underbrace{(x)}_{u}\underbrace{(\cos x\,dx)}_{dv} = \underbrace{(x)}_{u}\underbrace{(\sin x)}_{v} - \int \underbrace{(\sin x)}_{v}\underbrace{(dx)}_{du}$$

$$= \underline{x \sin x + \cos x + C.}$$

ERIC

Note: The functions e^x, sin x, and cos x are affected about the same by either integration or differentiation. On the other hand, polynomials are usually "complicated" by integration and made simpler by differentiation. This suggests the following guideline:

Let P(x) be any polynomial. All of the integrals

$$\int P(x)\, e^x\, dx, \qquad \int P(x) \sin x\, dx, \qquad \int P(x) \cos x\, dx$$

should be done by parts, with $u = P(x)$ *and dv the remainder.*

2. $\int x^2 \tan^{-1} x\, dx$

As in problem 1, there are two reasonable choices for u and dv:

$$\left|\begin{array}{l} u = x^2 \\ dv = \tan^{-1}x\, dx \end{array}\right| \quad \text{or} \quad \left|\begin{array}{l} u = \tan^{-1}x \\ dv = x^2\, dx. \end{array}\right|$$

Let's examine which choice will help more. In the first case we will have that $du = 2x\, dx$, which is rather nice. But we will have to integrate $dv = \tan^{-1}x\, dx$, and that is no simple matter. In the second case, we will have

$$du = \frac{1}{1+x^2}\, dx \quad \text{and} \quad v = \frac{1}{3}x^3.$$

Here du is *much* simpler than u, because we've replaced an inverse tangent by a rational function! With this choice we obtain

$$\int \underbrace{(\tan^{-1}x)}_{u} \underbrace{(x^2\, dx)}_{dv} = \underbrace{(\tan^{-1}x)}_{u} \underbrace{(\tfrac{1}{3}x^3)}_{v} - \int \underbrace{(\tfrac{1}{3}x^3)}_{v} \underbrace{\left(\frac{dx}{1+x^2}\right)}_{du}$$

The second integral can now be done by the procedure for rational functions. After using the procedure, we obtain

$$\int x^2 \tan^{-1}x\, dx = \underline{\frac{1}{3}x^3 \tan^{-1}x - \frac{1}{6}x^2 + \frac{1}{6}\ln(1+x^2) + C.}$$

3. $\int \sin^{-1}x\, dx$

This integrand can be considered as a product, if we write the problem as $\int (\sin^{-1}x)(1\, dx)$. Since, as in problem 2, we obtain the greatest simplification by differentiating an inverse trigonometric function, we set

$$\left|\begin{array}{l} u = \sin^{-1}x \\ dv = 1\, dx \end{array}\right| \quad \text{so that} \quad \left|\begin{array}{l} du = \dfrac{dx}{\sqrt{1-x^2}} \\ v = x. \end{array}\right|$$

Then

$$\int \underbrace{(\sin^{-1}x)}_{u} \underbrace{(1\, dx)}_{dv} = \underbrace{(\sin^{-1}x)}_{u} \underbrace{(x)}_{v} - \int \underbrace{(x)}_{v} \underbrace{\left(\frac{dx}{\sqrt{1-x^2}}\right)}_{du}$$

$$= \underline{x \sin^{-1}x + \sqrt{1-x^2} + C.}$$

4. $\int (\ln x)^2\, dx.$

Like problem 3, this can be done by parts if we write it as $\int [(\ln x)^2][1\, dx]$. With

$$\left|\begin{array}{l} u = (\ln x)^2 \\ dv = 1\, dx \end{array}\right| \quad \text{and} \quad \left|\begin{array}{l} du = \frac{2}{x}\ln x\, dx \\ v = x \end{array}\right|$$

we obtain

$$\int \underbrace{(\ln x)^2}_{u} \underbrace{(1\, dx)}_{dv} = \underbrace{(\ln x)^2}_{u} \underbrace{(x)}_{v} - \int \underbrace{(x)}_{v} \underbrace{\left(\frac{2}{x}\ln x\, dx\right)}_{du}$$

$$= x(\ln x)^2 - 2\int \ln x\, dx.$$

We haven't solved the problem, but we've simplified it: we now have to integrate $\int (\ln x\, dx)$ instead of $\int (\ln x)^2\, dx$. A second integration by parts with $U = \ln x$, $dV = 1\, dx$ gives

$$x(\ln x)^2 - 2\int \underbrace{(\ln x)}_{U} \underbrace{(1\, dx)}_{dV} = x(\ln x)^2 - 2\Big[\underbrace{(\ln x)}_{U} \underbrace{(x)}_{V} - \int \underbrace{(x)}_{V} \underbrace{(\tfrac{1}{x}dx)}_{dU}\Big]$$

$$= \underline{x(\ln x)^2 - 2x(\ln x) + 2x + C.}$$

Note: Like many problems in integration, this can be done in more than one way. The substitution $W = \ln x$ (or $e^W = x$), transforms $\int (\ln x)^2 dx$ to $\int W^2 e^W dW$, which is done by parts (twice)

Section 3

TRIGONOMETRIC FUNCTIONS

There are many special techniques for integrating combinations of the trigonometric functions, and trying to keep track of all of them can be difficult. Instead we can keep some general guidelines for approaching trigonometric integrals in mind. The basic idea is to exploit the relationships among the trigonometric functions themselves, in order to simplify the integrand.

The first kind of manipulation we look for is a simple substitution of the kind $u = \sin x$, $u = \cos x$, etc. For this kind of substitution to be successful, the integrand should consist of an expression involving one trigonometric function, multiplied by the derivative of that function. For example,

$$\int \frac{\cos x\, dx}{1+\sin^2 x} \text{ is of the form } \int f(\sin x)[d(\sin x)],$$

where $f(\sin x) = \dfrac{1}{1+\sin^2 x}$ and $d(\sin x) = \cos x\, dx$.

In this problem we would make the substitution $u = \sin x$. Similarly, if an integral can be expressed as

$\int f(\sec x)(\sec x \tan x\, dx)$, we would set $u = \sec x$.

Our first object, then, is to manipulate an integral into the form $\int f(\sin x)(\cos x\, dx)$, etc. To do this, we try to exploit the "twin pairs" of trigonometric functions: $\sin x$ and $\cos x$, $\sec x$ and $\tan x$, and $\csc x$ and $\cot x$. The "twin pair" relationships are summarized in the table on page 32. We discuss how to use them in sample problems 1 and 2.

Twin Pairs of Trigonometric Functions

$\frac{d}{dx}(\sin x) = \cos x$	$\frac{d}{dx}(\sec x) = \sec x \tan x$	$\frac{d}{dx}(\csc x) = -\csc x \cot x$
$\frac{d}{dx}(\cos x) = -\sin x$	$\frac{d}{dx}(\tan x) = \sec^2 x$	$\frac{d}{dx}(\cot x) = -\csc^2 x$
$\sin^2 x + \cos^2 x = 1$	$\tan^2 x + 1 = \sec^2 x$	$1 + \cot^2 x = \csc^2 x$

If we are unable to exploit the "twin pairs", we turn to a different approach. The next thing we try to do is to reduce the powers of the trigonometric functions appearing in the integrand. This is usually done with the help of the formulas

$\sin^2 x = \frac{1}{2}(1-\cos 2x)$	$\cos^2 x = \frac{1}{2}(1+\cos 2x)$

or by a reduction formula obtained by using integration by parts. See sample problems 3 and 4.

Finally, there is a "last resort" technique based on the substitution $u = \tan \frac{x}{2}$. Admittedly, this formula seems to come "out of the blue". However, if nothing else seems to work when you are trying to integrate a rational function of $\sin x$ and $\cos x$, the substitutions

$u = \tan(\frac{x}{2})$,	$\sin x = \dfrac{2u}{1+u^2}$,	$\cos x = \dfrac{1-u^2}{1+u^2}$,	$dx = \dfrac{2\,du}{1+u^2}$

will transform the integrand to a rational function of u. It can then be finished by the techniques of section 1. See problem 5. In sum,

INTEGRATING TRIGONOMETRIC FUNCTIONS

(1) Exploit "twin pairs" to prepare for substitutions. Try to obtain $\int f(\sin x)(\cos x\, dx)$, etc.

(2) Reduce powers of trig functions in the integrand, by half-angle formula or integration by parts.

(3) As a last resort, the substitution $u = \tan(\frac{x}{2})$ transforms rational functions of $\sin x$ and $\cos x$ to rational functions of u.

SAMPLE PROBLEMS

1. $\int \cos^5 x\, dx$ 2. $\int \sec^3 x \tan^3 x\, dx$ 3. $\int \sin^4 x\, dx$

4. $\int \sec^3 x\, dx$ 5. $\int \frac{dx}{2 + \sin x}$

SOLUTIONS

1. $\int \cos^5 x\, dx$

As a first approach to the problem, we should try to exploit the "twin pair" of $\sin x$ and $\cos x$. Thus we should try to obtain either

(a) a function of $\cos x$, multiplied by $(-\sin x)$, or

(b) a function of $\sin x$, multiplied by $(\cos x)$.

Notice that we can achieve (b). *Since $\cos^2 x$ can be expressed in terms of $\sin^2 x$, then any even power of $\cos x$ can be expressed in terms of powers of $\sin^2 x$.* In this problem we can write $\cos^5 x = (\cos^4 x)(\cos x)$, which gives us

$$\int \cos^5 x\, dx = \int (\cos^4 x)(\cos x\, dx) = \int (1-\sin^2 x)^2 (\cos x\, dx)$$

$$= \int [\, 1 - 2\sin^2 x + \sin^4 x\,](\cos x\, dx).$$

This is now in the form $\int f(\sin x)(\cos x\, dx)$, and the substitutions $u = \sin x$, $du = \cos x\, dx$ give us

$$\int [\, 1 - 2u^2 + u^4\,]\, du = u - \frac{2}{3}u^3 + \frac{1}{5}u^5 + C$$

$$= \sin x - \frac{2}{3}\sin^3 x + \frac{1}{5}\sin^5 x + C.$$

Note: This technique will work exactly in this manner for any *odd* powers of $\cos x$ and $\sin x$.

2. $\int \sec^3 x \tan^3 x\, dx$

Since this integrand involves $\sec x$ and $\tan x$, we should see if we can express it as

(a) a function of $\sec x$, multiplied by $(\sec x \tan x)$, or

(b) a function of $\tan x$, multiplied by $(\sec^2 x)$.

In this case we can achieve (b), since factoring out the term $(\sec x \tan x)$ leaves us with $(\sec^2 x \tan^2 x)$, and the even power of $\tan x$ can be expressed in terms of secant. We have

$$\int \sec^3 x \tan^3 x\, dx = \int (\sec^2 x)(\tan^2 x)(\sec x \tan x\, dx)$$

$$= \int (\sec^2 x)(\sec^2 x - 1)(\sec x \tan x\, dx),$$

and the substitution $u = \sec x$ gives us

$$\int u^2(u^2-1)\, du = \int (u^4 - u^2)\, du = \frac{1}{5}u^5 - \frac{1}{3}u^3 + C$$

$$= \frac{1}{5}\sec^5 x - \frac{1}{3}\sec^3 x + C.$$

3. $\int \sin^4 x\, dx$

Technique (1) doesn't help us in this problem: if we try to separate out $(\sin x\, dx)$, we're left with the term $(\sin^3 x\, dx)$, which can't be expressed as a polynomial in its twin, $\cos x$. Instead we turn to technique (2) and use the double-angle formula:

$$\int \sin^4 x\, dx = \int (\sin^2 x)^2\, dx = \int [\tfrac{1}{2}(1-\cos 2x)]^2\, dx$$

$$= \frac{1}{4}\int (1 - 2\cos 2x + \cos^2 2x)\, dx.$$

The first two terms in this expression can be integrated easily, and we can again call on a double-angle formula to express $\cos^2 2x = \frac{1}{2}(1 + \cos 4x)$. This gives us

$$\frac{1}{4}\int dx - \frac{1}{2}\int \cos 2x\, dx + \frac{1}{8}\int dx + \frac{1}{8}\int \cos 4x\, dx =$$

$$\frac{x}{4} - \frac{\sin 2x}{4} + \frac{x}{8} + \frac{\sin 4x}{32} + C = \frac{\sin 4x - 8\sin 2x + 12x}{32} + C.$$

4. $\int \sec^3 x\,dx$

This is a difficult problem. We'll go through it slowly and in detail, so that the reasoning for it and problems like it becomes apparent. We begin by noticing that technique (1) doesn't work for us here and that the double-angle formulas don't apply, so we decide to use integration by parts. There are three reasonable choices:

(a) $\int \underbrace{(\sec^3 x)}_{u}\underbrace{(1\,dx)}_{dv}$, (b) $\int \underbrace{(\sec^2 x)}_{u}\underbrace{(\sec x\,dx)}_{dv}$, and

(c) $\int \underbrace{(\sec x)}_{u}\underbrace{(\sec^2 x\,dx)}_{dv}$

In choice (a), setting dv=dx would lead to v=x, and the term (v du) would involve both x and a combination of trig functions. Integrating that looks difficult, so we go on to try something else. In choice (b), setting dv = (sec x dx) leads to v = ln|sec x + tan x|, which is nasty. Thus we examine choice (c). Since

$$\left|\begin{array}{l} u = \sec x \\ dv = \sec^2 x\,dx \end{array}\right| \text{ gives } \left|\begin{array}{l} du = \sec x \tan x\,dx \\ v = \tan x \end{array}\right|$$

and this is the best of the three alternatives, we proceed :

$$\int \underbrace{(\sec x)}_{u}\underbrace{(\sec^2 x\,dx)}_{dv} = \underbrace{(\sec x)}_{u}\underbrace{(\tan x)}_{v} - \int \underbrace{(\tan x)}_{v}\underbrace{(\sec x \tan x\,dx)}_{du}$$

$$= \sec x \tan x - \int (\sec x)(\tan^2 x\,dx).$$

We can use the identity $(\tan^2 x = \sec^2 x - 1)$ to obtain

$$\int \sec^3 x\,dx = \sec x \tan x - \int (\sec x)(\sec^2 x - 1)\,dx, \text{ or}$$

$$(*)\quad \int \sec^3 x\,dx = \sec x \tan x - \int \sec^3 x\,dx + \int \sec x\,dx.$$

For a moment it looks as if we've gone around in circles, because we now have the term

$$U = \int \sec^3 x\,dx$$

on *both sides* of equation (*). Notice, however, that U appears with a *negative* sign on the right-hand side of (*). We can then consider (*) as an *algebraic* equation,

$$U = (\sec x \tan x) - U + \left(\int \sec x\,dx\right).$$

Solving this equation for U, we obtain

$$2U = \sec x \tan x + \int \sec x\,dx$$

$$= \sec x \tan x + \ln|\sec x + \tan x| + C.$$

Dividing both sides of this equation by 2, and replacing U by $\int \sec^3 x\,dx$, we finally obtain

$$\int \sec^3 x\,dx = \frac{1}{2}\left[\sec x \tan x + \ln|\sec x + \tan x|\right] + C',$$

where C'=C/2.

Note: This is a long and involved procedure. With minor modifications, it will provide reduction formulas for powers of all the trigonometric functions. Because of its complexity, however, you should only consider using it after checking that technique (1) and the double-angle formulas don't help.

5. $\int \frac{dx}{2 + \sin x}$

In this problem, neither the "twin pairs" or reduction formulas seem to help, so we make use of the "last resort" substitution given in technique (3). The substitutions

$$u = \tan(\tfrac{x}{2});\quad \sin x = \frac{2u}{1+u^2};\quad \cos x = \frac{1-u^2}{1+u^2};\quad dx = \frac{2\,du}{1+u^2}$$

transform the integral to

$$\int \frac{\frac{2\,du}{1+u^2}}{2+\frac{2u}{1+u^2}} = \int \frac{2\,du}{2u^2+2u+2} = \int \frac{du}{u^2+u+1}.$$

This is a rational function, and is done by completing the square in the denominator:

$$\int \frac{du}{u^2+u+1} = \int \frac{du}{(u+1/2)^2+(3/4)} = \frac{1}{\sqrt{3/4}}\tan^{-1}\left[\frac{u+1/2}{\sqrt{3/4}}\right] + C$$

$$= \frac{1}{\sqrt{3/4}}\tan^{-1}\left[\frac{\tan(\frac{x}{2}) + (1/2)}{\sqrt{3/4}}\right] + C.$$

Section 4

SPECIAL FUNCTIONS

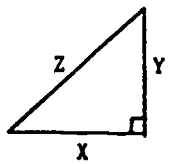
In this section we will discuss three kinds of substitutions which occur often enough that they are worth singling out for special mention. The first type of substitution deals with terms of the form

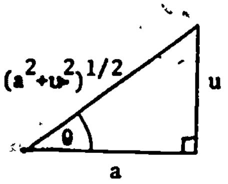
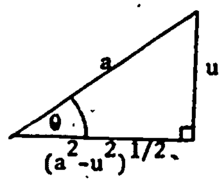
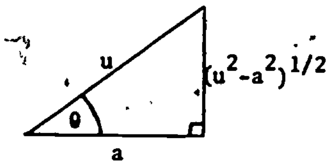
$$(a^2+u^2)^{n/2}, \quad (a^2-u^2)^{n/2}, \quad \text{and} \quad (u^2-a^2)^{n/2}.$$

We deal with functions like these by making a trigonometric substitutions for one of the terms

$$(*) \qquad (a^2+u^2)^{1/2}, \quad (a^2-u^2)^{1/2}, \quad \text{or} \quad (u^2-a^2)^{1/2}.$$

The substitutions can be memorized, but I find it easier to draw a triangle and derive them. All of the substitutions come from the Pythagorean theorem, which says that $X^2 + Y^2 = Z^2$ in the triangle to the right. If we place the sides a and u on the triangle carefully, we can make the third side of the triangle be any of the terms in (*). See the triangles below.



for $(a^2+u^2)^{1/2}$	for $(a^2-u^2)^{1/2}$	for $(u^2-a^2)^{1/2}$
		

Once the triangle has been drawn and labeled, we can "read" whatever substitution we need from it. Follow this procedure.

To obtain an expression for u, use the trigonometric function that involves u and a. Once you have u as a trigonometric function of θ, differentiate to find du.

To obtain an expression for $(\text{something})^{1/2}$, *use the trigonometric function that involves the sides* $(\text{something})^{1/2}$ *and a in the triangle.*

Make the substitutions. The result will be a trigonometric integral, which you can solve in terms of θ. To express the answer in terms of x, "read" the functions from the triangle.

Sample problems 1 and 2 will illustrate how to use this procedure. See page 41 for the second and third kinds of substitutions we discuss in this section.

SAMPLE PROBLEMS

1. $\displaystyle\int \frac{dx}{(x^2+9)^{3/2}}$ 2. $\displaystyle\int \frac{x^2\,dx}{\sqrt{4-9x^2}}$

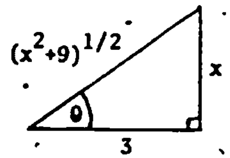
SOLUTIONS

1. $\displaystyle\int \frac{dx}{(x^2+9)^{3/2}}$

Before we try a trigonometric substitution, we should check for any SIMPLIFICATIONS. Unfortunately there are none, so we draw a triangle. In this case the term we wish to substitute for is

$$(x^2+9)^{1/2},$$

so we draw the triangle



ERIC

To obtain the substitution for x, we use the function that involves the sides (3) and (x). In this case

$$\tan\theta = \frac{x}{3}, \quad \text{so} \quad x = 3\tan\theta \quad \text{and} \quad dx = 3\sec^2\theta\, d\theta.$$

To obtain the substitution for $(x^2+9)^{1/2}$, we use the function that involves $(x^2+9)^{1/2}$ and (3). This gives us

$$\sec\theta = \frac{(x^2+9)^{1/2}}{3}, \quad \text{or} \quad (x^2+9)^{1/2} = 3\sec\theta.$$

We are now ready to substitute these into the problem. We get

$$\int \frac{dx}{(x^2+9)^{3/2}} = \int \frac{dx}{[(x^2+9)^{1/2}]^3} = \int \frac{3\sec^2\theta\, d\theta}{(3\sec\theta)^3}$$

$$= \int \frac{3\sec^2\theta\, d\theta}{27\sec^3\theta} = \frac{1}{9}\int \cos\theta\, d\theta$$

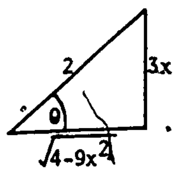
$$= \frac{1}{9}\sin\theta + C.$$

We now return to the triangle to obtain the value of sin θ. This gives us the final answer

$$\frac{1}{9}\left(\frac{x}{(x^2+9)^{1/2}}\right) + C.$$

2. $\int \frac{x^2\, dx}{\sqrt{4-9x^2}}$

In this problem the term $\sqrt{4-9x^2}$ is of the form $\sqrt{a^2-u^2}$, and suggests a triangle with hypotenuse 2 and leg 3x, like the one drawn to the right.



We first determine x and dx by using the trigonometrics function involving (3x) and (2). This gives us

$$\sin\theta = \frac{3x}{2}, \quad \text{so} \quad x = \frac{2}{3}\sin\theta \quad \text{and} \quad dx = \frac{2}{3}\cos\theta\, d\theta.$$

To substitute for $\sqrt{4-9x^2}$, we use the trigonometric function involving that term and the constant. Here

$$\cos\theta = \frac{\sqrt{4-9x^2}}{2}, \quad \text{so} \quad \sqrt{4-9x^2} = 2\cos\theta.$$

At this point we're ready to substitute in the integral. We obtain

$$\int \frac{x^2\, dx}{\sqrt{4-9x^2}} = \int \frac{(\frac{2}{3}\sin\theta)^2\, (\frac{2}{3}\cos\theta\, d\theta)}{2\cos\theta}$$

$$= \frac{4}{27}\int \sin^2\theta = \frac{4}{27}\int (\tfrac{1}{2})(1 - \cos 2\theta)\, d\theta$$

$$= \frac{2}{27}\int d\theta - \frac{2}{27}\int \cos 2\theta\, d\theta$$

$$= \frac{2}{27}\theta - \frac{1}{27}\sin 2\theta + C$$

$$= \frac{2}{27}(\theta - \sin\theta\cos\theta) + C.$$

To complete the problem, we need only read off the values of the functions of θ from the triangle. Sin θ and cos θ are $\frac{3x}{2}$ and $\frac{\sqrt{4-9x^2}}{2}$, respectively. To find θ, we can use the function sin θ: since $\sin\theta = \frac{3x}{2}$, $\theta = \sin^{-1}(\frac{3x}{2})$.

Thus

$$\int \frac{x^2\, dx}{\sqrt{4-9x^2}} = \frac{2}{27}\left[\sin^{-1}(\tfrac{3x}{2}) - (\tfrac{3x}{2})(\tfrac{\sqrt{4-9x^2}}{2})\right] + C.$$

ERIC

The second and third types of substitution we discuss in this section are really special cases of a suggestion we discussed in Chapter 1, where we noted that it is often worth considering substitutions for the "nasty" terms in integrands. Expressions involving e^x and $\sqrt[n]{ax+b}$ occur often enough to justify listing these substitutions.

We frequently encounter integrals like

$$\int \frac{1}{e^x+1}\,dx, \int \frac{1}{e^x - e^{-x}}\,dx, \quad \text{and} \quad \int \frac{e^{3x}+1}{e^{2x}+1}\,dx,$$

which are rational functions of e^x. At first glance it looks like the substitution $u = e^x$ will not be of assistance, because the term $du = e^x\,dx$ is missing. *You should make the substitution anyway!*

$$\text{If } u = e^x, \text{ then } du = e^x\,dx = u\,dx, \text{ so } dx = \frac{1}{u}\,du.$$

If you are trying to integrate a rational function of e^x, make the substitutions

$$e^x = u \quad \text{and} \quad dx = \frac{1}{u}\,du.$$

The result will be a rational function of u.

A similar comment holds for integrals which include terms of the form

$$\sqrt[n]{ax+b}.$$

If we set $u = \sqrt[n]{ax+b}$, then $u^n = ax+b$, and $x = \frac{1}{a}(u^n-b)$. Differentiating, we obtain $dx = \frac{n}{a}u^{n-1}\,du$.

If you are trying to solve an integral which is a rational function of x and $\sqrt[n]{ax+b}$, make the substitutions

$$\sqrt[n]{ax+b} = u, \quad x = \frac{1}{a}(u^n-b), \quad \text{and} \quad dx = \frac{n}{a}u^{n-1}\,dx.$$

The result of these substitutions will be a rational function of u.

See sample problems 3 and 4 for these substitutions. The table on page 42 summarizes this section.

INTEGRATING SPECIAL FUNCTIONS

(1) *If the integrand includes terms of the form*

$$(a^2-u^2)^{n/2}, \quad (u^2-a^2)^{n/2}, \quad \text{or} \quad (a^2+u^2)^{n/2},$$

(a) Draw a right triangle.
(b) Place a and u so that the third side of the triangle is the term you want.
(c) "Read" the substitutions from the triangle.

(2) *If the integrand is a rational function of e^x,* make the substitutions

$$e^x = u \quad \text{and} \quad dx = \frac{1}{u}\,du.$$

(3) *If the integrand is a rational function of x and $\sqrt[n]{ax+b}$,* make the substitutions

$$\sqrt[n]{ax+b} = u, \quad x = \frac{1}{a}(u^n-b), \quad \text{and} \quad dx = \frac{n}{a}u^{n-1}\,du.$$

SAMPLE PROBLEMS

3. $\int \frac{1}{e^x - e^{-x}}\,dx$

4. $\int \frac{1}{x}\sqrt[3]{2x+1}\,dx$

SOLUTIONS

3. $\int \frac{1}{e^x - e^{-x}}\,dx$

The integrand in this problem is a rational function of e^x. Therefore we should make the substitutions

$$e^x = u, \quad dx = \frac{1}{u}\,du,$$

even though $(e^x dx)$ does not appear in the numerator. Since $e^{-x} = \frac{1}{e^x} = \frac{1}{u}$, the integral becomes

$$\int\left(\frac{1}{u-\frac{1}{u}}\right)\left(\frac{1}{u}\,du\right) = \int\frac{du}{(u-\frac{1}{u})(u)} = \int\frac{du}{u^2-1}.$$

Using partial fractions or the formula on page 17, this is

$$\int\frac{1}{2}\left(\frac{1}{u-1}-\frac{1}{u+1}\right)du = \frac{1}{2}\left[\ln|u-1| - \ln|u+1|\right]+C$$

$$= \frac{1}{2}\ln\left|\frac{u-1}{u+1}\right|+C = \frac{1}{2}\ln\left|\frac{e^x-1}{e^x+1}\right|+C.$$

4. $\int \frac{1}{x}\sqrt[3]{2x+1}\,dx$

The integrand in this problem is a rational function of x and $\sqrt[3]{2x+1}$, so we should make the substitutions

$$u = \sqrt[3]{2x+1}\ ;\quad u^3 = 2x+1\ ;\quad x = \tfrac{1}{2}(u^3-1)\ ;\quad dx = \tfrac{3}{2}u^2\,du.$$

The integral then becomes

$$\int\left(\frac{1}{x}\right)\left(\sqrt[3]{2x+1}\right)(dx) = \int\left(\frac{1}{\frac{1}{2}(u^3-1)}\right)(u)\left(\frac{3}{2}u^2\,du\right)$$

$$= 3\int\frac{u^3}{u^3-1}\,du.$$

Using the technique for rational functions, this becomes

$$\int\left(3+\frac{3}{u^3-1}\right)du = \int\left(3+\frac{1}{u-1}-\frac{u+2}{u^2+u+1}\right)du$$

$$= 3u + \ln|u-1| - \frac{1}{2}\ln|u^2+u+1| - (\sqrt{3})\tan^{-1}\left(\frac{2u+1}{\sqrt{3}}\right) + C,$$

where $u = \sqrt[3]{2x+1}$.

EXERCISES FOR CHAPTER 2

PART 1: The purpose of these exercises is to give you practice in the SIMPLIFY and CLASSIFY steps of the General Procedure. *DO NOT SOLVE THE INTEGRALS AT THIS POINT.* Examine them for simplifications, classify them, and *decide which technique you would use to solve them.* Then compare your reasoning with mine, in the solutions manual. Remember:

[1. sol. 6] means that solution #6 presents a discussion of exercise 1.

1. (sol. 6) $\int\frac{7}{x\sqrt{x^2+4}}\,dx$

2. (sol. 10) $\int 2\tan^4 x\,dx$

3. (sol. 4) $\int\frac{4}{e^x-1}\,dx$

4. (sol. 1) $\int\frac{6x^2}{\sqrt{3x+1}}\,dx$

5. (sol. 13) $\int\frac{9x}{\sqrt{x^2+4}}\,dx$

6. (sol. 8) $\int x\tan^{-1}x\,dx$

7. (sol. 14) $\int\frac{5x^3}{x^4-1}\,dx$

8. (sol. 11) $\frac{5}{\sqrt{x^2+6x}}\,dx$

9. (sol. 9) $\int\frac{\cos(\ln x)}{x}\,dx$

10. (sol. 2) $\int\frac{9}{2+\cos x}\,dx$

The exercises are continued on page 45...

11. sol. 16 $\int \csc^3 x \cot^3 x \, dx$

12. sol. 7 $\int \frac{x^3 + x^2}{x^2 + x - 2} \, dx$

13. sol. 5 $\int \frac{\tan^{-1} x}{x^2 + 1} \, dx$

14. sol. 15 $\int \csc^2 x \cot^3 x \, dx$

15. sol. 3 $\int (\sin^2 x - \cos^2 x) \, dx$

16. sol. 12 $\int \frac{x^4}{x^3 - 1} \, dx$

*** *PART 2* ***

Solve each of the exercises from part 1. Detailed solutions are in the solutions manual. This table lists the number of the solution to each exercise below the number of the exercise.

Exercise #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Solution #	22	26	20	17	29	24	30	27	25	18	32	23	21	31	19	28

Chapter 3

MODIFY!

Chapters 1 and 2 of this booklet contain the basic techniques necessary for solving most first-year calculus integration problems. Once we can SIMPLIFY or CLASSIFY an integrand, its solution is a routine (although not necessarily easy) matter.

We encounter the most difficuly with problems of unfamiliar form, those which resist classification by the methods of Chapter 2. With such problems our goal is to MODIFY the integrand, manipulating it until it is in a more convenient or recognizable form. Once this has been done, we return to the SIMPLIFY and CLASSIFY steps of the General Procedure to finish the problem.

The three sections of this chapter are:

(1) Problem Similarities: looking for and exploiting resemblances between the problem we are working on and problems we know how to integrate

(2) Special Manipulations: techniques for expressing complicated integrands in more convenient form

(3) Needs Analysis: looking to see what additional terms might help solve a problem, and modifying the integrand to include them.

Together, these form the third step of the General Procedure:

Step 3: MODIFY		
Problem Similarities	Special Manipulations	Needs Analysis

Section 1

PROBLEM SIMILARITIES

Some integrals can be classified easily, but look so complicated that the standard procedures for solving them promise to be very messy. Other integrals may not fit into the classification scheme of Chapter 2, and we may not know an appropriate way to solve them. One way to approach such problems is to look for similarities between them and problems we know how to do. If the form of a difficult problem resembles that of a "standard" problem, there are two possibilities. We might be able to reduce the difficult problem to that "standard" form. Or, the techniques we would use on the easier problem might help us solve the more difficult one. The sample problems will illustrate this kind of approach. Summarized in table form, we have

PROBLEM SIMILARITIES
(1) Look for easy problems similar to the one you are working on.
(2) Try to reduce the difficult problem to the form of the easy similar problem.
(3) Try the techniques you would use on the similar problem.

SAMPLE PROBLEMS

Use the suggestions given on page 47 to try to solve these problems. Then compare your solution with mine.

1. $\int \frac{x}{1+x^4}\,dx$

2. $\int \frac{x^2}{x^6 - 9x^3 + 8}\,dx$

3. $\int \frac{1}{(x+1)\sqrt{x^2+2x}}\,dx$

SOLUTIONS

1. $\int \frac{x}{1+x^4}\,dx$

The integrand in this problem is a rational function, so we could solve the problem by the procedures of chapter 2. The denominator is difficult to factor, however, so we look for another approach.

The problem would be easy of the denominator were $(1 + x^2)$ instead of $(1 + x^4)$; can that be arranged? Yes, because of the x term in the numerator. Making the substitutions

$$u = x^2, \quad du = 2x\,dx,$$

we get

$$\int \frac{x}{1+x^4}\,dx = \frac{1}{2}\int \frac{2x\,dx}{1+x^4} = \frac{1}{2}\int \frac{du}{1+u^2} = \frac{1}{2}\tan^{-1}u + C$$

$$= \frac{1}{2}\tan^{-1}(x^2) + C.$$

2. $\int \frac{x^2}{x^6 - 9x^3 + 8}\,dx$

This problem, like problem 1, can be solved directly by the procedure for rational functions. The denominator factors without difficulty to give us

$$\int \frac{x^2\,dx}{(x^3-1)(x^3-8)}.$$

We *could* continue factoring the denominator, use partial fractions, and then integrate term by term. That promises to be a very involved procedure, however. We should stop and look for other alternatives.

Note that the integrand *resembles* a simple rational function with a quadratic denominator: instead of $(x-1)(x-8)$, we have $(x^3-1)(x^3-8)$. Can we simplify the denominator? Yes, since the term $(x^2\,dx)$ appears in the numerator. With the substitutions

$$u = x^3; \quad du = 3x^2\,dx,$$

we obtain

$$\frac{1}{3}\int \frac{du}{(u-1)(u-8)} = \frac{1}{3}\int \left(\frac{(-\frac{1}{7})}{u-1} + \frac{(\frac{1}{7})}{u-8} \right) du$$

$$= \frac{1}{21}\int \left(\frac{1}{u-8} - \frac{1}{u-1} \right) du$$

$$= \frac{1}{21}\left(\ln|u-8| - \ln|u-1| \right) + C = \frac{1}{21}\ln\left|\frac{u-8}{u-1}\right| + C$$

$$= \frac{1}{21}\ln\left|\frac{x^3-8}{x^3-1}\right| + C.$$

Note: Substitutions like this might have occurred to you after working through chapter 1. If so, terrific! Our guiding principle is: *at every stage of a problem, look for easy alternatives.* As you gain experience, your catalogue of SIMPLIFYING techniques will grow.

3. $\int \frac{1}{(x+1)\sqrt{x^2+2x}}\,dx$

As a preliminary simplification, we might consider a substitution for the "nasty" term in the denominator: $u = x^2+2x$. This leads to $du = (2x+2)dx$, and at first glance this looks promising. Unfortunately, the term $(x+1)$ is in the denominator, instead of the numerator, where we would like it! So we abandon this substitution temporarily, in the hope we can find something easier.

Looking for similarities, we can ask: are there any "standard forms" that include square roots in the denominator? Yes, terms like $\int \frac{du}{\sqrt{u^2 \pm a^2}}$, $\int \frac{du}{u\sqrt{u^2 \pm a^2}}$, etc. This suggests completing the square, in the hope that we get something easier to handle. We have $[x^2+2x] = [(x+1)^2 - 1]$, which suggests the substitution $u = (x+1)$. Then

$$\int \frac{dx}{(x+1)\sqrt{x^2+2x}} = \int \frac{du}{u\sqrt{u^2-1}} = \sec^{-1}u + C$$

$$= \sec^{-1}(x+1) + C.$$

Section 2

SPECIAL MANIPULATIONS

In this section we discuss four techniques designed to express complicated integrands in more convenient form for integration. They are

SPECIAL MANIPULATIONS
A. Rationalizing denominators of quotients
B. Special use of trigonometric identities
C. "Common denominator" substitutions
D. "Desperation" substitutions

These techniques often involve complex manipulations. It may not be clear that they are helping to solve a problem until we have done some complicated calculations. For that reason these techniques differ from the simplifications of Chapter 1. When we first examine an integral, we look for fast and easy ways to solve it. If that fails, we try to classify it and use standard techniques. Only if that fails, or if the standard techniques look <u>very</u> complicated, do we look for alternatives such as these. With practice you will discover which approaches to integrals you can examine rapidly, and which are time-consuming. This knowledge should govern the order in which you apply them.

A. RATIONALIZING DENOMINATORS

This technique is based on the relation $(A+B)(A-B) = A^2-B^2$. If we replace A by $\sqrt{u}$ and B by $\sqrt{v}$, we obtain

$$(\sqrt{u} + \sqrt{v})(\sqrt{u} - \sqrt{v}) = u - v.$$

If the integrand is a fraction whose denominator is of the form $(\sqrt{u} \pm \sqrt{v})$, multiply both the numerator and denominator by its "conjugate," $(\sqrt{u} \mp \sqrt{v})$. The denominator of the resulting fraction is simply $(u - v)$.

See sample problems 1 and 2.

B. SPECIAL USE OF TRIGONOMETRIC IDENTITIES

The basic trigonometric identities, like the terms discussed in *(A)*, can be written as the difference of two squares. For example,

$$(1+\cos x)(1-\cos x) = 1-\cos^2 x = \sin^2 x;$$

$$(1+\sin x)(1-\sin x) = 1-\sin^2 x = \cos^2 x;$$

$$(\sec x + \tan x)(\sec x - \tan x) = \sec^2 x - \tan^2 x = 1;$$

$$(\csc x + \cot x)(\csc x - \cot x) = \csc^2 x - \cot^2 x = 1.$$

The terms paired above, like $(1+\cos x)$ and $(1-\cos x)$, are called *conjugates*.

If the integrand contains any of the terms $(1 \pm \cos x)$, $(1 \pm \sin x)$, $(\sec x \pm \tan x)$, or $(\csc x \pm \cot x)$, either in the denominator of a fraction or inside a square root, consider multiplying and dividing the integrand by its conjugate.

See sample problems 3 and 4.

C. "COMMON DENOMINATOR" SUBSTITUTIONS

When an integrand involves a single term like $\sqrt[n]{x} = x^{1/n}$, we make the substitution $u = x^{1/n}$, or equivalently, $u^n = x$. The result of this substitution is an integrand which has *integer* (whole number) powers of u instead of *fractional* powers of x.

Some integrands involve more than one fractional power of x, like

$$\int \frac{x^{1/3} + 4}{x^{1/2} + x^{2/3}}\,dx$$

To solve an integral like this, we would like to find a substitution $u = x^{1/N}$ such that *all* of the fractional powers of x are replaced by integer powers of u. We choose N as follows.

Let N be the smallest common denominator of all the fractional powers of x which appear in the integrand. Make the substitution $u = x^{1/N}$, so that $x = u^N$ and $dx = N u^{N-1}\,du$. The integrand which results from this substitution will be a rational function of u.

In the problem above, the smallest common denominator of $\frac{1}{3}$, $\frac{1}{2}$, and $\frac{2}{3}$ is 6. Thus we should make the substitutions

$$u = x^{1/6}; \quad x = u^6; \quad dx = 6u^5\,du.$$

The integral then becomes

$$\int \left(\frac{u^2 + 4}{u^3 + u^4}\right)(6u^5\,du),$$

which can be solved by the procedure for rational functions. See sample problem 5.

D. "DESPERATION" SUBSTITUTIONS

Our guideline in Chapter 1 was that we should only consider substitutions that are quick and easy to use, and we postponed looking at any substitutions that looked complicated or unpromising. If neither the SIMPLIFY nor CLASSIFY steps help us solve a problem, we should now consider more complicated substitutions in the *hope* that they will prove helpful. At this stage we have little to lose. For example, to solve

$$\int \left(\sqrt{1 + \sqrt{x}}\right) dx,$$

we might try $u = 1 + \sqrt{x}$ or even $u = \sqrt{1 + \sqrt{x}}$. See problem 6. To solve

$$\int \left(\sqrt{\frac{x+1}{x}}\right) dx,$$

we might try $u = \frac{x+1}{x}$ or even $u = \sqrt{\frac{x+1}{x}}$.

REMEMBER: Our goal is to manipulate the integrand until it takes a familiar or convenient form. As soon as we succeed, we return to the SIMPLIFY and CLASSIFY techniques of chapters 1 and 2.

SAMPLE PROBLEMS

Try each problem before you read the solution. Then compare your reasoning with mine.

1. $\int \frac{dx}{\sqrt{x+1} + \sqrt{x-1}}$

2. $\int \frac{x\,dx}{1 - \sqrt{1-x}}$

3. $\int \frac{dx}{1 + \cos x}$

4. $\int \sqrt{1 - \cos x}\,dx$

5. $\int \frac{dx}{x^{1/3} - x^{1/2}}$

6. $\int \sqrt{1 + \sqrt{x}}\,dx$

ERIC Full Text Provided by ERIC

SOLUTIONS

1. $$\int \frac{dx}{\sqrt{x+1} + \sqrt{x-1}}$$

To solve this problem, we multiply both numerator and denominator by the conjugate term $(\sqrt{x+1} - \sqrt{x-1})$. This gives us

$$\int \frac{(\sqrt{x+1} - \sqrt{x-1})\,dx}{(\sqrt{x+1} + \sqrt{x-1})(\sqrt{x+1} - \sqrt{x-1})} = \int \frac{(\sqrt{x+1} - \sqrt{x-1})\,dx}{(x+1) - (x-1)} =$$

$$\int \frac{(\sqrt{x+1} - \sqrt{x-1})\,dx}{2} = \frac{1}{2}\int (x+1)^{1/2}dx - \frac{1}{2}\int (x-1)^{1/2}dx$$

$$= \frac{1}{3}(x+1)^{3/2} - \frac{1}{3}(x-1)^{3/2} + C.$$

2. $$\int \frac{x\,dx}{1 - \sqrt{1-x}}$$

Here too we multiply numerator and denominator by the conjugate term, $(1 + \sqrt{1-x})$. This gives us

$$\int \frac{x(1 + \sqrt{1-x})\,dx}{(1-\sqrt{1-x})(1+\sqrt{1-x})} = \int \frac{x(1 + \sqrt{1-x})\,dx}{1 - (1-x)} =$$

$$\int \frac{x(1 + \sqrt{1-x})\,dx}{x} = \int (1 + (1-x)^{1/2})\,dx$$

$$= x - \frac{2}{3}(1-x)^{3/2} + C.$$

3. $$\int \frac{dx}{1 + \cos x}$$

Since the integrand is a rational function of cos x, we could use the substitution $u = \tan \frac{x}{2}$ to transform it to a rational function of u. Since working with conjugates in this case is fairly easy, we can try that first and see what happens. We get

$$\int \frac{dx}{1+\cos x} = \int \frac{(1-\cos x)\,dx}{(1+\cos x)(1-\cos x)} = \int \frac{(1-\cos x)\,dx}{\sin^2 x} =$$

$$\int \frac{dx}{\sin^2 x} - \int \frac{\cos x\,dx}{\sin^2 x} = \int \csc^2 x\,dx - \int \frac{du}{u^2}, \text{ (where } u = \sin x)$$

$$= -\cot x + \frac{1}{u} + C = -\cot x + \frac{1}{\sin x} + C$$

$$= -\cot x + \csc x + C.$$

4. $$\int \sqrt{1 - \cos x}\,dx$$

In this problem the "nasty" term is inside the square root. If we multiply (1-cos x) by its conjugate (1+cos x), we obtain $\sin^2 x$, and the square root of that is just sin x. For that reason we can try the technique, in the hope that the result is simpler to work with. If it isn't, we would look for something else.

$$\int \sqrt{1-\cos x}\,dx = \int \sqrt{\frac{(1-\cos x)(1+\cos x)}{(1+\cos x)}}\,dx =$$

$$\int \sqrt{\frac{\sin^2 x}{1+\cos x}} = \int \frac{\sqrt{\sin^2 x}}{\sqrt{1+\cos x}}\,dx = \int \frac{\sin x\,dx}{\sqrt{1+\cos x}}.$$

This may look as complicated as the integral we started with, but is much easier and can be done by the techniques of Chapter 1. We have the term cos x in the denominator,

ERIC

and (almost) its derivative in the numerator. Making the substitution $u = \cos x$, the integral becomes

$$\int \frac{-du}{\sqrt{1+u}} = \int -(1+u)^{-1/2} = -2(1+u)^{1/2} + C$$

$$= \underline{-2\sqrt{1+\cos x} + C.}$$

5. $\displaystyle\int \frac{dx}{x^{1/3} - x^{1/2}}$

This problem involves fractional exponents. The least common denominator of $\frac{1}{2}$ and $\frac{1}{3}$ is 6, so we make the substitution

$$u = x^{1/6}, \quad \text{so} \quad x = u^6 \quad \text{and} \quad dx = 6u^5\,du.$$

The integral becomes

$$\int \frac{6u^5\,du}{u^2 - u^3} = 6\int \frac{u^3\,du}{1-u} = -6\int \frac{u^3\,du}{u-1} =$$

$$-6\int \left(u^2 + u + 1 + \frac{1}{u-1} \right) du =$$

$$-\left(2u^3 + 3u^2 + u + \ln|u-1|\right) + C =$$

$$\underline{-\left(2x^{1/2} + 3x^{1/3} + 6x^{1/6} + 6\ln\left|x^{1/6} - 1\right|\right) + C.}$$

6. $\displaystyle\int \sqrt{1 + \sqrt{x}}\;dx$

This problem can be done by sequential substitutions $u = \sqrt{x}$; $v = 1+u$; $w = \sqrt{v}$. As an example of a "desperation" substitution, however, we might try

$u = \sqrt{1 + \sqrt{x}}$. Then $u^2 = 1 + \sqrt{x}$; $x = (u^2-1)^2$; and $dx = 4u(u^2-1)\,du$. Then

$$\int \sqrt{1 + \sqrt{x}}\;dx = \int (u)[4u(u^2-1)]\,du = \int (4u^4 - 4u^2)\,du =$$

$$\frac{4}{5}u^5 - \frac{4}{3}u^3 + C = \underline{\frac{4}{5}(1 + \sqrt{x})^{5/2} - \frac{4}{3}(1 + \sqrt{x})^{3/2} + C.}$$

Section 3

NEEDS ANALYSIS

The technique of needs analysis has been implicit in much of our work so far, and we now state it formally as an integration technique. It consists of asking what might enable us to solve a problem, and then either adding it (and compensating for it) or changing something in the problem to it. Needs analysis explains the reasoning behind our exploiting "twin pairs" of trigonometric functions, for example. If an integrand is a complicated expression involving sin x, we search for a way to introduce the term (cos x dx). Conversely, if (cos x dx) appeared in the integrand, we might seek to express the rest of the integrand in terms of sin x. For an integrand involving e^x, we might seek to introduce $(e^x dx)$. [This is done automatically by the substitutions $u=e^x$; $du=e^x dx$; $dx = \frac{1}{u}du$. An alternate strategy is given in sample problem 1.] If the integrand involves x^n, we can look for a way to introduce $[nx^{n-1}\,dx]$. As usual, we summarize in table form.

NEEDS ANALYSIS
(1) Look for a term, or a form of the integral, that would enable you to solve it.
(2) Try to modify the integral to produce the term or form you need.
(3) Try to introduce the term you need, and compensate for it.

SAMPLE PROBLEMS

Try to solve each of these problems using a needs analysis. Then compare your solution with mine.

1. $\int \frac{dx}{e^x - e^{-x}}$

3. $\int \frac{dx}{x(ax^n+b)}$

2. $\int \frac{\sec^2 x\, dx}{\sqrt{5 - \sec^2 x}}$

4. $\int \frac{dx}{(\sin x + 6)(\cos x)}$

SOLUTIONS

1. $\int \frac{dx}{e^x - e^{-x}}$

We solved this problem before on page 42, where the procedure for special functions called for the substitutions

$$e^x = u \quad \text{and} \quad dx = \frac{1}{u}du.$$

Needs analysis provides another route to a solution. Since the integrand is a rational function of e^x, I would like to make the substitution $u = e^x$. This would work most easily *if the term $du = e^x\,dx$ were present in the integrand.* I can obtain it, if I multiply numerator and denominator of the integrand by e^x. This gives me

$$\int \frac{e^x\,dx}{(e^x)(e^x - e^{-x})} = \int \frac{e^x\,dx}{e^{2x} - 1},$$

and *now* the substitution $u = e^x$ gives

$$\int \frac{du}{u^2 - 1} = \frac{1}{2}\int\left(\frac{1}{u-1} - \frac{1}{u+1}\right)du = \frac{1}{2}\left(\ln|u-1| - \ln|u+1|\right) + C$$

$$= \frac{1}{2}\ln\left|\frac{u-1}{u+1}\right| + C = \frac{1}{2}\ln\left|\frac{e^x-1}{e^x+1}\right| + C.$$

2. $\int \frac{\sec^2 x\,dx}{\sqrt{5 - \sec^2 x}}$

Since this integral involves a function of sec x, our first reaction is: we need the term (sec x tan x dx). We can multiply numerator and denominator by tan x to obtain

$$\int \frac{(\sec x)(\sec x \tan x\,dx)}{(\tan x)\sqrt{5-\sec^2 x}},$$

but this looks very nasty. Instead, we can ask: We have the term $(\sec^2 x\,dx)$ in the numerator. Can the rest of the integral be expressed in terms of tan x? Yes, since $\sec^2 x = \tan^2 x + 1$. Using this in the denominator, we obtain

$$\int \frac{\sec^2 x\,dx}{\sqrt{4 - \tan^2 x}} = \int \frac{du}{\sqrt{4-u^2}} \qquad [u = \tan x].$$

$$= \sin^{-1}\left(\frac{u}{2}\right) + C = \sin^{-1}\left(\frac{\tan x}{2}\right) + C.$$

3. $\int \frac{dx}{x(ax^n+b)}$

One way to handle this problem might be a "desperation" substitution, $u = (ax^n+b)$. Another way is to focus on the term causing diffuculty, the x^n in the denominator. To make a substitution like $u = x^n$, *we would need nx^{n-1} in the numerator.* We can get it, if we multiply numerator and denominator by nx^{n-1}. The integral becomes

$$\int \frac{nx^{n-1}\,dx}{(nx^{n-1})(x)(ax^n+b)} = \frac{1}{n}\int \frac{nx^{n-1}\,dx}{(x^n)(ax^n+b)} = \frac{1}{n}\int \frac{du}{(u)(au+b)},$$

where $u = x^n$. We can now solve the problem by partial fractions, obtaining

$$\frac{1}{n}\int\left(\frac{(1/b)}{u} - \frac{(a/b)}{au+b}\right)du = \frac{1}{nb}\int\left(\frac{1}{u} - \frac{a}{au+b}\right)du$$

ERIC

$$= \frac{1}{nb}\left(\ln|u| - \ln|au+b|\right) + C = \frac{1}{nb}\ln\left|\frac{u}{au+b}\right| + C$$

$$= \frac{1}{nb}\ln\left|\frac{x^n}{ax^n+b}\right| + C.$$

4. $$\int \frac{dx}{(\sin x + 6)(\cos x)}$$

Since this integral involves sin x and cos x, we need either

(a) (sin x dx) in the numerator, with all the rest expressed in terms of cos x, or

(b) (cos x dx) in the numerator, with all the rest expressed in terms of sin x.

If we try (a), we obtain

$$\int \frac{\sin x\, dx}{(\sin^2 x + 6\sin x)(\cos x)} .$$

That doesn't help, because we can't express the denominator easily in terms of cos x. So we try (b):

$$\int \frac{dx}{(\sin x + 6)(\cos x)} = \int \frac{\cos x\, dx}{(\sin x + 6)(\cos^2 x)} = \int \frac{\cos x\, dx}{(\sin x + 6)(1-\sin^2 x)}.$$

Here the numerator is (cos x dx) and the denominator is a function of sin x. *Now* the substitution $u = \sin x$ gives us

$$\int \frac{du}{(u+6)(1-u^2)} = \int \frac{du}{(u+6)(1+u)(1-u)} =$$

$$\int \left(\frac{-1/35}{u+6} + \frac{1/10}{1+u} + \frac{1/14}{1-u}\right) du$$

$$= \frac{-1}{35}\ln|u+6| + \frac{1}{10}\ln|1+u| - \frac{1}{14}\ln|1-u| + C$$

$$= \frac{-1}{35}\ln|\sin x + 6| + \frac{1}{10}\ln|1+\sin x| - \frac{1}{14}\ln|1-\sin x| + C.$$

Whew!

EXERCISES FOR CHAPTER 3

PART 1: *Examine each of these integrals and DECIDE how you would solve it. Then compare your chosen approach with mine, which is given in the solutions manual.*

1. (sol. 5) $\int \frac{x}{x^4-3x^2+2}\,dx$

2. (sol. 10) $\int \frac{\tan x}{\sec x + 2}\,dx$

3. (sol. 3) $\int \frac{x^{2/3}}{x+1}\,dx$

4. (sol. 9) $\int \sqrt{\frac{x}{x+1}}\,dx$

5. (sol. 4) $\int \frac{x}{\sqrt{1+x} + \sqrt{1-x}}\,dx$

6. (sol. 7) $\int \frac{x^5}{\sqrt{1+x^3}}\,dx$

7. (sol. 8) $\int \frac{1}{(x+4)\sqrt{x^2+8x}}\,dx$

8. (sol. 1) $\int \frac{1}{\sqrt{1+\sqrt{x}}}\,dx$

9. (sol. 2) $\int x^{1/2}(1+x^{1/3})\,dx$

10. (sol. 6) $\int \frac{1}{\sec x + \tan x}\,dx$

PART 2: *Solve each of the exercises given above. Detailed solutions are in the solutions manual. The solution numbers are given below.*

Exercise #	1	2	3	4	5	6	7	8	9	10
Solution #	15	20	13	19	14	17	18	11	12	16

Appendix I

Pre-Test

You should be able to do all of these problems without difficulty. If you have a lot of trouble, practice these types of problems before you try to work through the booklet. Answers are on the opposite page.

(1) $\frac{d}{dx}(5x^2+3x^{1/2}+1)$	(8) $\int(x^2+1)^2dx$
(2) $\frac{d}{dx}(x^2+1)^3$	(9) $\int x\ e^{x^2}\ dx$
(3) $\frac{d}{dx}(\sqrt{\sin x})$	(10) $\int \frac{2x\ dx}{x^2+6}$
(4) $\frac{d}{dx}(e^{7x^2+1})$	(11) $\int(2x+1)(x^2+x)^7\ dx$
(5) $\frac{d}{dx}(\tan 6x)$	(12) $\int \csc^2 2x\ dx$
(6) $\frac{d}{dx}(\log(\sin 2x))$	(13) $\int \cos 5x\ dx$
(7) $\frac{d}{dx}(\tan^{-1}4x)$	(14) $\int \sec x \tan x\ dx$

Appendix

Answers to Pre-Test

(1) $10x + \frac{3}{2}x^{-1/2}$	(8) $\frac{1}{5}x^5 + \frac{2}{3}x^3 + x + C$; NOT $\frac{1}{3}(x^2+1)^3 + C$.
(2) $6x(x^2+1)^2$	(9) $\frac{1}{2}e^{x^2} + C$
(3) $\frac{\cos x}{2\sqrt{\sin x}}$	(10) $\ln\lvert x^2+6\rvert + C$
(4) $14x\ e^{7x^2+1}$	(11) $\frac{1}{8}(x^2+x)^8 + C$
(5) $6\sec^2 6x$	(12) $\frac{-1}{2}\cot 2x + C$
(6) $\frac{2\cos 2x}{\sin 2x}$	(13) $\frac{1}{5}\sin 5x + C$
(7) $\frac{4}{1+16x^2}$	(14) $\sec x + C$

This table contains the formulas which are ESSENTIAL for integration. You should know them so well that you never have to refer to the table when solving problems.

ESSENTIAL FORMULAS

Trigonometry

(a) $\sin^2 x + \cos^2 x = 1$

(b) $\sin 2x = 2 \sin x \cos x$

(c) $\cos 2x = \begin{cases} \cos^2 x - \sin^2 x \\ 2\cos^2 x - 1 \\ 1 - 2\sin^2 x \end{cases}$

Integration

(1) $\int u^n du = \frac{u^{n+1}}{n+1} + C$

(2) $\int \frac{du}{u} = \ln u + C$

(3) $\int e^u du = e^u + C$

(4) $\int \sin u \, du = -\cos u + C$

(5) $\int \cos u \, du = \sin u + C$

(6) $\int \sec^2 u \, du = \tan u + C$

(7) $\int \csc^2 u \, du = -\cot u + C$

(8) $\int \sec u \tan u \, du = \sec u + C$

(9) $\int \csc u \cot u \, du = -\csc u + C$

This table contains the formulas which are USEFUL for integration. For short-term use (on tests, for example) memorizing them will save you time and trouble. For long-term or occasional use, you can look them up or derive them when you need them.

USEFUL FORMULAS

Trigonometry

(d) $\tan^2 x + 1 = \sec^2 x$

(e) $1 + \cot^2 x = \csc^2 x$

(f) $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$

(g) $\cos^2 x = \frac{1}{2}(1 + \cos 2x)$

Integration

(10) $\int \frac{du}{\sqrt{a^2 - u^2}} = \sin^{-1}\left(\frac{u}{a}\right) + C$

(11) $\int \frac{du}{a^2 + u^2} = \frac{1}{a} \tan^{-1}\left(\frac{u}{a}\right) + C$

(12) $\int \frac{du}{u\sqrt{u^2 - a^2}} = \frac{1}{a} \sec^{-1}\left(\frac{u}{a}\right) + C$

(13) $\int \tan u \, du = -\ln|\cos u| + C$

(14) $\int \cot u \, du = \ln|\sin u| + C$

(15) $\int \sec u \, du = \ln|\sec u + \tan u| + C$

(16) $\int \csc u \, du = -\ln|\csc u + \cot u| + C$

SIMPLIFY!

EASY ALGEBRAIC MANIPULATIONS

(1) Break integrals into *Sums*.

(2) Exploit *Identities*.

(3) Reduce rational functions to *Proper Fractions* by division.

OBVIOUS SUBSTITUTIONS

(1) Substitute for the *"Inside Terms"* in complex expressions.

(2) Try to substitute for *"Nasty" Terms or Denominators* (brief tries only).

CLASSIFY!

INTEGRATING RATIONAL FUNCTIONS

(1) Reduce to "proper fractions" by division.

(2) Factor the denominator.

(3) Decompose by *partial fractions* into a sum of "basic" rational functions.

(4) If the denominator is $(ax+b)$ or $(ax+b)^n$, use the substitution $u = (ax+b)$.

(5) If a quadratic denominator does not factor easily, complete the square. For the terms

i: (a^2+u^2), integrate directly to obtain a logarithm and/or arctangent.

ii: $\pm(u^2-a^2)$, break into a sum and use the formula on p.17, or use partial fractions.

INTEGRATING PRODUCTS

Consider *integration by parts*. The formula is

$$\int u\,dv = uv - \int v\,du ,$$

and your choice of u and dv should be governed by two things:

(1) You must be able to integrate the term dv.

(2) You want $\int v\,du$ to be easier than the original integral. This often happens when u is simplified by differentiation.

INTEGRATING TRIGONOMETRIC FUNCTIONS

(1) Exploit *twin pairs* to prepare for substitutions. Try to obtain integrals of the form

$$\int f(\sin x)(\cos x\,dx), \text{ etc.}$$

(2) Use half-angle formulas or integration by parts to reduce powers of trigonometric functions in the integrand.

(3) As a last resort, the substitution $u = \tan(\frac{x}{2})$ transforms rational functions of sin x and cos x to rational functions of u. (see p.32)

INTEGRATING SPECIAL FUNCTIONS

(1) *If the integrand includes terms of the form*

$$(a^2-u^2)^{n/2},\ (u^2-a^2)^{n/2},\text{ or } (a^2+u^2)^{n/2},$$

(a) Draw a right triangle

(b) Place a and u so that the third side of the triangle is the term you want.

(c) "Read" the substitutions from the triangle.

(2) If the integrand *is a rational function of* e^x, make the substitutions

$$e^x = u \quad\text{and}\quad dx = \frac{1}{u}\,du.$$

(3) If the integrand is *a rational function of* x *and* $\sqrt[n]{ax+b}$, make the substitutions

$$\sqrt[n]{ax+b} = u;\quad x = \frac{1}{a}(u^n-b),\text{ and } dx = \frac{n}{a}u^{n-1}\,du.$$

MODIFY!

PROBLEM SIMILARITIES

(1) Look for easy problems similar to the one you are working on.

(2) Try to reduce the difficult problem to the form of the easy similar problem.

(3) Try the techniques you would use on the similar problem.

SPECIAL MANIPULATIONS

(1) Rationalizing denominators of quotients.

(2) Special uses of trigonometric identities.

(3) "Common denominator" substitutions.

(4) "Desperation" substitutions

NEEDS ANALYSIS

(1) Look for a term, or a form of the integral, that would enable you to solve it.

(2) Try to modify the integral to produce the term or form you need.

(3) Try to introduce the term you need; compensate for it.

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.)

73

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.)

This table contains the formulas which are ESSENTIAL for integration. You should know them so well that you never have to refer to the table when solving problems.

This table contains the formulas which are USEFUL for integration. For short-term use (on tests, for example) memorizing them will save you time and trouble. For long-term or occasional use, you can look them up or derive them when you need them.

ESSENTIAL FORMULAS

Trigonometry

(a) $\sin^2 x + \cos^2 x = 1$

(b) $\sin 2x = 2 \sin x \cos x$

(c) $\cos 2x = \begin{cases} \cos^2 x - \sin^2 x \\ 2\cos^2 x - 1 \\ 1 - 2\sin^2 x \end{cases}$

Integration

(1) $\int u^n du = \frac{u^{n+1}}{n+1} + C$

(2) $\int \frac{du}{u} = \ln u + C$

(3) $\int e^u du = e^u + C$

(4) $\int \sin u\, du = -\cos u + C$

(5) $\int \cos u\, du = \sin u + C$

(6) $\int \sec^2 u\, du = \tan u + C$

(7) $\int \csc^2 u\, du = -\cot u + C$

(8) $\int \sec u \tan u\, du = \sec u + C$

(9) $\int \csc u \cot u\, du = -\csc u + C$

USEFUL FORMULAS

Trigonometry

(d) $\tan^2 x + 1 = \sec^2 x$

(e) $1 + \cot^2 x = \csc^2 x$

(f) $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$

(g) $\cos^2 x = \frac{1}{2}(1 + \cos 2x)$

Integration

(10) $\int \frac{du}{\sqrt{a^2-u^2}} = \sin^{-1}(\frac{u}{a}) + C$

(11) $\int \frac{du}{\sqrt{a^2+u^2}} = \frac{1}{a} \tan^{-1}(\frac{u}{a}) + C$

(12) $\int \frac{du}{u\sqrt{u^2-a^2}} = \frac{1}{a} \sec^{-1}(\frac{u}{a}) + C$

(13) $\int \tan u\, du = -\ln|\cos u| + C$

(14) $\int \cot u\, du = \ln|\sin u| + C$

(15) $\int \sec u\, du = \ln|\sec u + \tan u| + C$

(16) $\int \csc u\, du = -\ln|\csc u + \cot u| + C$

umap

Units 203, 204, 205

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

Alan H. Schoenfeld

INTEGRATION:

Getting It All Together

Solutions Manual

June 1977

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

ED 036 478

1. (Exercise 4) Part (a) can be solved easily.

There are no easy algebraic manipulations in either part of the problem, so we look for substitutions. In both (a) and (b), the "nasty" term is $\tan^{-1}x$. If we try

$$u = \tan^{-1}x, \quad \text{then} \quad du = \frac{1}{x^2+1}\,dx,$$

and this term does appear in (a). Using this substitution in part (a), we obtain

$$\int \frac{\tan^{-1}x\,dx}{x^2+1} = \int (\tan^{-1}x)\left(\frac{1}{x^2+1}\,dx\right) = \int u\,du =$$

$$\frac{1}{2}u^2 + C = \frac{1}{2}(\tan^{-1}x)^2 + C.$$

2. (Exercise 2) Part (b) can be solved easily.

We begin by looking for algebraic simplifications. Both integrals (a) and (b) can be broken into sums, but that doesn't look terribly promising at this point. There are no identities which apply to either problem. But we notice that part (b) is an "improper fraction", so we should divide to reduce it to a proper fraction. The quotient is (x^2), and the remainder is (1). Now (b) is easy to finish:

$$\int \frac{x^3 + x^2 + 1}{x+1}\,dx = \int \left(x^2 + \frac{1}{x+1}\right)dx$$

$$= \frac{1}{3}x^3 + \ln|x+1| + C.$$

3. (Exercise 5) Part (a) can be solved easily.

In this problem, a simple algebraic manipulation is all we need. In part (a), $\ln(e^x) = x$, so

$$\int \ln(e^x)\,dx = \int x\,dx = \frac{1}{2}x^2 + C$$

4. (Exercise 6) Part (a) can be solved easily.

Since both parts of this exercise involve a 5th power of a rather nasty term, algebraic simplifications are out of the question. Both because it is "nasty" and an "inside" function, the term $(1 + \sqrt{x})$ commands our attention. If we try

$$u = (1 + \sqrt{x}), \quad \text{then} \quad du = \frac{1}{2\sqrt{x}}\,dx.$$

The term $\sqrt{x}$ appears in the denominator in part (a) of the exercise, so part (a) looks promising. We obtain

$$\int \frac{dx}{(\sqrt{x})(1+\sqrt{x})^5} = \int \frac{1}{(1+\sqrt{x})^5}\left(\frac{1}{\sqrt{x}}\,dx\right) = 2\int \frac{1}{(1+\sqrt{x})^5}\left(\frac{1}{2\sqrt{x}}\,dx\right) =$$

$$2\int u^{-5}\,du = \frac{2}{-4}u^{-4} + C$$

$$= \frac{-1}{2(1+\sqrt{x})^4} + C$$

RECEIVED
JAN 21 1977
Project CALC

5. (Exercise 1) Part (b) can be solved easily.

We might try exploring with trig identities in the hope of simplifying either part of this exercise. If we do, a short amount of exploration convinces us that this approach is unpromising. In both parts of this problem the "nasty" term is the denomintor, (2+sin x). If we try

$$u = (2+\sin x), \quad \text{then} \quad du = \cos x\, dx.$$

Since (du) is the numerator in part (b), (b) is the easier problem to solve. We get

$$\int \frac{\cos x\, dx}{2+\sin x} = \int \frac{du}{u} = \ln|u| + C$$

$$= \ln|2+\sin x| + C$$

6. (Exercise 7) Part (b) can be solved easily.

In both parts of this example we have that the integrand is a rational function of e^x. While we might be tempted to jump into the substitution $u = e^x$, let's follow the procedure. The first of our algebraic manipulations calls for breaking an integral of a sum into a sum of integrals. If we examine (b), we see that this almost finishes the problem. We obtain

$$\int \frac{e^{5x}+1}{e^x}\,dx = \int\left(\frac{e^{5x}}{e^x} + \frac{1}{e^x}\right)dx = \int\left(e^{4x} + e^{-x}\right)dx$$

$$= \frac{1}{4}e^{4x} - e^{-x} + C.$$

7. (Exercise 8) Part (b) can be solved easily.

As in Sample Problem 6, the moral here is: look before you leap! We can factor the denominator into (x-1)(x-3), which means that both parts of Exercise 8 can be solved by the technique of Partial Fractions. In both (a) and (b), however, the "nasty" term is the denominator, (x^2-4x+3). If we try

$$u = x^2-4x+3, \quad \text{then} \quad du = (2x-4)dx,$$

which is twice the numerator of part (b). Part (b) can then be solved almost immediately:

$$\int \frac{(x-2)dx}{x^2-4x+3} = \frac{1}{2}\int\frac{(2x-4)\,dx}{x^2-4x+3} = \frac{1}{2}\int\frac{du}{u} = \frac{1}{2}\ln|u| + C$$

$$= \frac{1}{2}\ln|x^2-4x+3| + C$$

8. (Exercise 3) Part (b) can be solved easily.

We might try to exploit the relationship $\tan^2 x + 1 = \sec^2 x$ in either part of this exercise, but manipulations with this may get complicated. We should hold off using this until we have checked for anything easier. In part (a) we have tan x as the "inside" function, which suggests

$$u = \tan x; \quad du = \sec^2 x.$$

Unfortunately, we don't have $\sec^2 x$ in part (a), or any easy way of getting it. So we go on to part (b). There the "inside" function is sec x, suggesting

$$u = \sec x; \quad du = \sec x \tan x\, dx.$$

At first this doesn't look helpful either, until we realize that we can "borrow" a (sec x) from $(\sec^4 x) = (\sec^3 x)(\sec x)$. Then

$$\int \sec^4 x \tan x\, dx = \int [\sec^3 x][\sec x \tan x\, dx]$$

$$= \int u^3\, du = \frac{1}{4}u^4 + C = \frac{1}{4}\sec^4 x + C$$

1. (Exercise 4) $\int \frac{6x^2\, dx}{\sqrt{3x+1}}$

The SIMPLIFYing techniques of Chapter 1 don't seem to help here. Our clue to approaching the problem is the term $\sqrt{3x+1}$ in the numerator. This is one of the special functions we studied in section 4, $\sqrt[n]{ax+b}$, and suggests the substitutions

$$u = \sqrt{3x+1}\ ; \quad u^2 = 3x+1\ ; \quad x = \tfrac{1}{3}(u^2-1)\ ; \quad dx = \tfrac{2}{3}u\, du\ .$$

2. (Exercise 10) $\int \frac{9\, dx}{2+\cos x}$

The methods of Chapter 1 don't seem to apply. The integrand is a combination of trig functions, so we check for the appropriate technique there. There doesn't seem to be any way to exploit "twin pairs", and there are no powers to reduce, so we are left with the "last resort" substitutions based on $u = \tan(\frac{x}{2})$.

3. (Exercise 15) $\int (\sin^2 x - \cos^2 x)\, dx$

We could use the techniques of the Trigonometric Functions section, but we should check for easy alternatives first. If we remember the trigonometric identity,

$$\cos 2x = \cos^2 x - \sin^2 x\ ,$$

the problem can be done easily by the methods of Chapter 1.

4. (Exercise 3) $\int \frac{4\, dx}{e^x+1}$

There are no apparent simplifications, and the term e^x in the denominator indicates that we should consider the substitutions $u = e^x$; $du = e^x\, dx$; $dx = \frac{du}{u}$.

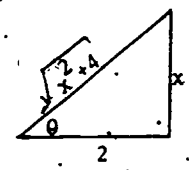
ERIC

5. (Exercise 13) $\int \frac{\tan^{-1}x \cdot dx}{x^2+1}$

The methods of Chapter 1 apply here. The "nasty" term is $\tan^{-1}x$; and if we set $u = \tan^{-1}x$, then $du = \frac{dx}{x^2+1}$. From this point on the problem is easy.

6. (Exercise 1) $\int \frac{7\, dx}{x\sqrt{x^2+4}}$

There are no apparent simplifications for this problem. Our clue for approaching it is the term $\sqrt{x^2+4}$, which is one of the special forms we studied in section 4. It suggests trig substitutions based on the triangle to the right.



7. (Exercise 12) $\int \frac{(x^3+x^2)\, dx}{x^2+x-2}$

The integrand in this exercise is a rational function, so we should follow the procedure for rational functions.

8. (Exercise 6) $\int x \tan^{-1}x\, dx$

There are no apparent simplifications. Here the integrand is a product of dissimilar functions, so integration by parts is a likely technique. The two choices we have are

(a): $u = x$; $dv = \tan^{-1}x\, dx$.

(b): $u = \tan^{-1}x$; $dv = x\, dx$.

Choice (a) doesn't look promising, because we would have to integrate $dv = \tan^{-1}x\, dx$. In choice (b), we <u>differentiate</u> $u = \tan^{-1}x$ to obtain $du = \frac{dx}{x^2+1}$, which is much simpler. So we use integration by parts,

with $u = \tan^{-1}x$; $dv = x\, dx$.

9. (Exercise 9) $\int \frac{\cos(\ln x)\, dx}{x}$

Since this integrand contains an "inside" function, $(\ln x)$, our first approach should be to try the substitution $u = \ln x$.

10. (Exercise 2) $\int 2 \tan^4 x\, dx$

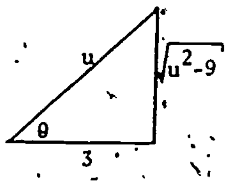
There are no apparent simplifications. Since the problem involves trigonometric functions, we should first try to exploit the relationship between tan x and its "twin", sec x. If that fails, we might look for a reduction formula.

11. (Exercise 8) $\int \frac{5\, dx}{\sqrt{x^2+6x}}$

As a preliminary simplification, we might factor the term in the denominator to obtain

$$\int \frac{5\, dx}{\sqrt{(x)(x+6)}}$$

but this doesn't seem to help much. What can we integrate? Terms of the form $\int \frac{du}{\sqrt{u^2 \pm a^2}}$; so we should consider completing the square in the denominator to obtain $\int \frac{5\, dx}{\sqrt{(x+3)^2 - 9}}$. With $u = (x+3)$, this is $\int \frac{5\, du}{\sqrt{u^2-3^2}}$, and a trig substitution is suggested, with the help of the diagram given above.



12. (Exercise 16) $\int \frac{x^4\, dx}{x^3-1}$

Since the integrand here is a rational function, we should follow the procedure given in section 1 for integrating rational functions.

13. (Exercise 5) $\int \frac{9\, x\, dx}{\sqrt{x^2+4}}$

While the denominator suggests a trig substitution, we should be careful and check for simplifications first. Since the integrand contains an "inside" function, (x^2+4), we can try $u = x^2+4$ and see what happens: $du = 2x\, dx$, and we're in luck. The problem can be done by the means of Chapter 1.

14. (Exercise 7) $\int \frac{5x^3\, dx}{x^4-1}$

The integrand is a rational function, but we shouldn't rush into the techniques of Chapter 2 until we've checked for simplifications. The "nasty" term is the denominator, and if we try $u = x^4-1$, then $du = 4x^3\, dx$. Since our numerator is $(5x^3 dx)$, we can finish the problem easily with the techniques of Chapter 1.

15. (Exercise 14) $\int \csc^2 x \cot^3 x\, dx$

This problem can be done directly by the means of Chapter 1. If we set $u = \cot x$, then $du = -\csc^2 x\, dx$, and the integral becomes $-\int u^3 du$.

16. (Exercise 11) $\int \csc^3 x \cot^3 x\, dx$

This problem can't be done immediately by the means of Chapter 1. We have to exploit "twin pairs", as in Sample Problem 10.

17. (Exercise 4) $\int \frac{6x^2\,dx}{\sqrt{3x+1}}$

See solution [1] for our reasoning. With the substitutions $u = \sqrt{3x+1}$; $u^2 = 3x+1$; $x = \frac{1}{3}(u^2-1)$; $dx = \frac{2}{3}u\,du$, the integral becomes

$$\int \frac{6\,[\frac{1}{3}(u^2-1)]^2[\frac{2}{3}u\,du]}{u} = \frac{4}{9}\int (u^2-1)^2\,du$$

$$= \frac{4}{9}\int (u^4-2u^2+1)\,du = \frac{4}{9}\,[\frac{1}{5}u^5 - \frac{2}{3}u^3 + u] + C$$

$$= \frac{4}{9}\,[\frac{1}{5}(3x+1)^{5/2} - \frac{2}{3}(3x+1)^{3/2} + (3x+1)^{1/2} + C$$

18. (Exercise 10) $\int \frac{9\,dx}{2 + \cos x}$

See solution [2] for our reasoning. With the substitutions $u = \tan(\frac{x}{2})$; $\sin x = \frac{2u}{1+u^2}$; $\cos x = \frac{1-u^2}{1+u^2}$; $dx = \frac{2\,du}{1+u^2}$, the integral becomes

$$\int \frac{9\left(\frac{2\,du}{1+u^2}\right)}{2 + \left(\frac{1-u^2}{1+u^2}\right)} = 18\int \frac{du}{u^2+3}$$

Now if we remember formula 11 from the table of useful integrals, this is

$$18\,[\frac{1}{\sqrt{3}}\tan^{-1}(\frac{u}{\sqrt{3}}) + C$$

$$= \frac{18}{\sqrt{3}}\tan^{-1}\left(\frac{\tan(\frac{x}{2})}{\sqrt{3}}\right) + C.$$

19. (Exercise 15) $\int (\sin^2 x - \cos^2 x)\,dx$

See solution [3] for our reasoning. We have

$$\int (\sin^2 x - \cos^2 x)\,dx = -\int \cos 2x\,dx = -\frac{1}{2}\sin 2x + C.$$

20. (Exercise 3) $\int \frac{4\,dx}{e^x-1}$

See solution [4] for our reasoning. With the substitutions $u = e^x$; $du = e^x\,dx$; $dx = \frac{du}{u}$, the integral becomes

$$\int \frac{4[\frac{du}{u}]}{u-1} = 4\int \frac{du}{(u)(u-1)}.$$ Using partial fractions, this is

$$4\int\left(\frac{1}{u-1} - \frac{1}{u}\right)du = 4\left(\ln|u-1| - \ln|u|\right) + C$$

$$= 4\ln\left|\frac{u-1}{u}\right| + C = 4\ln\left|\frac{e^x-1}{e^x}\right| + C.$$

21. (Exercise 13) $\int \frac{\tan^{-1}x\,dx}{x^2+1}$

See solution [5] for our reasoning. With the substitutions $u = \tan^{-1}x$; $du = \frac{dx}{x^2+1}$, the integral becomes

$$\int \tan^{-1}x\left(\frac{dx}{x^2+1}\right) = \int u\,du = \frac{1}{2}u^2 + C$$

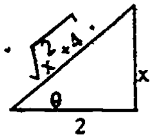
$$= \frac{1}{2}[\tan^{-1}x]^2 + C.$$

RECEIVED
JAN 24 1977
Project CALC.

ERIC

22. (Exercise 1) $\int \frac{7\,dx}{x\sqrt{x^2+4}}$

See solution [6] for our reasoning. Based on the triangle to the right, we obtain the substitutions $x = 2\tan\theta$; $dx = 2\sec^2\theta\,d\theta$; and $\sqrt{x^2+4} = 2\sec\theta$. the integral becomes



$$7\int \frac{2\sec^2\theta\,d\theta}{(2\tan\theta)(2\sec\theta)} = \frac{7}{2}\int \frac{\sec\theta\,d\theta}{\tan\theta}$$

$$= \frac{7}{2}\int \frac{[1/\cos\theta]\,d\theta}{[\sin\theta/\cos\theta]} = \frac{7}{2}\int \frac{d\theta}{\sin\theta} = \frac{7}{2}\int \csc\theta\,d\theta.$$

Using formula (16), we obtain

$$\frac{-7}{2}\ln|\csc\theta + \cot\theta| + C.$$

Going back to the triangle to "translate" $\csc\theta$ and $\cot\theta$ in terms of x, we obtain

$$\frac{-7}{2}\ln\left|\frac{\sqrt{x^2+4}}{x} + \frac{2}{x}\right| + C.$$

23. (Exercise 12) $\int \frac{(x^3+x^2)dx}{x^2+x-2}$

Following the procedure for rational functions, we obtain

$$\int \frac{(x^3+x^2)\,dx}{x^2+x-2} = \int\left(x + \frac{2x}{x^2+x-2}\right)dx$$

$$\int\left(x + \frac{2x}{(x+2)(x-1)}\right)dx = \int\left(x + \frac{(4/3)}{x+2} + \frac{(2/3)}{x-1}\right)dx$$

$$= \frac{1}{2}x^2 + \frac{4}{3}\ln|x+2| + \frac{2}{3}\ln|x-1| + C.$$

24. (Exercise 6) $\int x\tan^{-1}x\,dx$

See solution [8] for our reasoning. Using integration by parts with

$$\begin{cases} u = \tan^{-1}x & dv = x\,dx \\ du = \frac{dx}{x^2+1} & v = \frac{1}{2}x^2 \end{cases}$$, we obtain

$$\underbrace{(\tan^{-1}x)}_{u}\underbrace{(x\,dx)}_{dv} = \underbrace{(\tan^{-1}x)}_{u}\underbrace{(\tfrac{1}{2}x^2)}_{v} - \underbrace{(\tfrac{1}{2}x^2)}_{v}\underbrace{(\tfrac{dx}{x^2+1})}_{du}$$

$$= \frac{1}{2}x^2\tan^{-1}x - \frac{1}{2}\int\frac{x^2\,dx}{x^2+1} = \frac{1}{2}x^2\tan^{-1}x - \frac{1}{2}\int\left(1 - \frac{1}{x^2+1}\right)dx$$

$$= \frac{1}{2}x^2\tan^{-1}x - \frac{1}{2}x + \frac{1}{2}\tan^{-1}x + C.$$

25. (Exercise 9) $\int \frac{\cos(\ln x)\,dx}{x}$

See solution [9] for our reasoning. With the substitutions $u = \ln x$; $du = \frac{1}{x}dx$, the integral becomes

$$\int \cos(\ln x)\left[\left(\frac{1}{x}\right)dx\right] = \int \cos u\,du = \sin u + C$$

$$= \sin(\ln x) + C.$$

26. (Exercise 2) $\int 2\tan^4 x\,dx$

See solution [10] for our reasoning. Exploiting the relationships between tan x and its "twin," sec x, we obtain

$$2\int \tan^4 x\,dx = 2\int(\tan^2 x)(\sec^2 x - 1)\,dx =$$

$$2\int(\tan^2 x)(\sec^2 x)dx - 2\int\tan^2 x\,dx =$$

$$2\int(\tan^2 x)(\sec^2 x\,dx) - 2\int(\sec^2 x - 1)\,dx =$$

$$2\int(\tan^2 x)(\sec^2 x\,dx) - 2\int\sec^2 x\,dx + 2\int dx =$$

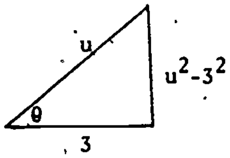
$$\frac{2}{3}\tan^3 x - 2\tan x + 2x + C.$$

ERIC Full Text Provided by ERIC

27 (exercise 8) $\int \frac{5\,dx}{\sqrt{x^2+6x}}$

See solution [11] for our reasoning. After completing the square in the denominator and making the substitution $u = x+3$, the above integral becomes

$\int \frac{5\,du}{\sqrt{u^2-3^2}}$. From the triangle to the right, we obtain the substitutions



$$u = 3\sec\theta;\quad du = 3\sec\theta\tan\theta\,d\theta;\ \sqrt{u^2-3^2} = 3\tan\theta.$$

This transforms the integral to

$$5\int \frac{3\sec\theta\tan\theta\,d\theta}{3\tan\theta} = 5\int \sec\theta\,d\theta =$$

$$5\ln|\sec\theta + \tan\theta| + C = 5\ln\left|\frac{u}{3} + \frac{\sqrt{u^2+3^2}}{3}\right| + C$$

$$= 5\ln\left|\frac{x+3}{3} + \frac{\sqrt{x^2+6x}}{3}\right| + C.$$

28 (Exercise 16) $\int \frac{x^4dx}{x^3-1}$

Following the procedure for rational functions, we obtain

$$\int \frac{x^4dx}{x^3-1} = \int\left(x + \frac{x}{x^3-1}\right)dx$$

$$= \int\left(x + \frac{x}{(x-1)(x^2+x+1)}\right)dx$$

$$= \int\left(x + \frac{(1/3)}{x-1} + \frac{(-1/3)x + (1/3)}{x^2+x+1}\right)dx$$

$$= \int x\,dx + \frac{1}{3}\int\frac{dx}{x-1} - \frac{1}{3}\int\frac{(x-1)\,dx}{(x+\frac{1}{2})^2+\frac{3}{4}}$$

We now make the substitution $u = (x+\frac{1}{2})$ in the third integral, to obtain

$$\int x\,dx + \frac{1}{3}\int\frac{dx}{x-1} - \frac{1}{3}\int\frac{(u-3/2)\,du}{u^2+3/4}$$

$$= \int x\,dx + \frac{1}{3}\int\frac{dx}{x-1} - \frac{1}{3}\int\frac{u\,du}{u^2+3/4} + \frac{1}{2}\int\frac{du}{u^2+3/4}$$

$$= \frac{1}{2}x^2 + \frac{1}{3}\ln|x-1| - \frac{1}{6}\ln\left|u^2+\frac{3}{4}\right| + \frac{1}{2}\left(\frac{1}{\sqrt{3/4}}\tan^{-1}\left(\frac{u}{\sqrt{3/4}}\right)\right) + C$$

$$= \frac{1}{2}x^2 + \frac{1}{3}\ln|x-1| - \frac{1}{6}\ln|x^2+x+1| + \frac{1}{\sqrt{3}}\tan^{-1}\left(\frac{2x+1}{\sqrt{3}}\right) + C.$$

29 (Exercise 5) $\int \frac{9x\,dx}{\sqrt{x^2+4}}$

See solution [13] for our reasoning. With the substitution $u = x^2+4$, we have

$$\int\frac{9x\,dx}{\sqrt{x^2+4}} = \frac{9}{2}\int\frac{2x\,dx}{\sqrt{x^2+4}} = \frac{9}{2}\int\frac{du}{\sqrt{u}} = \frac{9}{2}\int u^{-1/2}\,du$$

$$= 9u^{1/2} + C = 9\sqrt{x^2+4} + C.$$

30 (Exercise 7) $\int \frac{5x^3\,dx}{x^4-1}$

See solution [14] for our reasoning. With the substitution $u = (x^4-1)$, the integral is

$$\int\frac{5x^3\,dx}{x^4-1} = \frac{5}{4}\int\frac{4x^3\,dx}{x^4-1} = \frac{5}{4}\int\frac{du}{u} = \frac{5}{4}\ln|u| + C$$

$$= \frac{5}{4}\ln|x^4-1| + C.$$

ERIC Full Text Provided by ERIC

31. (Exercise 14) $\int \csc^2 x \cot^3 x \, dx.$

See solution [15] for our reasoning. With the substitutions $u = \cot x$; $du = -\csc^2 x \, dx$, the above becomes

$$-\int (\cot^3 x)(-\csc^2 x \, dx) = -\int u^3 \, du = \frac{-1}{4} u^4 + C$$

$$= \frac{-1}{4} \cot^4 x + C$$

32. (Exercise 11) $\int \csc^3 x \cot^3 x \, dx$

$$\int \csc^3 x \cot^3 x \, dx = -\int (\csc^2 x)(\cot^2 x)(-\csc x \cot x \, dx)$$

$$= -\int (\csc^2 x)(\csc^2 x - 1)(-\csc x \cot x \, dx).$$

At this point the integrand has been expressed in terms of $\csc x$ and its derivative. With the substitutions $u = \csc x$, $du = -\csc x \cot x \, dx$, we obtain

~~$-\int (u^2)(u^2-1)(du)$~~

$$= \int (-u^4 + u^2) \, du$$

$$= \frac{-1}{5} u^5 + \frac{1}{3} u^3 + C$$

$$= \frac{-1}{5} \csc^5 x + \frac{1}{3} \csc^3 x + C.$$

1. (Exercise 8) $\int \frac{dx}{\sqrt{1 + \sqrt{x}}}$

Like Sample Problem (9), this problem can be approached a number of ways. With such a nasty expression, we might be tempted to make a "desperation" substitution with

$u = \sqrt{1 + \sqrt{x}}$.

2. (Exercise 9) $\int x^{1/2}(1 + x^{1/3}) \, dx$

This is not a "common denominator" substitution problem. If we multiply the two terms in the integrand, the problem can be handled easily by the methods of Chapter 1.

3. (Exercise 3) $\int \frac{x^{2/3} \, dx}{x+1}$

This is a "common denominator" substitution problem, where the terms in the integrand are $x^{2/3}$ and x^1. The common denominator is 3, so we should make the substitution

$$u = x^{1/3}.$$

4. (Exercise 5) $\int \frac{x \, dx}{\sqrt{1+x} + \sqrt{1-x}}$

In this problem we should rationalize the denominator, and then see whatever else is called for.

5. (Exercise 1) $\int \frac{x \, dx}{x^4 - 3x^2 + 2}$

The form of this problem is similar to rational functions with quadratic denominators. We could obtain a quadratic denominator by setting $u = x^2$.

6. (Exercise 10) $\int \frac{dx}{\sec x + \tan x}$

The easiest way to handle this problem is to recall the identity: $\tan^2 x + 1 = \sec^2 x$, or $\sec^2 x - \tan^2 x = 1$. We can multiply the denominator by its conjugate, sec x - tan x.

7. (Exercise 6) $\int \frac{x^5\, dx}{\sqrt{1+x^3}}$

There might be any of a number of approaches to this problem. The key observation to make is that the numerator, $x^5 dx$, can be written as $\frac{1}{3}(x^3)(3x^2 dx)$. This makes the substitution $u = x^3$ look promising as a beginning; we can go on from there.

8. (Exercise 7) $\int \frac{dx}{(x+4)\sqrt{x^2+8x}}$

As in Sample Problem 3, we need a way to get started on this problem. Perhaps completing the square in the denominator will give us a lead.

9. (Exercise 4) $\int \sqrt{\frac{x}{x+1}}\, dx$

There doesn't seem to be any easy way to approach this problem. It might be worth trying a desperation substitution,

$$u = \sqrt{\frac{x}{x+1}}$$

10. (Exercise 2) $\int \frac{\tan x\, dx}{\sec x + 2}$

Since the integrand contains an expression involving sec x in the denominator, we can ask: what do we need to integrate such an expression? The derivative of sec x, [sec x tan x dx]. We can obtain this by multiplying both numerator and denominator by sec x.

11. (Exercise 8) $\int \frac{dx}{\sqrt{1+\sqrt{x}}}$

See solution [1] for our reasoning. With the substitution $u = \sqrt{1+\sqrt{x}}$, we have $u^2 = 1 + \sqrt{x}$; $u^2 - 1 = \sqrt{x}$; $x = (u^2-1)^2$; and $dx = 4u(u^2-1)du$. Then the integral becomes

$$\int \frac{4u(u^2-1)du}{u} = 4\int (u^2-1)\, du = \frac{4}{3}u^3 - 4u + C.$$

$$= \frac{4}{3}[1+\sqrt{x}]^{3/2} - 4[1+\sqrt{x}]^{1/2} + C.$$

12. (Exercise 9) $\int x^{1/2}(1 + x^{1/3})\, dx$

$$= \int (x^{1/2} + x^{5/6})dx = \frac{2}{3}x^{3/2} + \frac{6}{11}x^{11/6} + C.$$

13. (Exercise 3) $\int \frac{x^{2/3}\, dx}{x+1}$

See solution [3] for our reasoning. With the substitution $u = x^{1/3}$, we have $u^3 = x$ and $(3u^2\, du) = dx$. The integral becomes

$$\int \frac{(x^{1/3})^2\, dx}{x+1} = \int \frac{(u^2)(3u^2\, du)}{u^3+1} = \int \frac{3u^4\, du}{u^3+1}.$$

If we now follow the procedure for rational functions, we obtain

$$\int \left(3u - \frac{3u}{u^3+1}\right) du = \int \left(3u - \frac{3u}{(u+1)(u^2-u+1)}\right) du$$

$$= \int \left(3u - \left(\frac{-1}{u+1} + \frac{u+1}{u^2-u+1}\right)\right) du = \int \left(3u + \frac{1}{u+1} - \frac{u+1}{(u-\frac{1}{2})^2 + (3/4)}\right) du.$$

For the third integral, we set $w = (u - \frac{1}{2})$. This gives us

$$\int 3u\, du + \int \frac{du}{u+1} - \int \frac{(w+3/2)\, dw}{w^2 + (3/4)}$$

(continued...)

$$= \int 3u\,du + \int \frac{du}{u+1} - \int \frac{w\,dw}{w^2+(3/4)} - \frac{3}{2}\int \frac{dw}{w^2+(3/4)}$$

$$= \frac{3}{2}u^2 + \ln|u+1| - \frac{1}{2}\ln\left|w^2+\frac{3}{4}\right| - \frac{3}{2}\left(\frac{1}{\sqrt{3/4}}\tan^{-1}\left(\frac{w}{\sqrt{3/4}}\right)\right) + C$$

$$= \frac{3}{2}u^2 + \ln|u+1| - \frac{1}{2}\ln|u^2-u+1| - \sqrt{3}\tan^{-1}\left(\frac{2u-1}{\sqrt{3}}\right) + C$$

$$= \frac{3}{2}x^{2/3} + \ln|x^{1/3}+1| - \frac{1}{2}\ln|x^{2/3}-x^{1/3}+1| - \sqrt{3}\tan^{-1}\left(\frac{2x^{1/3}-1}{\sqrt{3}}\right) + C$$

14. (Exercise 5) $\int \frac{x\,dx}{\sqrt{1+x}+\sqrt{1-x}}$

To rationalize the denomimator in this problem, we multiply both numerator and denominator by $[\sqrt{1+x} - \sqrt{1-x}]$. This yields

$$\int \frac{[\sqrt{1+x} - \sqrt{1-x}]\,x\,dx}{[\sqrt{1+x} - \sqrt{1-x}][\sqrt{1+x} + \sqrt{1-x}]} = \int \frac{[\sqrt{1+x} - \sqrt{1-x}]\,x\,dx}{(1+x)-(1-x)}$$

$$= \int \frac{[\sqrt{1+x} - \sqrt{1-x}]\,x\,dx}{2x} = \frac{1}{2}\int [\sqrt{1+x} - \sqrt{1-x}]\,dx$$

$$= \frac{1}{3}\left((1+x)^{3/2} + (1-x)^{3/2}\right) + C.$$

15. (Exercise 1) $\int \frac{x\,dx}{x^4-3x^2+2}$

See solution [5] for our reasoning. With the substitutions $u = x^2$, $du = 2x\,dx$, this integral becomes

$$\frac{1}{2}\int \frac{2x\,dx}{(x^2)^2-3(x^2)+2} = \frac{1}{2}\int \frac{du}{u^2-3u+2} = \frac{1}{2}\int \frac{du}{(u-2)(u-1)}$$

$$= \frac{1}{2}\int \left(\frac{1}{u-2} - \frac{1}{u-1}\right)du = \frac{1}{2}\left(\ln|u-2| - \ln|u-1|\right) + C$$

$$= \frac{1}{2}\ln\left|\frac{u-2}{u-1}\right| + C = \frac{1}{2}\ln\left|\frac{x^2-2}{x^2-1}\right| + C.$$

16. (Exercise 10) $\int \frac{dx}{\sec x + \tan x}$

See solution [6] for our reasoning. Multiplying numerator and denominator of the above integral by $[\sec x - \tan x]$, we obtain

$$\int \frac{[\sec x - \tan x]\,dx}{[\sec x - \tan x][\sec x + \tan x]} = \int \frac{[\sec x - \tan x]\,dx}{\sec^2 x - \tan^2 x}$$

$$= \int \frac{[\sec x - \tan x]\,dx}{1} = \int [\sec x - \tan x]\,dx$$

$$= \ln|\sec x + \tan x| + \ln|\cos x| + C$$

NOTE: As usual, there is more than one way to approach this problem. If we don't notice that we can multiply by the conjugate of the denominator, or if we feel uncomfortable with sec x and tan x, we can express the integrand in terms of sin x and cos x. This gives us

$$\int \frac{dx}{\frac{1}{\cos x} + \frac{\sin x}{\cos x}} = \int \frac{\cos x\,dx}{\sin x + 1} = \ln|\sin x + 1| + C.$$

We can show easily that these are the same answer. In this case the second alternative gives us a faster solution than the use of conjugates. That can happen; the important thing to have is an organized, logical procedure for approaching integrals.

17. (Exercise 6) $\int \frac{x^5\,dx}{\sqrt{1+x^3}}$

See solution [7] for our reasoning. With the substitutions $u = x^3$, $du = 3x^2\,dx$, this integral is

$$\frac{1}{3}\int \frac{(x^3)(3x^2\,dx)}{\sqrt{1+x^3}} = \frac{1}{3}\int \frac{u\,du}{\sqrt{1+u}}.$$

The substitutions $v = (1+u)$, $dv = du$, reduce this to

$$\int \frac{(v-1)\,dv}{\sqrt{v}} = \int\left(\sqrt{v} - \frac{1}{\sqrt{v}}\right) dv = \int [v^{1/2} - v^{-1/2}]\, dv$$

$$= \frac{2}{3} v^{3/2} - 2\, v^{1/2} + C = \frac{2}{3}(1+u)^{3/2} - 2(1+u)^{1/2} + C$$

$$= \underline{\frac{2}{3}(1+x^3)^{3/2} - 2\,(1+x^3)^{1/2} + C}$$

18. (Exercise 7) $\int \frac{dx}{(x+4)\sqrt{x^2+8x}}$

See solution [8] for our reasoning. Completing the square in the denominator, we obtain

$$\int \frac{dx}{(x+4)\sqrt{(x+4)^2-4^2}}$$, and the substitution $u = x+4$ yields

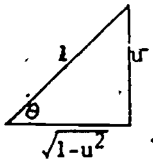
$$\int \frac{du}{u\sqrt{u^2-4^2}} = \frac{1}{4}\sec^{-1}\left(\frac{u}{4}\right) + C = \underline{\frac{1}{4}\sec^{-1}\left(\frac{x+4}{4}\right) + C.}$$

19. (Exercise 4) $\int \sqrt{\frac{x}{x+1}}\, dx$

This is the hardest problem in these materials; don't get dismayed if you had a lot of trouble! With the desperation substitution $u = \sqrt{\frac{x}{x+1}}$, we obtain $u^2 = \frac{x}{x+1}$, so $u^2(x+1) = x$. Solving for x, we obtain $x = \left(\frac{u^2}{1-u^2}\right)$, so $dx = \left(\frac{2u\,du}{(1-u^2)^2}\right)$. The integral becomes

$$\int (u)\left(\frac{2u\,du}{(1-u^2)^2}\right) = 2\int \frac{u^2\,du}{(1-u^2)^2}$$

The term in the denominator suggests the substitutions $\sqrt{1-u^2} = \cos\theta$; $u = \sin\theta$; $du = \cos\theta\, d\theta$, which we derived from the triangle to the right.



These substitutions transform the integral to

$$2\int \frac{(\sin^2\theta)(\cos\theta\, d\theta)}{\cos^4\theta} = 2\int \left(\frac{\sin\theta}{\cos\theta}\right)^2\left(\frac{1}{\cos\theta}\right) d\theta$$

$$= 2\int \tan^2\theta \sec\theta\, d\theta = 2\int (\sec^2\theta - 1)(\sec\theta)\, d\theta$$

$$= 2\int \sec^3\theta\, d\theta - 2\int \sec\theta\, d\theta.$$

Finally, we can see our way to the end of the problem. The first integral can be done by parts, the second by formula 15. We obtain

$$2\left(\frac{1}{2}[\sec\theta\tan\theta + \ln|\sec\theta + \tan\theta|]\right) - 2\left(\ln|\sec\theta + \tan\theta|\right) + C$$

$$= (\sec\theta\tan\theta) - \ln|\sec\theta + \tan\theta| + C$$

$$= \left(\frac{1}{\sqrt{1-u^2}}\right)\left(\frac{u}{\sqrt{1-u^2}}\right) - \ln\left|\frac{1}{\sqrt{1-u^2}} + \frac{u}{\sqrt{1-u^2}}\right| + C$$

$$\underline{\frac{u}{1-u^2} - \frac{1}{2}\ln\left|\frac{1+u}{1-u}\right| + C, \text{ where } u = \sqrt{\frac{x}{x+1}}.}$$

20. (Exercise 2) $\int \frac{\tan x\, dx}{\sec x + 2}$

See solution [10] for our reasoning. Multiplying numerator and denominator by (sec x) and making the substitution u = sec x, we obtain

$$\int \frac{\sec x \tan x\, dx}{(\sec x)(\sec x + 2)} = \int \frac{du}{u(u+2)}.$$

$$= \frac{1}{2}\int\left(\frac{1}{u} - \frac{1}{u+2}\right) du = \frac{1}{2}\left(\ln|u| - \ln|u+2|\right) + C = \frac{1}{2}\ln\left|\frac{u}{u+2}\right| + C$$

$$= \underline{\frac{1}{2}\ln\left|\frac{\sec x}{\sec x + 2}\right| + C.}$$

ERIC

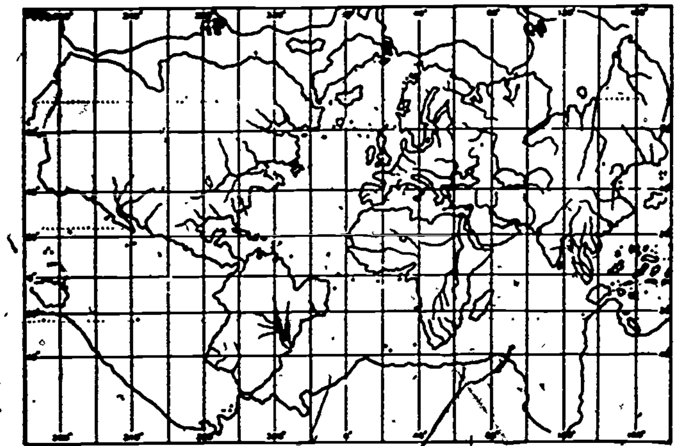
umap

UNIT 206

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT

MERCATOR'S WORLD MAP AND THE CALCULUS

by Philip M. Tuchinsky



APPLICATIONS OF CALCULUS TO GEOGRAPHY

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

ED 036 478

MERCATOR'S WORLD MAP AND THE CALCULUS

Philip M. Tuchinsky
Department of Mathematical Sciences
Ohio Wesleyan University
Delaware, Ohio 43015

6/26/78

TABLE OF CONTENTS

Intermodular Documentation Sheet: UMAP *Unit 206*

Title: MERCATOR'S WORLD MAP AND THE CALCULUS

Author: Philip M. Tuchinsky
Department of Mathematical Sciences
Ohio Wesleyan University
Delaware, Ohio 43015

Review Stage/Date: IV 6/26/78

Classification: APPL CALC/GEOGRAPHY

Suggested Support Material: None is essential. A Mercator wall map and globe are helpful. A spherical blackboard (a globe painted dull black on which chalk lines can be drawn) is an outstanding classroom aid for this module. Physics, geology, and/or geography departments often have such a globe and will lend it.

References: See Section 7 of text.

Prerequisite Skills:

For the basic application in Sections 1, 2, 3:
Definition of the trigonometric relationships in a triangle.
The identity $\sin^2 x + \cos^2 x = 1$.
Recognition of integral sums and the Riemann integrals they approach.
Definition of the natural logarithm function, basic knowledge of latitude and longitude.

For Section 4 add:
Integration of x^{-1} to $\ln |x|$.
Partial fractions integration.
Change of variables in integration.
Derivatives of the trigonometric functions.
Trigonometric relations like $\sin(\frac{\pi}{2} - x) = \cos x$.
Double angle formulas from trigonometry.
$\ln(ab) = \ln(a) + \ln(b)$.

For Section 5 add:
Convergence and sum of geometric series.
Integration of $\int \cos x \sin^4 x\, dx$.

For the exercises add:
Chain rule.
Difficult trigonometric identity work (you might give hints).

Output Skills: From Sections 1-3:

Describe an application which
a) caused $\int \sec x\, dx$ to be calculated before calculus was invented.
b) leads to an approximation sum for $\int \sec x\, dx$.
Sketch the basic frame work of a Mercator map.
Show a rhumb line path between any two cities on a Mercator map.
Describe the mathematical principles that make the Mercator map useful to sailors.

Discuss the advantages and shortcomings of the Mercator projection.
Discuss the historical need for and development of the Mercator map and $\int \sec x\, dx$ as interdependent problems.

From Section 4:
Three calculations of $\int \sec x\, dx$.
Explain why $\sqrt{x^2} = |x|$, not x.
Integrate using both radians and degrees.
Confidently use the easier trigonometric identities.

From Section 5:
Approximate $\int \sec x\, dx$, $\int \tan x\, dx$ and arctan x numerically.
Integrate a series term by term.

From Boxes:
Briefly discuss the achievements of Gerhardus Mercator, James Gregory, and John Wallis.

Suggested Uses: The unit can be done all at once or in several pieces. Section 1-3 plus 4.1 and exercises 1b, 2, 5, 6 are appropriate as soon as $\int \sec x\, dx$ is discussed. Section 4.2, 4.3 and exercises 1, 3, 7, 8 call for more knowledge of integration. Section 5 and exercises 9-13 require knowledge of the series portion of calculus and might be done much later than Sections 1-4. The unit is also appropriate for independent reading by honors students and for seminar presentation by advanced undergraduates.

Other Related Units:

UMAP editor for this module: William U. Walton

The author is indebted to V. Frederick Rickey of the Mathematics Department at Bowling Green State University in Ohio for much of the material in this paper. His first introduction to this application was through Professor Rickey's presentation "An Application of Geography to Mathematics: $\int \sec\theta\, d\theta$ and its History" at the May, 1975 meeting of the Ohio Section of the Mathematical Association of America, and he has generously helped the author with source material.

© 1978 EDC/Project UMAP
All Rights Reserved

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
Paula M. Santillo	Administrative Assistant
Zachary Zevitas	Staff Assistant

NATIONAL STEERING COMMITTEE

W.T. Martin	M.I.T. (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

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

MERCATOR'S WORLD MAP AND THE CALCULUS

Philip Tuchinsky
Department of Mathematical Science
Ohio Wesleyan University
Delaware, Ohio 43015

6/26/78

1. MERCATOR'S ACHIEVEMENT

1.1 A Strategy for Navigation with Map and Compass

Imagine yourself piloting a ship at sea--how can you reliably get to your destination? Suppose you have brought the most basic of navigational aids: a magnetic compass and good maps. The simplest way to use your compass would be to hold its needle still by keeping your ship moving in a constant compass direction. Thus, if you travel steadily northeast, your compass needle (which points north)[1] will make a steady 45° angle with your direction of motion and the needle will stay still.

Figures 1-5 show such a northeastward journey (an airflight from the Galapagos Islands in the Pacific Ocean to Franz Josef Land in the Artic) as it would appear on five types of map. The airplane's course makes a 45° angle with all the meridians (the north-south lines, great circles through the north and south poles) on each map.

Which map would be the easiest one on which to lay out the course? Figure 1 may give the best overall view of the earth as a sphere, but the Mercator projection in Figure 5 is the best for navigation because your ship's

[1]In truth, the compass points to the north magnetic pole, not the North (geographic) pole. Discrepancies of this kind are discussed later in the Special Assistance Supplement. [*S-1*]

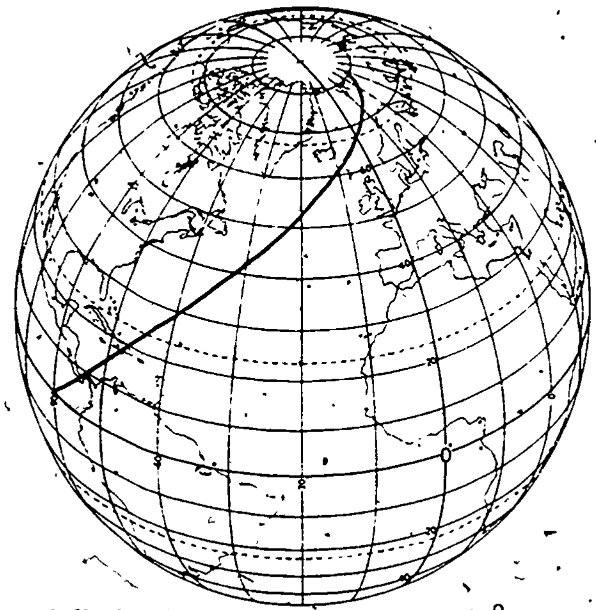


Figure 1. A flight with a constant bearing of 45° E of N from the Galapagos Islands in the Pacific to Franz Josef Land in the Arctic Ocean.

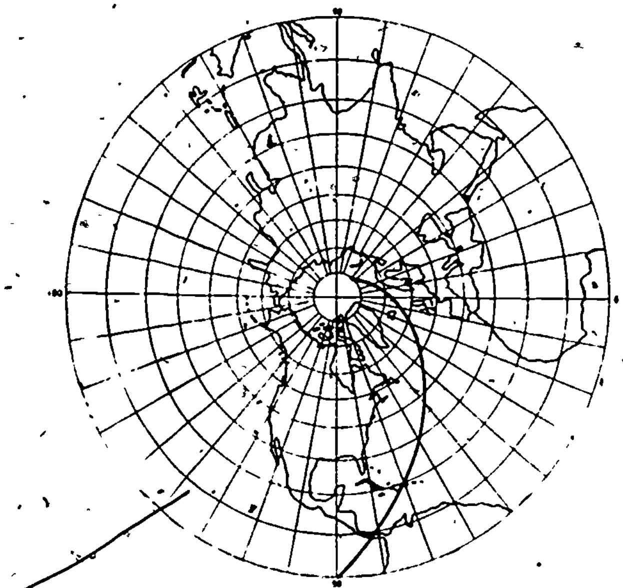


Figure 2. The flight of Figure 1 plotted on one form of conformal map. (The angles to the meridians are constant but because the meridians converge the path is curved and would be difficult to plot and measure.)

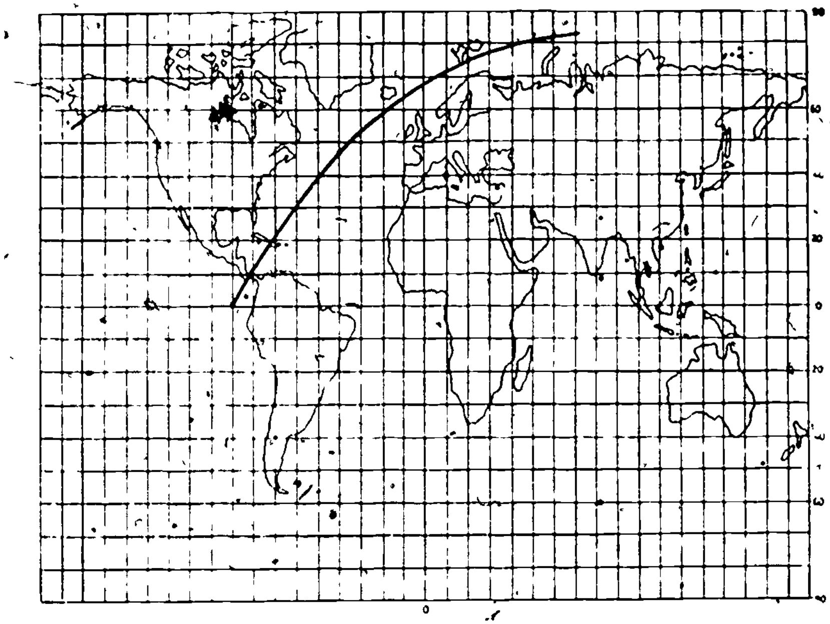


Figure 3. Plot of flight on "plane chart" such as was in use for charts of small areas in Mercator's time. Angles are not true and a straight line would not give a path of constant bearing.

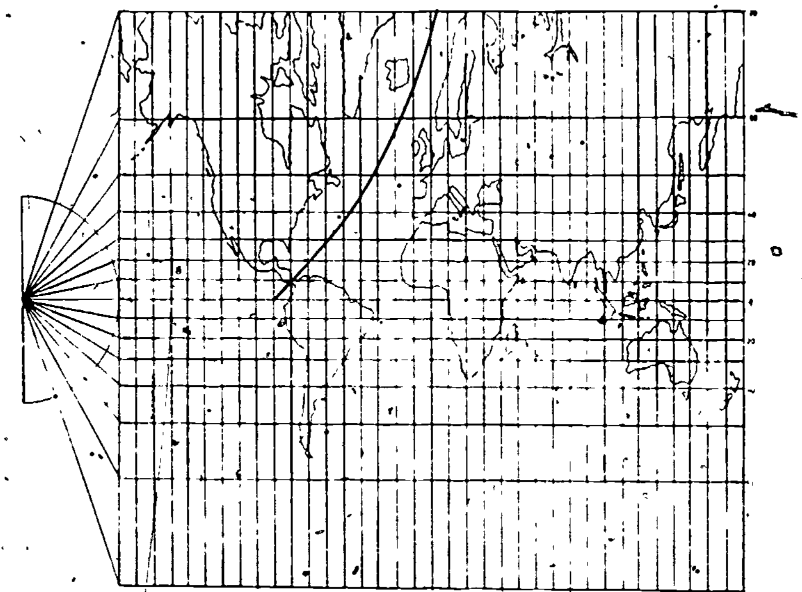


Figure 4. Flight on a cylindrical projection, a map often confused with Mercator's. Again, the path of constant bearing is not a straight line. See Exercise 5.

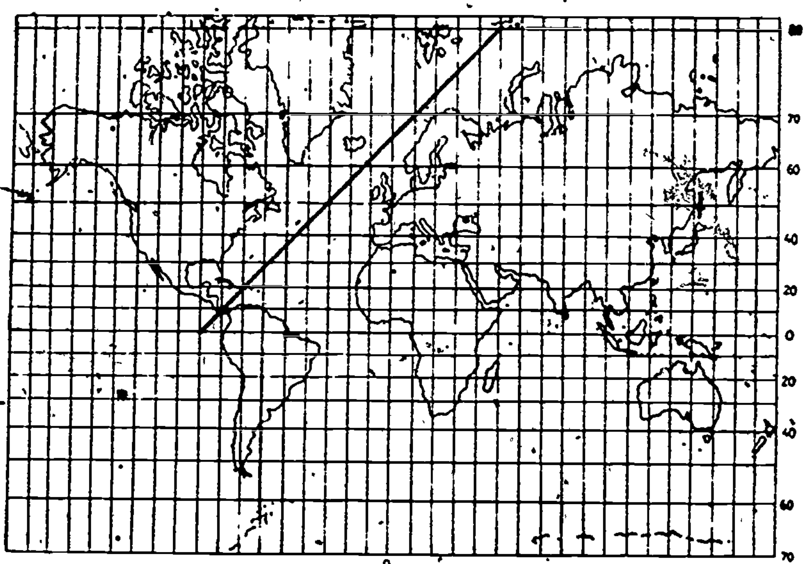


Figure 5. Flight at constant bearing on a Mercator projection. Straight line path is easily constructed, measured, and followed.

course appears there as a straight line, not a curve. It's easy to construct the course with a protractor (compass rose) and straight edge because it is a straight line course.

1.2 Rhumb Lines

Sailors have used the compass and followed lines of constant compass direction since at least the thirteenth century.[2] They called such paths on a map or chart *rhumb lines*. Cartographers and mathematicians found that the sailors' rhumb lines became spiral-like curves on the globe and named them "loxodromic curves" or *loxodromes* (from the Latin loxos--slant and drome--running) because they cut all the meridians they cross at the same slant angle (See Figure 1). You should trace a loxodrome on a globe to see that it spirals. As a rhumb line moves north and the meridians get closer together, the line must turn steadily toward the pole to cut all the meridians at the same angle. It spirals toward the pole without ever reaching it.

[2] Historians disagree as to the origins of the magnetic compass. You will find an interesting account of the compass and its history under "compass" in the *Encyclopedia Brittanica*.

1.3 The Need for a Map On Which Rhumb Lines Are Straight

If you wish to follow a rhumb line course, you must know what constant compass direction to use from your starting point S to your destination D. If you had a map on which the rhumb line path between any points S and D was simply the straight line between those points, you could draw that line with a ruler and read the compass direction by measuring (with a protractor) the angle at which meridians are cut. Before Mercator's time sailors attempted to use plain charts (charts in which the lines of latitude were equally spaced) for this purpose. Figures 6a and b and 7 show the error that arises when a straight line on a plain chart is assumed to be a rhumb line course.

1.4 Mercator's Successful Map

In the sixteenth century, Gerhardus Mercator recognized that such a map, on which all rhumb lines would appear as straight lines, would be very useful to sailors. He succeeded in creating such a map--his famous world map published in 1569. This map was recognized as a gigantic achievment, the first significant improvement in map design in 1400 years. A standard reference on cartography calls the Mercator projection a "radical departure and improvement over methods existing before his (Mercator's) time. In contemporary judgment he was styled as 'In cosmographia longe primus', which, translated, means: In cosmography by far the first." (Deetz and Adams, p. 104)

The Mercator world map has become such a fixture in our culture that it is familiar to every school child. I remember this map as a very unsatisfactory early view of our planet, because my teacher convinced me more of its shortcomings than of its value. The shortcomings are serious: distances are hard to measure on the map because: northern regions appear grossly exaggerated in area (compare Greenland to Africa on a globe and on a Mercator map--or in Figure 1 and Figure 5); the polar regions cannot be shown at all but must be inset as separate maps; distances are hard to

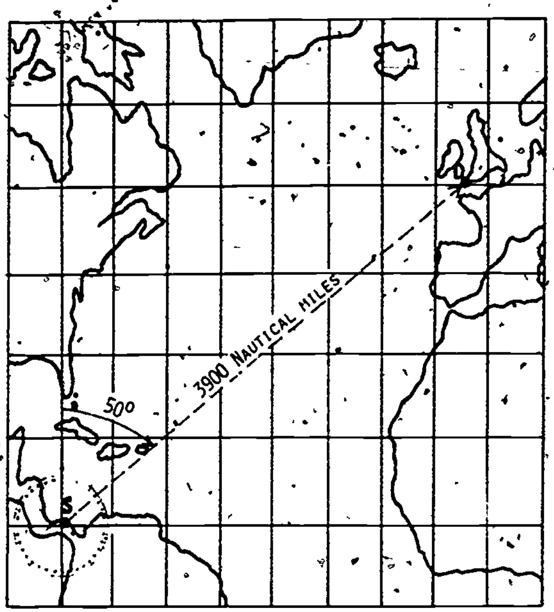


Figure 6a. A straight line course joining the Panama Canal to Land's End, England drawn on a plane chart. It advises us to use a compass bearing of 50° as shown.

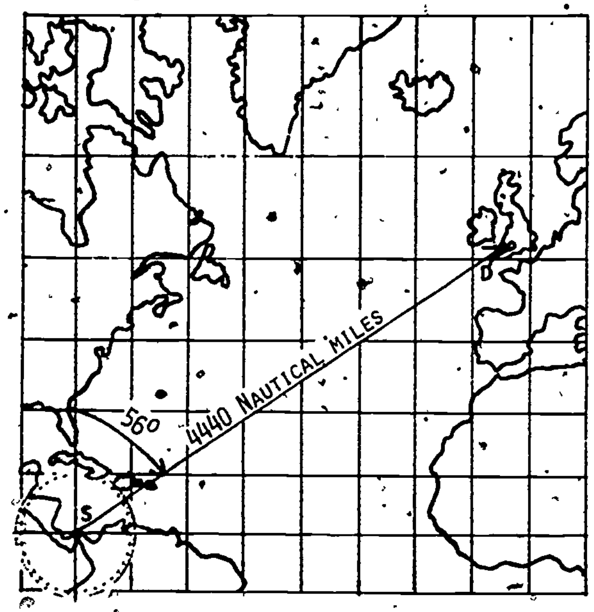


Figure 6b. The comparable straight line course on a Mercator map. The correct compass direction to follow is 56° east of compass north.

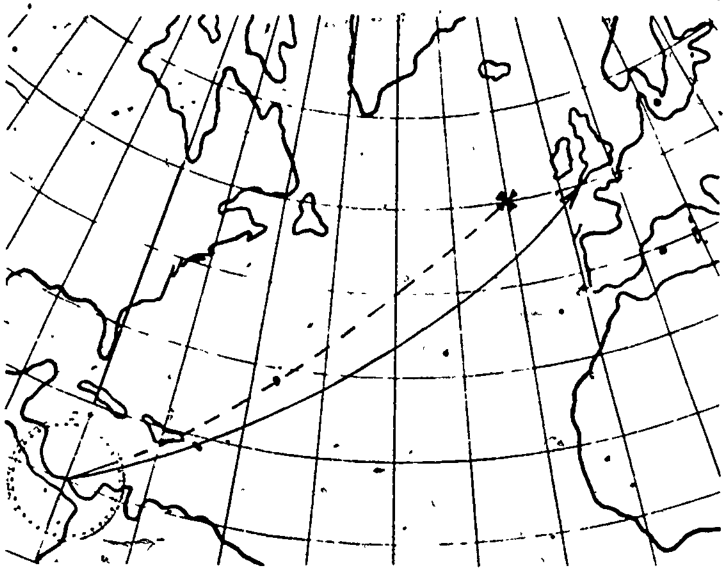


Figure 7. The navigational advice obtained from charts 6a and 6b leads to different results. The solid line shows the course from 6b, a rhumb line that does join the Panama Canal and Land's End. If we obeyed the plane chart in 6a and followed a constant 50° compass bearing we would be far off course, as the dashed path shows.

GERHARDUS MERCATOR is the Latinized name of Gerhard Krämer, born in Flanders in 1512. He was the expert engraver of map-sections for a globe made by Gemma Frisius in 1536, a craftsman of mathematical and astronomical instruments, and a land surveyor. His major achievements as a cartographer include a globe in 1541, a large (132 x 159 cm) map of Europe (1554) which made his reputation and was reprinted with corrections in 1572, a map of the British Isles in 1564 and the great world map of 1569. His major work was done at Duisberg, Germany where he was cosmographer to the Duke of Cleves. Mercator spent his final years creating a collection of maps of west and south Europe, of high accuracy for the period. It was published in 1595, a year after his death, as *Atlas - or Cosmographic Meditations on the Structure of the World.* Thus the word "atlas" was first applied to a collection of maps. He should not be confused with Nicholas Mercator, 1614-1687, mathematician and astronomer, nor were they related. (Source: *Dictionary of Scientific Biography* Vol. IX, Am. Council of Learned Societies, 1974, p. 309.)

measure because the scale changes as we look along vertical lines

In the schoolroom where students are learning about the relative sizes and locations of countries the map is at its worst. As a navigational aid, the map has been unsurpassed for 400 years because loxodromes appear as straight lines and angles measured on the map are the same as those measured on the globe.

1.5 Modern Navigators Use Mercator Charts

The shortest path between two points on a sphere is the great circle route [*S-2*]. Modern air and sea navigators naturally prefer to follow that shortest route. To do so, they begin by plotting the course with a straightedge on a gnomonic map (Figure 8) on which all great circle routes appear as straight lines. However, the compass direction changes continually along the great circle route (which, except in special cases is not a rhumb line), and pilots still expect to be told to follow a fixed compass direction. It is thus convenient and usual to replace the great circle route with a sequence of rhumb lines.

Because angles cannot be measured readily on a gnomonic or great circle chart, the navigator selects convenient intersections with the meridians along his great circle course and plots these points on a large scale Mercator map. Straight lines drawn between these points on the Mercator projection give rhumb lines which are easy to follow as a course and which usefully approximate the great circle path. Figure 9 shows the resulting course on the Mercator map. The extra distance involved in following the rhumb line pieces instead of the great circle itself is minor in comparison to the improved ease and certainty of navigation.

1.6 The Integral $\int \sec\phi \, d\phi$ Is Involved

In this paper we'll explore the construction of the Mercator map in some detail. We will see why, a century before Newton and Leibnitz created the calculus, Mercator

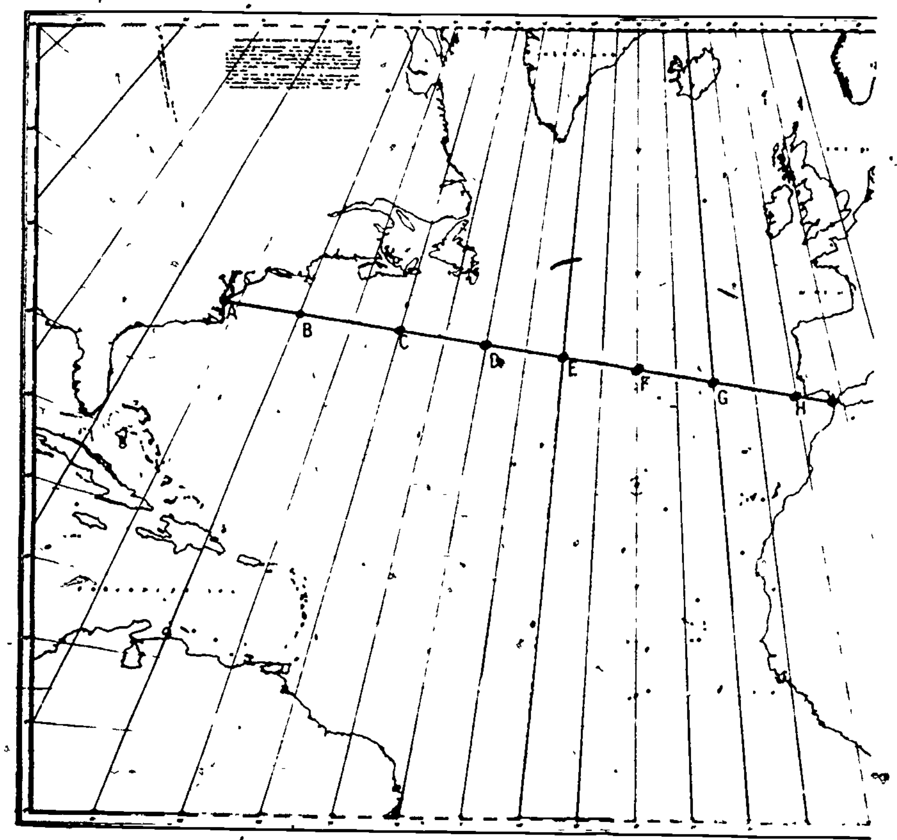


Figure 8. A great circle route appears as a straight line on this gnomonic projection. (The path appears curved because of an optical illusion; sight along it to verify that it is a straight line.)

found himself in need of the integral

$$\int_0^\phi \sec\phi \, d\phi .$$

We'll briefly cover the mathematical history of this integral as well. For all practical purposes this integral was evaluated long before the invention of the calculus although no proof appeared until 1668, when the calculus was newborn but known.

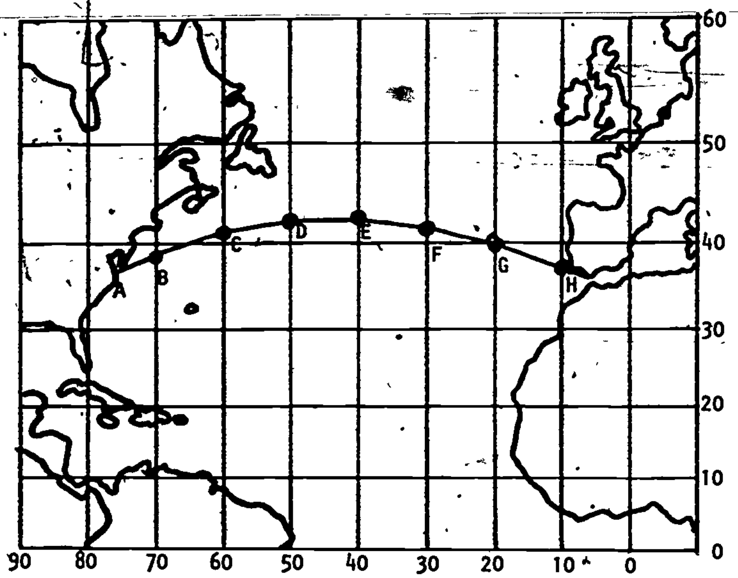


Figure 9. A series of rhumb line paths (straight line segments on this Mercator map) approximating the great circle route of Figure 8.

2. CALCULUS AND THE MERCATOR MAP

2.1 The Framework of the Mercator Map

Let's begin to create Mercator's map. The equator is a rhumb line in the east-west direction and will have to be a straight line on the map; let's place it horizontally across the middle. The meridians of longitude are the north-south rhumb lines and must also appear on the map as straight lines. Let's place them vertically, and space them evenly. This gives us accurate right angles between the north-south and east-west meridians and the equally divided equator on the map. The other east-west rhumb lines include the arctic and antarctic circles, the tropics of Cancer and Capricorn and all the other parallels of latitude. As we will see; our main problem is to place them as horizontal lines with such a spacing that rhumb lines will turn out to be straight lines on the map.

Two of our assumptions should be made explicit:

Our map is in "one-to-one" scale: we will duplicate distances along the equator mile for mile (although other distances will be distorted). That does not yield a pocket size map but scaling it down to printable size is an easy matter. Thus we'll soon talk of "stretching" earth distances to put them on the map!

We take the earth as a sphere. Cartographers can include the planet's equatorial bulge, but we will not attempt to do that here.

2.2 Horizontal Distances at Latitude ϕ Get Stretched

So far we have placed a family of parallel meridians on the map at right angles to the horizontal equator. Our troubles begin when we try to place the parallels of latitude on the map. In Figure 10, distances along the parallels of latitude between specified meridians are seen to shrink to zero as we move toward the poles, but those distances will have to be equal on the map because, there, meridians are parallel lines. Thus horizontal distances on the map will have to be longer than the true earth distances, and the stretching will have to increase as we move toward the poles. The vertical placement of these stretched horizontal lines will have to be skillfully done to keep the rhumb lines straight.

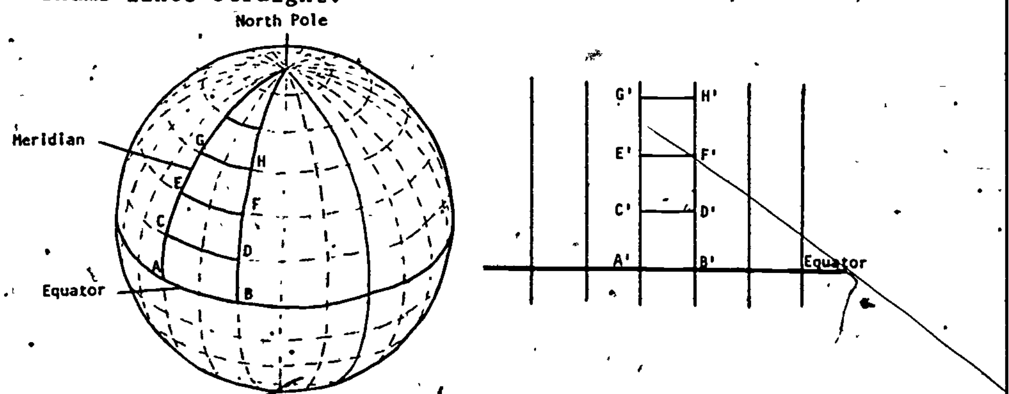


Figure 10. Corresponding points on meridians and map: EF on the globe stretches to E'F' on the map.

To see how to place the parallels of latitude, we must study the horizontal stretching. A wedge of the earth and its associated map are shown in Figure 11. Segment AB is a part of the equator spanning θ (radians) longitude. If R (= approximately 4000 miles) is the earth's radius, then Rθ is the actual length of AB and of the corresponding A'B' on the map.

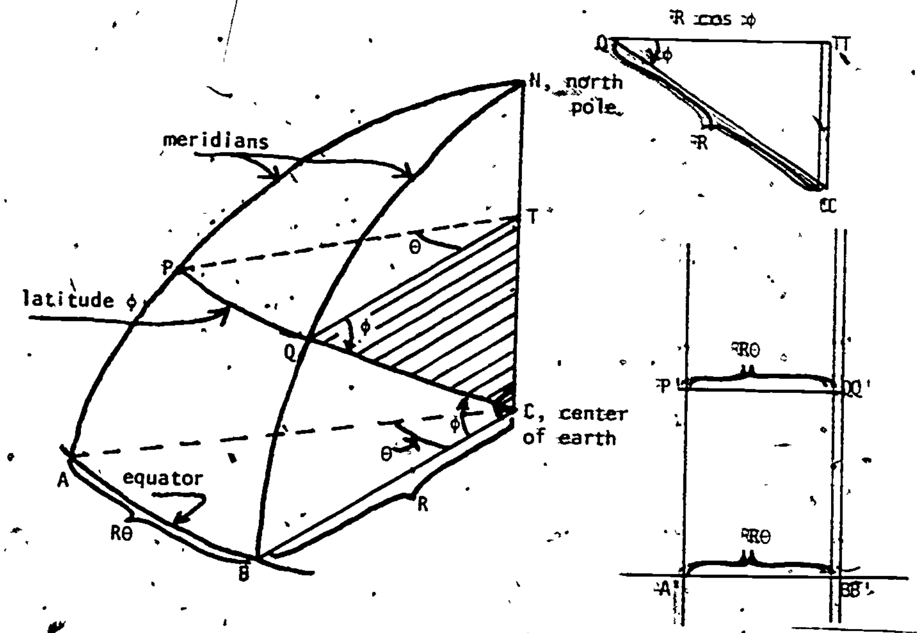


Figure 11. A wedge of the earth and its corresponding part of the map.

Now consider PQ at φ radians north latitude. PQ is a part of a circle centered at T, where T is directly north of the earth's center C, inside the earth. Since QT and BC are parallel, the angle φ of latitude appears in triangle QTC where shown. Since QC is an earth-radius, we see that PQ is a part of a circle of radius QT = R cos φ. The sector PTQ has the same central angle θ as does sector ABC; thus the actual length of PQ is

$$PQ = (QT)\ \theta$$

or

$$PQ = (R \cos \phi)\theta.$$

The horizontal stretching can now be understood. An east-west length PQ = (R cos φ) θ at latitude φ must be stretched into P'Q' = Rθ on the map. (Why does P'Q' = A'B' = Rθ?) We must stretch PQ by the factor $\frac{P'Q'}{PQ}$ to convert it into P'Q' for the map, because

$$P'Q' = \left[\frac{P'Q'}{PQ}\right] PQ.$$

Since

$$\frac{P'Q'}{PQ} = \frac{R\,\theta}{(R \cos \phi)\,\theta} = \frac{1}{\cos \phi} = \sec \phi$$

we get as the length of P'Q' on the map

$$P'Q' = (PQ) \sec \phi .$$

2.5 Mercator's Insight: Vertical Distances Must Also Be Stretched

Mercator's great insight was that each piece *of vertical meridian* at latitude φ must be stretched, when put on the map, by the same factor sec φ. As we shall see, he thus succeeded in preserving angles from the earth onto the map. That is, if any two lines meet at an angle on the earth, their images copied onto the map will meet at that same angle. This will be true for all angles everywhere on globe and map. (A map that preserves angles is called *conformal*; the study of just which globe-to-map functions yield conformal maps is an important part of advanced mathematics and cartography.)

Why does making the map conformal cause the rhumb lines to appear as straight lines? On the earth, recall, a rhumb line cuts all the meridians at a constant angle. If the map is conformal, the rhumb line on the map will cut all the vertical *parallel* meridian lines at that fixed angle and will thus be a straight line, for straight lines are exactly the curves that cut a family of parallel lines all at the same angle in plane geometry. Thus to have rhumb lines plot as straight lines, the whole secret is to space out the horizontal lines correctly, placing them at such distances from the equator line on the map that angles will be preserved. (Of course, stretching the meridian lines and spacing the parallels of latitude around the equator are two names for the same task.)

ERIC

2.4 The Vertical and Horizontal Stretching at Latitude ϕ Must Be Equal

Let's explore the stretching further. Figure 12 shows a rhumb line cutting a meridian at angle α. It cuts the horizontal parallel of latitude at the complementary angle

$$\beta = \frac{\pi}{2} - \alpha.$$

Suppose we move a small distance Δz along this rhumb line away from the crossing point. This movement is the combined effect of simultaneously moving $\Delta z \cos\beta$ units horizontally (eastward) and $\Delta z \sin\beta$ units vertically (northward).[3] What will this movement look like on our Mercator map? If an initial point was at latitude ϕ, the $\Delta z \cos\beta$ horizontal portion of the movement is stretched by a factor $\sec\phi$. If the angles α and β are going to be preserved on the map, the vertical component of the motion *must* also be stretched by the same factor $\sec\phi$, becoming $(\Delta z \sin\beta)\sec\phi$. Then Δz on the earth is mapped as $\Delta z \sec\phi$ on the map. (See Fig. 13)

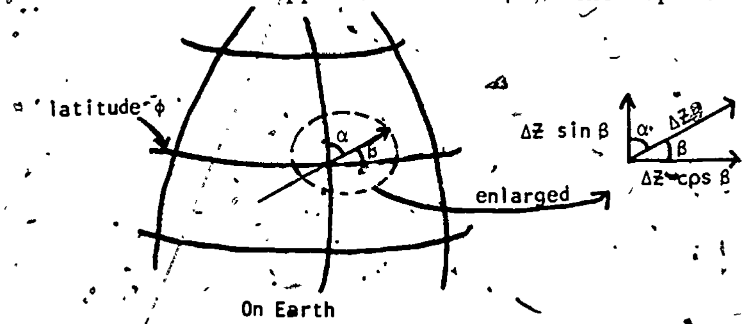


Figure 12. The local scene on a globe at latitude ϕ.

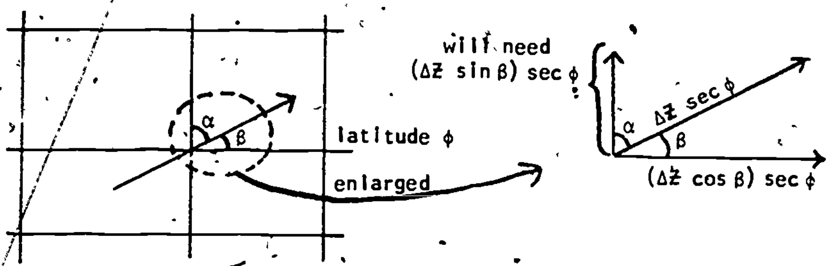


Figure 13. Same local scene on the map.

[3]Recalling that $\cos^2\beta + \sin^2\beta = 1$ may help here.

2.5 Summary: How We Get Straight Rhumb Lines

A concise summary of our logic now reads as follows:

1. To get rhumb lines to appear as straight lines on the map, we need to preserve angles from the earth onto the map.
2. Horizontal distances at latitude ϕ are stretched by a factor $\sec\phi$ as they are shifted from globe to map.
3. To preserve angles, we must also stretch the vertical lengths along the meridians by the same factor $\sec\phi$ at latitude ϕ.

2.6 How To Place The Parallels of Latitude

As we move north along a meridian, the latitude changes continually. What will it mean "to stretch the vertical lengths along the meridian by the same factor $\sec\phi$ at latitude ϕ"?

Integral calculus provides a method. Let's try to calculate $D(\phi_0)$, the distance *on the map* along the meridian from the equator to the parallel at latitude ϕ_0. (If we knew the number $D(\phi_0)$ we would know how to locate the parallel at latitude ϕ_0 on the Mercator map.) First, we cut the interval from 0 to ϕ_0 into many small pieces: let $\Delta\phi$ represent a bit of angle located near ϕ, where $0 \le \phi \le \phi_0$. This small bit of latitudinal angle subtends a bit of meridian $R\Delta\phi$ on the globe (Figure 14), a length of meridian located roughly at latitude ϕ. As this

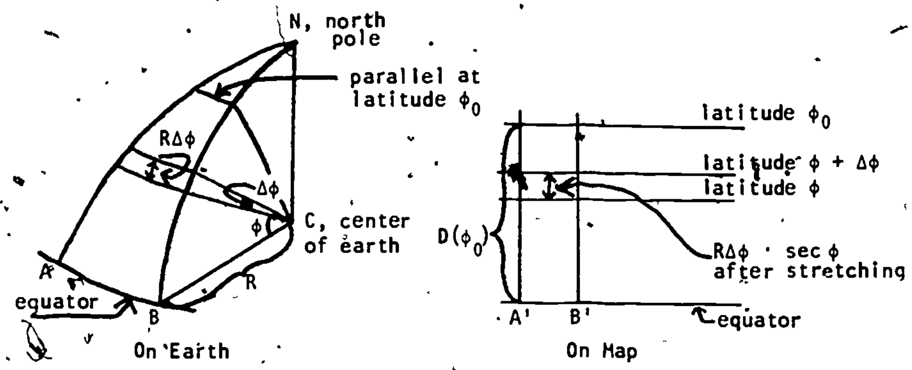


Figure 14. Setting up the interval for $D(\phi_0)$.

bit of meridian is shifted from globe to map, it is stretched by the factor $\sec\phi$ and has length $R\Delta\phi\sec\phi$ on the map.

Thus $D(\phi_0)$ is approximately the sum of such bits of length $R\sec\phi\,\Delta\phi$ as ϕ moves from 0 to ϕ_0:

(1) $$D(\phi_0) \approx \sum R\sec\phi\,\Delta\phi.$$

If we let all the $\Delta\phi$ lengths tend to zero and use more and more of them, we get better and better approximations of $D(\phi_0)$; in the limit we get

(2) $$D(\phi_0) = \int_0^{\phi_0} R\sec\phi\,d\phi = R\int_0^{\phi_0}\sec\phi\,d\phi.$$

To place all the parallels of latitude on the Mercator map, we will need $D(\phi_0)$ for all values $0 \leq \phi_0 \leq \frac{\pi}{2}$. Thus we need $\int\sec\phi\,d\phi$ to construct the Mercator map!

3. MORE HISTORY[4]

3.1 Mercator's Map: Cartography In His Time

Mercator did not know that he needed the calculus to make his map. He did know that he must place the equally-spaced-on-earth parallels of latitude further and further apart. His map contained minor errors in the placing of the parallels of latitude; it also contained misplaced mountain ranges, rivers and continents, as the sketch (Figure 15) of the original 131 x 208 centimeter map shows very clearly. Mercator's sources were the written itineraries of travelers and the older maps of his day, both notoriously inaccurate. Where modern mapmakers spend their energy on the accurate accumulation of data, Mercator's main task was to reconcile the inevitably contradictory reports that reached him.

One severe example will show the inaccuracies of mapping at that time. Mercator's map constituted the first useful, dramatic improvement in mapping the known world since the time of Ptolemy (the great astronomer and geographer) 1400 years earlier. An important error on those early maps resulted from taking 1° as 56.5 miles

[4]The cartographic history in this section is taken from Crone (1966). The mathematical history is drawn from the notes of Professor V. F. Rickey and from Cajori (1915).

ERIC

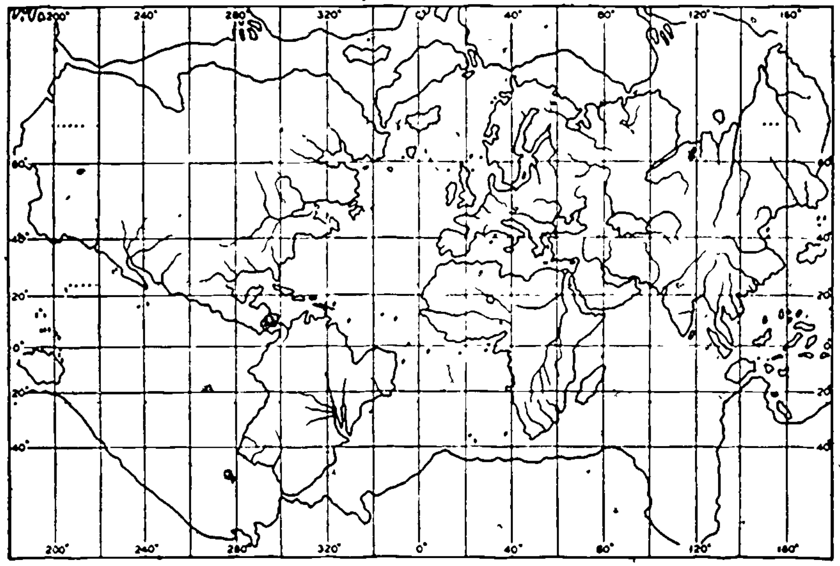


Figure 15. Sketch of Mercator's map of 1569.

on the earth's surface. An almost-correct value of 68.5 miles per degree was known to Eratosthenes (200 B.C.) but not accepted by Ptolemy. Thus distances were stretched across too many degrees of latitude and longitude. Ptolemy took the east-west length of the Mediterranean Sea as 62°. Mercator's value of 52° was a substantial improvement but a correct value of less than 42° was not known until after 1700 A.D. One result of this error is worth mention: geographers of the generation before Columbus had stretched the Europe-Asia land mass much too far around the globe; Columbus had reason to believe that a journey of reasonable length to the west would bring him to the orient. Maps of modern quality did not appear until nineteenth century explorers began to carry sophisticated instruments into the field.

Mercator, facing these complex problems, spaced his parallels of latitude as best he could on the map of 1569. His exact method is not known. And sailors, properly distrustful of mapmakers, did not adopt the Mercator map at once. By 1600 or so, Mercator maps of portions of the earth began to appear. These were of larger scale and incorporated corrections in the placement of the parallels of latitude due to the work of Edward Wright. Acceptance by sailors grew steadily. The first atlas of Mercator projections was the *Arcano del Mare* of Sir Robert Dudley, published in 1646. By that time Mercator maps were the navigator's standard.

3.2 Edward Wright's Discovery

The mathematical history that arose from Mercator's achievement is astonishing. As mentioned earlier, we do not know whether Mercator really understood where to place the parallels of latitude to straighten the rhumb lines. By the time Edward Wright published *Certaine Errors in Navigation* in 1599, the secret was out:

> "the parts of the meridian at every poynt of latitude must needs increase with the same proportion wherewith the Secantes or hypotenusae of the arke, intercepted betweene those pointes of latitude and the aequinoctiall (equator) do increase... by perpetuall addition of the Secantes answerable to the latitudes of each point...we may make a table which shall shew the sections and points of latitude in the meridians of the nautical planisphaere: by which sections, the parallels are to be drawne." [From Wright (1599, pp. 17-18) as quoted in Cajori (1915, pp. 312-313).]

Wright recognized that a sum of secants was needed; by his "perpetuall addition" we assume he meant the continuous summation of integration. He could not have known of integration as an anti-differentiation process, but the intuitive notion of a limit of integral sums was afloat in the intellectual seas of that time. To provide the navigational corrections, Wright published a table of summation-approximations of the integral (2) for ϕ between 0 and 45^o at intervals of 1 minute of latitude.

3.3 Later Mathematical History

Geographers really needed an integration formula for the integral so that lengthy summations could be avoided. The following fifty years saw a search for such a formula through non-calculus techniques. In 1614 Napier published tables of sines and logarithms of sines, although these were not quite logarithms as we know them today. In 1620 Edmund Gunter published a table of log(tan θ). By 1645, Henry Bond discovered by chance and published in Norwood's *Epitome of Navigation*, as Edmund Halley tells us half a century later,

> "that the Meridian Line was Analogous to a Scale of Logarithmick Tangents of half the complements of the Latitudes." [From Halley (1698, p.202) as quoted in Cajori (1915, p.314).]

Bond's discovery is that

(3) $$\int_0^{\phi} \sec\phi\, d\phi = -\ln \tan\left[\frac{1}{2}\left(\frac{\pi}{2} - \phi\right)\right],$$

a correct formula not usually seen by calculus students. Bond did not prove the formula, but led a number of prominent mathematicians, including John Collins, N. Mercator, William Oughtred and John Wallis, to attempt the proof. Bond's conjecture came from comparison of tables and graphs.

During the 1660's, Newton and Leibnitz produced a systematic calculus and by 1668, via a nastily complicated geometric argument, James Gregory proved the truth of (3).[5] During the next decade or two, simple calculations of $\int \sec\phi\, d\phi$ were found. Throughout this period, mathematicians were quite conscious that they were providing the theory necessary for an accurate Mercator projection and they consistently regarded the task as an important and worthy one.

Thus $\int \sec\phi\, d\phi$ is one case where an integral was first treated by summation long before the birth of the calculus

[5] I cannot resist including a quote, ascribed to Edmund Halley quoted from Cajori's article. Halley, in reviewing the research on our problem, says about Gregory's proof that it involved "a long train of Consequences and Complication of Proportions, whereby the evidence of the Demonstration is in a great measure lost, and the Reader wearied before he attain it."

and was eventually made part of the calculus mainstream. This integral is one of the most esoteric that calculus students are asked to handle because the usual integration methods given seem unmotivated and "magical." But the integral's importance for applications makes its study worthwhile, and we will next examine several techniques for calculating it. (If Sections 4 and/or 5 are omitted, it is still appropriate to read Section 6.)

Exercises

1. Differentiate to confirm that

a). $\int_0^x \sec x \, dx = -\ln \tan \left(\frac{1}{2}\left(\frac{\pi}{2} - x\right)\right)$

b). $\int_0^x \sec x \, dx = \ln (\sec x + \tan x)$ for $0 \le x < \pi/2$

c). $\int_0^x \sec x \, dx = \ln \tan \left(\frac{1}{2}\left(\frac{\pi}{2} + x\right)\right)$

2. Starting with blank graph paper, make part of a Mercator map, as follows: put in the equator, the meridians at 0°, 30°, 60°, ..., 330° latitude, and parallels of latitude at 10°, 20°, 30°, 40°, 50°, 60°, 70° and 80° north and south. (Arrange the scale so that these parallels do fit. Now, using a globe or non-Mercator world map as a source of data, sketch in Greenland and Africa. Do your results look about like Figure 5? Do Africa and Greenland have roughly equal areas, as the map seems to say? (look up the actual area in the almanac or atlas.)

3. The formulas in Exercise 1 are for x measured in radians. Convert any one of them so that it gives

$$\int \sec y \, dy$$

for y measured in degrees.

4. It is probable that Mercator constructed his map grid by using a table of secants at one degree intervals. Adding up the secants from one degree to 30 degrees would give him the approximate spacing of the 30° line of latitude in terms of the size of one

degree at the equator. Without knowing it, he was using the approximation

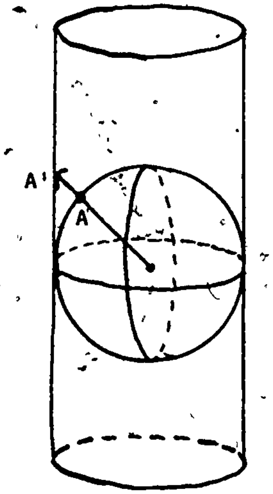
$$\int_0^{30^\circ} \sec \phi \, d\phi \cong \sum_{i=1}^{i=30} \sec \phi_i \, \Delta\phi_i .$$

Try this approximation yourself to find the distance in earth radii (radians) from the equator to 30° north latitude on the map. First use steps of 5° and then of 1° (Remember $1^\circ = \frac{\pi}{180}$ radians). Compare your results to the exact value given by equation 3. Do you think Mercator's probable method was sufficiently accurate for a small scale world map? When you have completed Section 5, compare your result to the value given by the series approximation of Exercise 12.

5. I was taught, erroneously, that a Mercator map is obtained when a paper cylinder is wrapped around the earth, tangent at the equator as in the sketch, and points on earth are projected onto the cylinder as though by a point-size light at the earth's center. What spacing of the longitude-circles does this projection involve (instead of the

$$D(\phi) = R\int_0^{\phi} \sec \phi \, d\phi$$

placement of the circle at latitude ϕ on the Mercator map)? Are the longitude lines more spread out on this map or on the Mercator map? Figure 4 shows a map made with this cylindrical projection.



CYLINDRICAL PROJECTION
Point A is mapped to A'

(Hints for Exercises 3 and 5 may be found on page 32.)

ERIC Full Text Provided by ERIC

6. Answer each question, supporting your answer with specific evidence from the unit:

a) Do we know enough about integration when we have learned to calculate integrals by antidifferentiation?

b) Did Newton and Leibnitz create the calculus in response to scientific questions as part of the intellectual growth of their age or did they create it out of thin air because of its internal logic and beauty?

c) Can $\int \sec(x)\,dx$ be calculated between limits x = a and x = b without the use of a "closed integration formula?"

d) What is the advantage of sailing a rhumb line course as opposed to another course? Are there disadvantages in sailing the rhumb line and, if so, what are they?

e) On an accurate Mercator map of the world, how or where is the north pole located?

f) On an accurate Mercator world map, does an inch of map distance along the parallel of latitude at 40° north represent the same earth-distance as and inch of map distance along the parallel at 30° north?

g) Is the Mercator map an easy one to use to measure the distance between Chicago and New Orleans?

4.* SEVERAL CALCULATIONS OF $\int \sec x\,dx$

4.1 The Usual Integration

A typical, sneaky calculation of this integral is

$$\int \sec x\,dx = \int \sec x \cdot \frac{\sec x + \tan x}{\sec x + \tan x}\,dx$$

$$= \int \frac{\frac{d}{dx}(\sec x + \tan x)}{\sec x + \tan x}\,dx$$

(4) $$= \ln|\sec x + \tan x| + c$$

*This section may be omitted. See Suggested Uses on inside of title page.

and no motivation other than "look, it works" seems possible. We have used the well-known result

$$\int \frac{dy}{y} = \ln|y| + c.$$

4.2 A Partial Fractions Integration

A little obvious trigonometry permits us to calculate the integral by partial fractions. Some equal signs have been marked for further comment:

$$\int \sec x\,dx = \int \frac{dx}{\cos x} \boxed{=} \int \frac{\cos x\,dx}{\cos^2 x}$$

$$\textcircled{=} \int \frac{\cos x\,dx}{1-\sin^2 x} = \int \frac{\cos x\,dx}{(1-\sin x)(1+\sin x)}.$$

The multiplication by 1 = cos x/ cos x at $\boxed{=}$ is done so that the next step $\textcircled{=}$ can be done, a modest example of planning ahead. Here are the partial fractions:

$$\frac{1}{(1-\sin x)(1+\sin x)} = \frac{1/2}{1-\sin x} + \frac{1/2}{1+\sin x}.$$

Thus

$$\int \sec x\,dx = \frac{1}{2}\int \frac{\cos x}{1-\sin x} + \frac{\cos x}{1+\sin x}\,dx$$

$$= \frac{1}{2}[-\ln(1-\sin x) + \ln(1+\sin x)] + c$$

$$= \frac{1}{2}\ln\frac{1+\sin x}{1-\sin x} + c$$

$$\boxed{=} \frac{1}{2}\ln\left(\frac{1+\sin x}{1-\sin x}\cdot\frac{1+\sin x}{1+\sin x}\right) + c$$

$$\textcircled{=} \frac{1}{2}\ln\frac{(1+\sin x)^2}{1-\sin^2 x} + c$$

$$\textcircled{=} \frac{1}{2}\ln\frac{(1+\sin x)^2}{\cos^2 x} + c$$

$$= \ln\sqrt{\left(\frac{1+\sin x}{\cos x}\right)^2} + c$$

$$\overset{\triangle}{=} \ln\left|\frac{1+\sin x}{\cos x}\right| + c$$

$$= \ln|\sec x + \tan x| + c.$$

Again, the decision at ⊟ leads to improvements at the following ⊜ steps.

The step marked △ rests on the fact that

(5) $$\sqrt{y^2} = |y|$$

although you might think more immediately of

(6) $$\sqrt{y^2} = y?$$

Both (5) and (6) are correct when $y \geq 0$ but (5) is still true when y is negative:

$$\sqrt{(-5)^2} = |-5| = 5$$

while (6) is not:

$$\sqrt{(-5)^2} \neq -5.$$

Thus, the absolute value bars arise very naturally in the integration formula.

The calculation involves no trigonometry more sophisticated than $\sin^2 x + \cos^2 x = 1$ and $1/\cos x = \sec x$, but requires a little algebraic organization. It was apparently first done by Isaac Barrow in about 1670 and may be the earliest use of partial fractions in integration.

4.3 Gregory's Form Of the Integral

It is not difficult to derive (3), Gregory's form of the integral. The needed trigonometry this time is that

$$\cos x = \sin\left(\frac{\pi}{2} - x\right)$$

and the double angle formula

$$\sin\left(\frac{\pi}{2} - x\right) = 2\sin\left(\frac{1}{2}\left(\frac{\pi}{2} - x\right)\right)\cos\left(\frac{1}{2}\left(\frac{\pi}{2} - x\right)\right).$$

Here it is:

$$\int \sec x\,dx = \int \frac{dx}{\cos x} = \int \frac{dx}{\sin\left(\frac{\pi}{2}-x\right)}$$

$$= \int \frac{dx}{2\sin\left(\frac{\pi/2-x}{2}\right)\cos\left(\frac{\pi/2-x}{2}\right)}$$

$$\overset{\boxminus}{=} \int \frac{dx}{2\frac{\sin y}{\cos y}\cos^2 y} \quad \text{where } y = (\pi/2-x)/2$$

$$= \frac{1}{2}\int \frac{\sec^2 y\,dx}{\tan y}.$$

Now change variables to y, using $dy = -\frac{1}{2}dx$, and get

$$= -\int \frac{\sec^2 y\,dy}{\tan y} = -\ln|\tan y| + c$$

(7) $$= -\ln\left|\tan \tfrac{1}{2}\left(\tfrac{\pi}{2}-x\right)\right| + c.$$

The algebra at ⊟ is sneakier than one might like, I admit. The minus sign here is not unreasonable. For our basic interval $x \in \left(0,\frac{\pi}{2}\right)$, $\sec x > 0$ and the integral > 0. But $(\pi/2-x)/2$ is between 0 and $\pi/4$, its tangent is between 0 and 1, and the ln would be negative. The minus sign straightens that out.

Another form of the integral is given in Problem 7. And it should be possible to convert $\ln(\sec\theta + \tan\theta)$ into $-\ln(\tan((\pi/2-x)/2))$ via trigonometry, should it not? You are asked to do so in Problem 8.

JAMES GREGORIE or Gregory, 1638-1675, created much more of importance in mathematics and optics than he was given credit for in his own day. A great technological achievement of that time was the design of efficient low-distortion telescopes. Gregory contributed experimental designs that influenced Newton's reflector telescopes; the Cassegrain design in 1672 was the ultimate successful result of this effort.

Gregory put much effort into finding the lengths, areas and volumes associated to the conic sections. These results were needed for engineering work such as design of optical instruments. Difficult integrations were involved, and were done by geometric methods using classical Greek knowledge of the conics, physical principles, etc. The calculation of $\int \sec x\, dx$ for use in the Mercator Projection is one example. His later mathematics centered on calculation of roots of polynomials. He used approximation methods that were rediscovered by Newton, Simpson, Taylor and Cotes some years later, and credited to them. He was also a pioneer in the use of infinite series; see Exercise 13.

His work was not influential because Gregory, teaching at isolated universities, was too much out of communication with his proper peers. His generation saw its work absorbed as small portions of the deeper, richer, systematic calculus developed by Newton and Leibniz. (Source: *Dictionary of Scientific Biography*, Vol. V, pp 524-530 and C.B. Boyer, *A History of Mathematics*.)

JOHN WALLIS (1616-1703) was the greatest English mathematician of the generation that preceded Newton. He was an important leader of the transition from Greek geometric methods to modern algebraic methods. Two books published by him in 1655, one on analytic geometry, the other on infinitesimal methods, were both influencial. He derived many fundamental results of the calculus, including

$$\int_0^a x^n\, dx = a^{n+1}/(n+1)$$

by algebra-based methods that were a great simplification of the geometric derivations used earlier by Cavalieri and others. Some of his "proofs" were incomplete or erroneous and were critized during his life even though the results were correct; he thus helped mathematics make enormous progress, leaving the rigorous and most efficient derivations to be found by others later. He was also a clergyman, and chaplain to King Charles II. (Source: C. B. Boyer, *A History of Mathematics*, John Wiley, New York, 1968, Chapter XVIII.)

266

Exercises:

7. Use ideas close to those we used to derive Gregory's formula,

$$\int \sec x\, dx = -\ln \tan \tfrac{1}{2}\left(\tfrac{\pi}{2} - x\right) + c,$$

in section 4 to derive this formula:

$$\int \sec x\, dx = \ln \tan \tfrac{1}{2}\left(\tfrac{\pi}{2} + x\right) + c.$$

8. a). Show via trigonometry that, if $0 \le x < \pi/2$,

$$\tan\left(\tfrac{1}{2}\left(x + \tfrac{\pi}{2}\right)\right) = \sec x + \tan x.$$

This may be done in many ways. My own method started from the formulas for tan (A/2) and tan (A+B).

b. Now show $\cot\left(\tfrac{1}{2}\left(\tfrac{\pi}{2} - x\right)\right) = \tan \tfrac{1}{2}\left(\tfrac{\pi}{2} + x\right)$. (Hint: draw graphs of the tangent and cotangent curves and give a geometric sort of proof.) Then explain why the two integrations listed in problem 7 are equivalent.

(Hint for Exercise 7 may be found on page 32.)

5.* A SERIES FOR $\int \sec x\, dx$

Recall that Mercator's need was to calculate

$$\int_0^{\phi} \sec \phi\, d\phi$$

for many values of ϕ, even every 1/60 of a degree. While Gregory's proof that a "logarithmick Tangent" formula was correct for this integral was valuable, it helped the task of computation only to the extent that tables of log (tan x) were available. In 1685, John Wallis published a series form of the integral, offering a wholly new and fairly convenient computational method.

*This section may be omitted. See Suggested Uses on inside of title page.

5.1 Derivation of Wallis' Series

This series is very easy to derive. From section 4.2, in the partial fractions derivation, we have

$$\int_0^x \sec x\,dx = \ldots = \int_0^x \frac{\cos x\,dx}{1-\sin^2 x}$$
$$= \int_0^x \cos x\,[1 + \sin^2 x + \sin^4 x + \sin^6 x + \ldots]dx.$$

All we have done here is to use the geometric series formula

$$\frac{1}{1-a} = 1 + a + a^2 + a^3 + a^4 + \ldots \text{ if } |a| < 1$$

with $a = \sin^2 x$, (which does satisfy $|a| < 1$ for the $x\epsilon(0,\pi/2)$ that concern us).

The next step is to convert this integral-of-an-infinite-sum into an infinite-sum-of-integrals, which we calculate term by term[6]. Continuing:

$$\int_0^x \sec x\,dx = \int_0^x \cos x \cdot 1\,dx + \int_0^x \cos x \cdot \sin^2 x\,dx +$$
$$\int_0^x \cos x \cdot \sin^4 x\,dx + \int_0^x \cos x \cdot \sin^6 x\,dx + \ldots$$
$$= \sin x + \frac{\sin^3 x}{3} + \frac{\sin^5 x}{5} + \frac{\sin^7 x}{7} + \ldots$$

(8) $$= y + \frac{y^3}{3} + \frac{y^5}{5} + \frac{y^7}{7} + \ldots \text{ with } y = \sin x.$$

This series is convergent for any $x\epsilon[0,\pi/2)$ as Exercise 11 asks you to show.

[6] The reader should be warned that, in general, it is not true that the integral of an infinite sum is equal to the term-by-term sum of the integrals. However, as you will see proven in more advanced courses, the calculation here is legal because the series involved is convergent for all values in a closed interval $[0,x]$ where $0 \le x < \frac{\pi}{2}$ and the functions involved, including the sum, are all continuous.

5.2 Numerical Approximation of the Integral

We can use (8) to approximate

$$\int_0^x \sec x\,dx$$

by getting $y = \sin x$ from a table and taking partial sums of (8) until the desired level of convergence is obtained. You are asked to do so on the computer in Exercise 12. The series is not an exceptionally fast-converging one. For $x = \pi/6$ some partial sums are:

Highest power of sin x included	Partial Sum
1	0.49999999
3	0.54166666
5	0.54791666
7	0.54903273
9	0.54924975
11	0.54929414
13	0.54930352
15	0.54930556
17	0.54930601

The correct value, for comparison, is

$$\int_0^{\pi/6} \sec x\,dx \doteq 0.54930614.$$

Many integrations that lead to obnoxious formulas can be converted into series calculations leading to convergent, computable answers. Wallis published a series for

$$\int \tan x\,dx$$

in 1685, too, and we include this one as Exercise 9 as one example. See also Exercise 10.

6. WHAT HAS CALCULUS CONTRIBUTED TO THE MERCATOR PROJECTION?

The map that Mercater published in 1569 was revolutionary because it simplified the task of navigation at sea -- sailors could plot rhumb line courses by the simple use of straight lines. By about 1600 corrected versions of Mercator world map were accurate enough to satisfy the practical requirements of navigation at sea and the map

ERIC

soon came into wide use. But all of this was accomplished before the invention of the calculus; what has calculus really added to the achievements of Mercator?

As more and more detailed Mercator charts of smaller and smaller parts of the earth's surface have been made over the centuries, a more and more accurate spacing of the parallels of latitude has been necessary. Once the precise mathematical calculation of $\int \sec\theta \, d\theta$ was known, this spacing could be immediately accomplished to any degree of accuracy. The only limitations in the production of Mercator maps were those imposed by the printing process, size and quality of paper, and so on. No mathematical barriers stood in the way of the cartographer, because methods had been provided to create the Mercator projection both in theory and to any level of accuracy in practice.

The influence of Mercator on the course of mathematical development was important. Along with many other technological problems of that age, the problem of refining the Mercator projection to a high level of accuracy inspired the mathematicians and guided their efforts in developing the calculus. They did not quit working on cartography-inspired problems once the Mercator problem had been solved, of course. Since 1600 the Mercator projection has been further refined (to take into account the equatorial bulge of the earth, for example) through use of more mathematics and many other projections have been developed on a sound mathematical basis.

Exercises

9. Use this start to get a series-form, also given by Wallis in 1685, for

$$\int \tan x \, dx = \int \frac{\sin x \, dx}{\cos x}$$

$$= \int \frac{\sin x \cos x \, dx}{\cos^2 x}$$

$$= \int \sin x \cos x \left(\frac{1}{1-\sin^2 x} \right) dx$$

The answer will be

$$\int_0^x \tan x \, dx = \frac{1}{2}\left(s^2 + \frac{s^4}{2} + \frac{s^6}{3} + \frac{s^8}{4} + \dots\right)$$

where $s = \sin (x)$.

For what x does this converge and why?

10. What goes wrong when we try to carry out Exercise 9 for

$$\int_0^x \cot x \, dx?$$

11. Give a proof that Wallis' series

$$y + \frac{y^3}{3} + \frac{y^5}{5} + \frac{y^7}{7} + \dots$$

with $y = \sin x$ for some $x \in [0, \pi/2)$, is convergent

a) by a comparison test

b) by another test

c) For what y (and then what x) does this series converge?

12. A computing project: use Wallis' series

$$\int_0^x \sec x \, dx = s + \frac{s^3}{3} + \frac{s^5}{5} + \dots \text{ where } s = \sin (x)$$

to calculate on the computer successive partial sum approximations of the integral. Your instructor will tell you what interval $[0,x]$ to use. Continue until you have the integral approximated within $.5 \times 10^{-6}$. How will you decide when you have calculated enough terms to the series and are ready to get off the machine?

13. a) Derive Gregory's series:

$$\arctan x = \int_0^x \frac{dx}{1 + x^2} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$$

Hint: Replace $1/(1+x^2)$ by a geometric series before integrating term by term.

b) For what x does this alternating series converge?

c) Use the computer and this series to get a table of arctan x for x-values that your instructor assigns. (How can you easily decide when to get off the machine, for $x \in [0,1]$? What is an estimate of the error if you stop after so many terms?)

d). What does the series tell you about a relationship between arctan x and arctan $(-x)$?

Hints for Exercises

3. Change variables using $y = \frac{180}{\pi} x$. The result will be an integral in x, where x is measured in radians.

5. In Figure 13, angle ϕ is the latitude of points A, A', and $D_2(\phi)$ gives the spacing of the parallel of latitude just as $D(\phi)$ did for the Mercator map. Try to find $D_2(\phi)$ and then show

$$D_2(\phi) \geq D(\phi).$$

7. Start with $\int \sec x \, dx = \int \frac{dx}{\cos x} = \int \frac{dx}{\sin(x + \frac{\pi}{2})}$.

9. Reread the beginning of Section 5.1.

10. The lower limit of integration, 0, causes the integral to be improper. Is it finite?

136

7. REFERENCES

Cajori, Florian, 1915: "On an Integration Ante-dating the Integral Calculus" Bibliotheca Mathematica, 3rd Series, 14 (1915) pp. 312-319.

Carslaw, H.S., 1924: "The Story of Mercator's Map. A Chapter in the History of Mathematics." The Mathematical Gazette XII (Jan. 1924) pp 1-7.

Crone, G.R., 1966: Maps and Their Makers. Capricorn Books, New York, 1966.

Deetz, Charles H. and Oscar S. Adams, 1944: Elements of Map Projection With Application to Map and Chart Construction, 5th Edition, revised, 1944. Special Publication No. 68 of The Coast and Geodetic Survey, U.S. Department of Commerce.

Halley, Edmund, 1698: "An Easie Demonstration of the Analogy of the Logarithmick Tangents to the Meridian Line or sum of the Secants: with various methods for Computing the same to the utmost Exactness", Philosophical Transactions XIX (1695-1697), London, 1698.

Hobbs, Richard R., 1976: Marine Navigation I; Piloting. Fundamentals of Naval Science Series, Naval Institute Press, Annapolis, Md., 1974. Reprinted 1976 with corrections.

Parsons, E.J.S. and W.F. Morris, 1939: "Edward Wright and his Work", Imagio Mundi 3 (1939) pp 61-71.

Taylor, E.G.R. 1971: The Haven-Finding Art. Hollis and Carter London 1971.

Sadler, D.H. and E.G.R. Taylor, 1953: "The Doctrine of Nauticall Triangles Compendious. Part I - Thomas Hariot's Manuscript (by Taylor). Part II - Calculating the Meridional Parts (by Sadler)." J. of the Institute of Navigation, London 6 (1953) pp 131-147.

137

Wright, Edward, 1599: Certaine Errors in Navigation, Arising either of the ordinarie erroneous making or using of the sea Chart, Compass, Cross Staffe and Tables of declination of the Sunne, and fixed Starres detected and corrected, London, printed by Valentine Sims, 1599.

Note: Historical material in this paper has been drawn almost totally from (Cajori, 1915) and (Crone, 1966) only. Other references are given to allow the reader quick access to the literature for further research.

In this brief paper, complex historical ideas have naturally been compressed more than they deserve. Any inaccurate impressions that may be conveyed as a result are the sole responsibility of the author of this paper.

8. ANSWERS TO EXERCISES

1. a) $\frac{d}{dx}\left[-\ln\tan\left(\frac{1}{2}\left(\frac{\pi}{2}-x\right)\right)\right]$

$$= -\frac{1}{\tan\left(\frac{1}{2}\left(\frac{\pi}{2}-x\right)\right)}\sec^2\left(\frac{1}{2}\left(\frac{\pi}{2}-x\right)\right)\cdot\frac{1}{2}\cdot(-1)$$

$$= \frac{1}{2}\,\frac{1/\cos^2 y}{\sin y/\cos y} \qquad \text{where } y=\frac{1}{2}\left(\frac{\pi}{2}-x\right)$$

$$= \frac{1}{2\cos y\,\sin y} = \frac{1}{\sin(2y)} = \csc(2y)$$

$$= \csc\left(\frac{\pi}{2}-x\right) = \sec x.$$

b) $\frac{d}{dx}\left(\ln(\sec x+\tan x)\right)$

$$= \frac{1}{\sec x+\tan x}\cdot(\sec x\,\tan x+\sec^2 x)$$

$$= \frac{1}{\sec x+\tan x}(\sec x+\tan x)\sec x$$

$$= \sec x.$$

c) $\frac{d}{dx}\left[\ln\tan\left(\frac{1}{2}\left(\frac{\pi}{2}+x\right)\right)\right]$

$$= \frac{1}{\tan\left(\frac{1}{2}\left(\frac{\pi}{2}+x\right)\right)}\sec^2\left(\frac{1}{2}\left(\frac{\pi}{2}+x\right)\right)\cdot\frac{1}{2}$$

$$= \frac{1}{2\sin z\,\cos z} \qquad \text{for } z=\frac{1}{2}\left(\frac{\pi}{2}+x\right) \text{ just as in (a) above.}$$

$$= \frac{1}{\sin(2z)} = \csc\left(\frac{\pi}{2}+x\right) = \sec x.$$

2. The area of Greenland is approximately 840,000 square miles, and the area of Africa is 11,706,727 square miles.

3. $$\int \sec y\,dy = \int \sec\left(\frac{180}{\pi}x\right)\frac{180}{\pi}\,dx$$

$$= \frac{180}{\pi}\int \sec\left(\frac{180}{\pi}x\right)dx$$

$$= \frac{180}{\pi}\cdot\frac{\pi}{180}\ln\left|\sec\left(\frac{180}{\pi}x\right) + \tan\left(\frac{180}{\pi}x\right)\right| + c$$

$$= \ln|\sec y + \tan y| + c.$$

Change of variables:

$$y = \frac{180}{\pi}x$$

$$dy = \frac{180}{\pi}dx$$

4. With 5° steps, $\int_0^{30^\circ}\sec\phi\,d\phi \approx 0.5564789339.$

With 1° steps, $\int_0^{30^\circ}\sec\phi\,d\phi \approx 0.5506730838.$

Equation (3) gives $\int_0^{30^\circ}\sec\phi\,d\phi = -\ln\tan\left(\frac{1}{2}\left(\frac{\pi}{2}-\phi\right)\right)\Big|_{\phi=0}^{\phi=\frac{\pi}{6}}$

$\approx 0.5493061443.$

5. From Figure 15,

$$\frac{D_2(\phi)}{R} = \tan\phi$$

so, $$D_2(\phi) = R\tan\phi$$

must be compared to

$$D(\phi) = R\int_0^\phi \sec\phi\,d\phi.$$

The easiest way to show $D_2(\phi) \geq D(\phi)$ is to notice that the derivatives are easy to compare:

$$\frac{d}{d\phi}D_2(\phi) = R\sec^2\phi \geq R\sec\phi = \frac{d}{d\phi}D(\phi).$$

Since $D_2(0) = D(0) = 0$, we can conclude that $D_2(\phi) \geq D(\phi)$ for all $\phi \varepsilon (0,\frac{\pi}{2})$.

6. (Sample answers)

a) The construction of the Mercator map leads us to discover $\int \sec x\,dx$ as a limit of integral sums; derivatives do not enter the problem. The integral was adequately approximated decades before an antidifferentiation formula was discovered.

b) Ideas that are now part of the calculus and problems calling for the calculus were "in the air" long before Leibnitz and Newton. For example: Cavalieri's integration of x^n; Mercator's need for $\int \sec x\,dx$; Gregory's geometric calculation of integrals.

c) Yes, as a finite integral sum $\Sigma(\sec x_i)\Delta x$ or by use of finitely many terms of Wallis' series.

d) A rhumb line is easy to sail because the pilot simply keeps an eye on his compass. He wants to keep the compass needle reasonably still. One disadvantage of the rhumb line path is its greater length in comparison to the great circle path. Extra distance costs time, fuel, and money.

e) The north pole needs to be located a distance

$$D\left(\frac{\pi}{2}\right) = R\int_0^{\pi/2}\sec\theta\,d\theta = \infty$$

away from the equator. It is off the map!

f) If an inch of map distance at 40° N represents A earth-miles and an inch at 50° N represents B miles, then

$$\frac{A}{B} = \frac{\sec 40^\circ}{\sec 50^\circ}$$

because of the "stretching" of earth distances as they are placed on the Mercator map. Thus $A \neq B$.

g) The scale changes continually along north-south lines of a Mercator map. Thus a mostly north-to-south distance like that from Chicago to New Orleans is quite hard to calculate from the map.

7. $$\int \sec x\,dx = \int \frac{dx}{\cos x} = \int \frac{dx}{\sin\left(x+\frac{\pi}{2}\right)}$$

$$= \int \frac{dx}{2\sin y\cos y}, \quad y = \frac{1}{2}\left(x+\frac{\pi}{2}\right)$$

$$= \int \frac{dx}{2\frac{\sin y}{\cos y}\cos^2 y} = \int \frac{\sec^2 y\,dy}{\tan y}$$

$$= \ln\tan y + c, \text{ done.}$$

8. a) Easier solution due to Ronald Shubert, Chairman, Department of Mathematics, Elizabethtown College, Elizabethtown, Pa.:

$$\tan\frac{A}{2} = \frac{\sin\frac{A}{2}}{\cos\frac{A}{2}} = \frac{2\sin^2\frac{A}{2}}{2\sin\frac{A}{2}\cos\frac{A}{2}} = \frac{1-\cos A}{\sin A}$$

$$= \csc A - \cot A,$$

and thus,

$$\tan\left(\frac{x}{2}+\frac{\pi}{4}\right) = \tan\frac{1}{2}\left(x+\frac{\pi}{2}\right)$$

$$= \csc\left(x+\frac{\pi}{2}\right) - \cot\left(x+\frac{\pi}{2}\right)$$

$$= \sec x + \tan x .$$

My own solution is longer: Recall that

$$(**)\quad \tan(A+B) = \frac{\tan A + \tan B}{1 - \tan A \tan B}$$

and

$$\tan\left(\frac{A}{2}\right) = \sqrt{\frac{1-\cos A}{1+\cos A}},$$

then

$$\tan\left(\frac{x}{2}+\frac{\pi}{4}\right) = \frac{\tan\frac{x}{2}+1}{1-\tan\frac{x}{2}\cdot 1}$$

$$= \frac{1+\sqrt{\frac{1-\cos x}{1+\cos x}}}{1-\sqrt{\frac{1-\cos x}{1+\cos x}}} \cdot \frac{1+\sqrt{\frac{1-\cos x}{1+\cos x}}}{1-\sqrt{\frac{1-\cos x}{1+\cos x}}}$$

$$= \frac{1+\frac{1-\cos x}{1+\cos x}+2\sqrt{\frac{1-\cos x}{1+\cos x}}}{1-\frac{1-\cos x}{1+\cos x}}$$

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

$$= \frac{2+2\cos x}{2\cos x} = \sec x + \tan x .$$

b) This is easy to see in the form

$$\cot\left(\frac{\pi}{4}-y\right) = \tan\left(\frac{\pi}{4}+y\right),\quad y = \frac{x}{2},$$

when the graphs are drawn. A proof will drag you into straight-forward use of formulas like (**) above in 8a.

9. $\int \tan x\,dx = \ldots.$

$$= \int \sin x \cos x\,(1+\sin^2 x+\sin^4 x+\sin^6 x+\ldots)dx.$$

Put in $y = \sin^2 x$:

$$= \frac{1}{2}\int (1+y+y^2+y^3+\ldots)dy$$

$$= \frac{1}{2}\left(c+y+\frac{y^2}{2}+\frac{y^3}{3}+\ldots\right)$$

$$= \frac{1}{2}\left(c+\sin^2 x+\frac{\sin^4 x}{2}+\frac{\sin^6 x}{3}+\ldots\right).$$

10. The integral is infinite and we get this absurd divergent series:

$$\int_0^x \cot x\,dx = \int \frac{\cos x}{\sin x}\,dx$$

$$= \int \frac{\cos x \sin x\,dx}{\sin^2 x}$$

$$= \int \frac{\cos x \sin x\,dx}{1-\cos^2 x}$$

$$= -\frac{1}{2}\int \frac{dy}{1-y} \quad \text{with } y = \cos^2 x$$

$$= -\frac{1}{2}\int (1+y+y^2+y^3+\ldots)dy$$

$$= -\frac{1}{2}\left(c+y+\frac{y^2}{2}+\frac{y^3}{3}+\frac{y^4}{4}+\ldots\right).$$

The left is $+\infty$ and the right side looks negative!

11. a) A term by term comparison with

$$y + y^3 + y^5 + y^7 + \ldots.$$

(a convergent geometric series with $y = \sin x < 1$) shows the convergence of Wallis' series.

b) The ratio test tells us to calculate, for our series

$$\sum_{n=0}^{\infty} \frac{y^{2n+1}}{2n+1},$$

$$R = \lim \left| \frac{y^{2n+3}/(2n+3)}{y^{2n+1}/(2n+1)} \right| = y^2.$$

Since R < 1, the series is convergent. Other tests may also be used.

c) The ratio test done in (b) tells us that the series converges when $y^2 < 1$ (i.e., $-1 < y < 1$) and diverges when $y^2 > 1$ (i.e., $y > 1$ or $y < -1$). For $y = 1$ we get a divergent series. For $y = -1$ we get a convergent alternating series. Thus we have convergence for $-1 \le y < 1$ exactly. Now, what x gives $-1 \le \sin x < 1$? All real x, except ....?

12. Computer results for x = 5°, 15°, 30°, 45°, 60°, 75°, and 80° are below.

	HIGHEST POWER OF SIN(X) INCLUDED IN PARTIAL SUM	PARTIAL SUM
x = 5° The correct integral is .08737743	1	0.08715574
	3	0.08737642
	5	0.08737743
x = 15° $\int$ = .26484224	1	0.25881904
	3	0.26459824
	5	0.26483692
	7	0.26484183
	9	0.26484221
x = 30° $\int$ = .54930614	1	0.49999999
	3	0.54166666
	5	0.54791666
	7	0.54903273
	9	0.54924975
	11	0.54929414
	13	0.54930392
	15	0.54930556
	17	0.54930601

	HIGHEST POWER OF SIN(X) INCLUDED IN PARTIAL SUM	PARTIAL SUM
x = 45° $\int$ = .88137358	1	0.70710678
	3	0.82495791
	5	0.86031325
	7	0.87294015
	9	0.87785062
	11	0.87985944
	13	0.88070933
	15	0.88107761
	17	0.88124009
	19	0.88131278
	21	0.88134566
	23	0.88136067
	25	0.88136758
	27	0.88137078
	29	0.88137226
	31	0.88137296
	33	0.88137329
x = 60° $\int$ = 1.31695789	1	0.86602540
	11	1.28128220
	21	1.31169008
	31	1.31604511
	41	1.31678696
	51	1.31692435
	61	1.31695109
	71	1.31695647
	79	1.31695747
x = 75° $\int$ = 2.02758941	1	0.96592582
	51	1.99537299
	101	2.02425779
	151	2.02717058
	201	2.02753181
	251	2.02758099
	301	2.02758805
	335	2.02758893

HIGHEST POWER OF SIN(X) INCLUDED IN PARTIAL SUM	PARTIAL SUM
1	0.98480775
51	2.27986109
101	2.39058930
151	2.42057205
201	2.43044352
251	2.43400011
301	2.43535054
351	2.43588126
401	2.43609499
451	2.43618263
501	2.43621907
551	2.43623439
601	2.43624088
651	2.43624365
701	2.43624482
751	2.43624532
801	2.43624552
811	2.43624554

$x = 80^\circ$

$\int = 2.43624604$

For x nearing 90°, sin x near 1, we need quite a few terms in a partial sum to get good accuracy!

13. a) $$\arctan x = \int_0^x \frac{1}{1+x^2}\,dx$$
$$= \int_0^x (1 - x^2 + x^4 - x^6 + x^8 - + \cdots)\,dx$$
$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + - \cdots$$

b) use the ratio test:
$$R = \lim \left| \frac{x^{2n+3}/(2n+3)}{x^{2n+1}/(2n+1)} \right| = x^2$$

Thus the series converges when $x^2 < 1$, diverges when $x^2 > 1$. For $x^2 = 1$ ($x = \pm 1$) we get a convergent alternating series. Thus the series converges for $-1 \le x \le 1$.

d) Plug in $-x$ for x and show $\arctan(-x) = -\arctan(x)$.

9. SPECIAL ASSISTANCE SUPPLEMENT

[*S-1*]

The system of latitude and longitude is based on the geographic north and south poles, but compass needles do not usually point to the geographic north pole. Instead they point to the north magnetic pole at 75° N. latitude, 101° W. longitude in far northern Canada, north of the Dakotas. Navigators are used to correcting compass readings for that discrepancy. Thus we will speak conveniently of compass north as the north pole, 90° N. when that is not true.

Compasses, in fact, must also be corrected for deviations due to the magnetic iron in a ships hull or cargo holds and even for iron ores in nearby land masses. The Encyclopedia Brittanica article under "compass" discusses this in more detail.

[*S-2*]

A *great circle* on a sphere is a circle of the largest possible circumference, like the Equator. All the meridians (north-south lines of constant longitude) are also great circles. The circles of constant latitude are (except for the equator, latitude 0°) not great circles because these east-west directed circles get smaller in size as we progress from the equator toward either pole.

Notice that there is a full set of meridian great circles reaching from the North Pole to all the other points on the sphere, and that all of them also pass through the South Pole, opposite the North Pole on the Sphere. Thus there are infinitely many great circles routes between the two poles. If we need to connect the North Pole N with any other point P except the South Pole, there is a unique great circle passing through N and P. The shorter arc between N and P along that unique great circle is the shortest path on the globe connecting N and P, the *great circle route* between them.

Similarly, there is a full set of great circles through any point Q on the globe. Between Q and its opposite point R there are infinitely many great circle routes. Connecting Q and any other point R (not opposite to Q) on the sphere, there is a unique great circle and along that circle lies the great circle route, again the shortest between Q and R.

The Project would like to thank Marjorie A. Fitting of San Jose State University, Barbara Juister of Elgin Community College, Roland Smith of Russell Sage College, and L.M. Larsen of Kearney State College for their reviews, and all others who assisted in the production of this unit.

This unit was field-tested and/or student reviewed in prelimary form by Alan Shuchat, Wellesley College, Wellesley, Massachusetts; Peter Nicholls, Northern Illinois University, DeKalb, Illinois; Richard G. Montgomery, Southern Oregon State College, Ashland, Oregon; Kurt Kreith, University of California at Davis; Robert L. Baker, Jr., U.S. Naval Academy, Annapolis, Maryland; James Bradley, Roberts Wesleyan College, Rochester, New York; and Jonathan Choate, The Groton School, Groton, Massachusetts, and has been revised on the basis of data received from these sites.

umap

UNIT 207

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

MANAGEMENT OF A BUFFALO HERD

by Philip M. Tuchinsky

AM = number of adult males
AF = number of adult females
YM, YF = numbers of male yearlings, female yearlings
CM, CF = numbers of male calves, female calves
QM' = number of adult males harvested "next year"
QF' = number of adult females harvested "next year"

$$\begin{pmatrix} AM' \\ AF' \\ YM' \\ YF' \\ CM' \\ CF' \end{pmatrix} = \begin{pmatrix} .95 & 0 & .75 & 0 & 0 & 0 \\ 0 & .95 & 0 & .75 & 0 & 0 \\ 0 & 0 & 0 & 0 & .6 & 0 \\ 0 & 0 & 0 & 0 & 0 & .6 \\ 0 & .48 & 0 & 0 & 0 & 0 \\ 0 & .42 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} AM \\ AF \\ YM \\ YE \\ CM \\ CF \end{pmatrix} - \begin{pmatrix} QM' \\ QF' \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

APPLICATIONS OF LINEAR ALGEBRA TO HARVESTING

(LESLIE-TYPE MODEL)

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

MANAGEMENT OF A BUFFALO HERD

by

Philip M. Tuchinsky
Department of Mathematical Sciences
Ohio Wesleyan University
Delaware, Ohio 43015

TABLE OF CONTENTS

Intermodular Description Sheet: *UMAP Unit 207*

Title: MANAGEMENT OF A BUFFALO HERD

Author: Philip M. Tuchinsky
Department of Mathematical Sciences
Ohio Wesleyan University
Delaware, Ohio 43015

Review Stage/Date: III 12/28/77

Classification: APPL LIN ALG/HARVESTING/LESLIE-TYPE MODEL

Suggested Support Material: Key exercises call for computer use.

References: See Section 6 of text.

Prerequisite Skills:

1. Matrix multiplication; matrix inverses; calculation of the inverse by row (Gaussian) elimination (optional); overdetermined linear equations; elementary matrix algebra.
2. No eigentheory is used. No background in biology/ecology/ranching is needed.

Output Skills:

1. Calculate with linear difference equations on computer.
2. Identify overdetermined linear equations and decide when they have a solution.
3. Set up and solve matrix equations from word problems.
4. Sum finite geometric series for matrix case.
5. Describe an application that uses linear equations to model birth, aging and death in a population. Specifically, detail use of a matrix to transform that population through time.
6. Differentiate between matrix level and entry level calculations and give examples where both are helpful.
7. Explain context where polynomial functions of a matrix inevitably arise.
8. Discuss major strengths and weaknesses of a linear model in a nonlinear reality.
9. Simulate several policies of harvesting by making varied use of a computer simulation.

Predicted Teaching Time: 2-3 class periods, including discussion of computer project results. This assumes that class time is mostly related to the math and students read the biological content for themselves.

Suggested Uses: Sections 4 and 5 are independent of each other; either can be done first. Sections 4.6 and 5.2 are harder than the other three application examples in 4.1, 4.3, 5.1. Section 4.2 may be omitted.

A wide range of basic linear algebra skills can be tied together by working through this module. Computer experience with a linear transformation is a key benefit of this module -- if at all possible, I recommend use of some of exercises 3 - 8, Section 2.6

This module is suitable for a first linear algebra course or a post-linear-algebra course in mathematical modeling. It is suitable for presentation by advanced students in a seminar.

The matrix is not diagonalizable: To pursue the calculations in Section 4.5 further, the natural path is to seek the eigenvalues of matrix M. But the more general

$$M_1 = \begin{pmatrix} a & 0 & b & 0 & 0 & 0 \\ 0 & a & 0 & b & 0 & 0 \\ 0 & 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 0 & c \\ 0 & d & 0 & 0 & 0 & 0 \\ 0 & e & 0 & 0 & 0 & 0 \end{pmatrix} \text{ has Jordan form } \left(\begin{array}{cc|cccc} 0 & 1 & & 0 & & \\ 0 & 0 & & & & \\ \hline & & a & & & \\ & 0 & & x & & \\ & & & & u & \\ & & & & & \bar{u} \end{array}\right),$$

where a,b,c,d,e are all $\in (0,1)$, and the eigenvalues are zero (twice), parameter a, real $x>a$, and a complex conjugate pair u, $\bar{u}$. The characteristic equation that yields these roots is:

$$\det(M_1 - \lambda I) = \lambda^2(a - \lambda)(-\lambda^3 + a\lambda^2 + bce) = 0$$

The eigenspace of zero is unfortunately one-dimensional; thus the Jordan form above. The square, cube, and higher powers of this Jordan form are diagonal.

© 1977 EDC/Project UMAP
All Rights Reserved.

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 Gardunkel	Associate Director/Consortium Coordinator
Felicia Weitzel	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Dianne Lally	Project Secretary
Paula M. Santillo	Secretary/Financial Assistant
Jack Alexander	Editorial Consultant
Edwina Michener	Editorial Consultant

NATIONAL STEERING COMMITTEE

W.T. Martin	MIT (Chairman)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
James D. Forman	Rochester Institute of Technology
Ernest J. Henley	University of Houston
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
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 Edward L. Keller for his review 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.

1. INTRODUCTION

1.1 What Harvest Should You Take?

Imagine yourself as the operator of a buffalo ranch.[1] You have a certain herd on hand, and each year you "harvest" a number of mature buffalo for their meat. You permit the remaining herd, for the next year, to replenish itself through its own natural breeding. The herd has a certain known structure: it is made up of known proportions of adult vs. immature animals, of females vs. males. Here are some questions you might ask while simultaneously trying to gain a good harvest and maintain the herd for good future harvests:

-- What harvest policy will lead to a herd next year that has the same size and structure as this year's herd?

-- What annual harvest will permit the herd to grow steadily so that in ten years it will have doubled in size while keeping the same proportional structure?

-- Do substantially different future herds result if more, the same number, or less females are harvested than males?

1.2 What Herd Should You Start With?

Next, imagine yourself as planning to enter the buffalo-ranching industry. You set goals (based on your costs, capital, desired income, etc.) for a desired harvest. That is, you select, as a basic parameter of your business, a number of mature animals that you intend to harvest each year. You might ask:

[1] Although buffalo management is not a major industry, this paper is developed in terms of it because a widely available computer program named BUFLO is based on the same model. The methods discussed here are the subject of research in human population dynamics; cattle, sheep, and other ranching industries; forest, fishing, and wildlife management. See the references.

-- What initial size and structure of herd will provide the desired harvest?

-- How should the quota be distributed among male and female animals to achieve a herd of smallest size that will continue to yield the quota?

1.3 Wildlife Management

Finally, imagine yourself as the manager of a game preserve. Conditions here are quite unlike those on a ranch because the buffalo herd lives among its natural predators, such as the wolf. You have a limited amount of land, and its vegetation must support the herd. What quotas of male and female buffalo should you license hunters to kill each year to maintain the herd at an appropriate size?

1.4 The Task Ahead

In this paper we will consider a mathematical model -- based on linear algebra -- of a buffalo herd. It will be possible to answer the questions above using the model, but the model is a much simplified version of the situation in nature. We will consider the underlying assumptions of the model and their limitations to some extent.

While we will look at the model mostly as a management tool, we will also be in a position (in the exercises in Section 2.6) to study historical issues concerning the destruction of the vast U.S. buffalo population that thrived on the Great Plains in the early 1800's.

2. THE MODEL

2.1 Herd Components and Their Survival Rates

We consider[2] six categories of buffalo within the herd: *calves* are in their first year of life, *yearlings* in their second, and all older buffalo are adults. Each age group is broken down in male and female categories.

[2]The model has been taken from computer program BUFLO. See the references for full acknowledgement.

Each 100 adult cows will bear approximately 48 male calves and 42 female calves each year in late spring. This 90% reproduction rate is almost unrelated to the number of adult males in the herd because male buffalo are polygamous.

Buffalo naturally suffer different death rates at different ages.[3] Because of deaths at birth and such natural enemies as the wolf and coyote, only about 60% of the calves survive to become next year's yearlings and about 75% of the yearlings become adults. Once they reach maturity, buffalo are quite safe from their enemies until they weaken from illness, injury, or old age: 95% of the adults survive from each year to the next. We will take these numbers to be the same for males and females and the same year after year.

2.2 Basic Model Equations

It is easy to organize this data into a mathematical model. Let

(1)
AM = number of adult males
AF = number of adult females
YM,YF = numbers of male yearlings, female yearlings
CM,CF = numbers of male calves, female calves

or more specifically, let these be the numbers of buffalo at the end of "this year" just after the harvest. Let AM', AF',..., CF' be the comparable counts for next year's herd, also at the completion of (next year's) harvest. Let

QM' = number of adult males harvested "next year"
QF' = number of adult females harvested "next year"
(It is our policy to harvest only adult buffalo.)

Then, the breeding process, followed by harvest, is contained in these equations:

[3]This will be true in the wild. On a ranch, survival to adulthood would be more likely. Data in this paragraph applies to wildlife. See Sections 2.5 and 3.1. The references offer similar data due to Fuller.

$$
(2)\quad
\begin{aligned}
AM' &= .95\ AM + .75\ YM - QM' \\
AF' &= .95\ AF + .75\ YF - QF' \\
YM' &= .60\ CM \\
YF' &= .60\ CF \\
CM' &= .48\ AF \\
CF' &= .42\ AF
\end{aligned}
$$

Let's read these equations in detail. Because of natural deaths, .95 AM represents the survivors next year among this year's adult males and .75 YM is the number of this year's yearlings who survive to become adult males. Thus .95 AM + .75 YM represents the total of adult males just *before* harvest next year. (We imagine that the harvest takes place at one specific moment each year, perhaps on a day in the fall.) Thus the *after*-harvest total, AM', is correctly given in the first equation of (2). The second equation treats the adult females similarly. The third and fourth equations say that 60% of this year's calves survive to become next year's yearlings. The last two equations say that, for each hundred adult cows after harvest this year, the herd will grow by 48 male calves and 42 female calves to be born next year.

2.3 The After-Harvest Model in Vector and Matrix Notation

Now label "this year" as "year 0", "next year" as 1 and so on. Define the vectors:

$\vec{G}_j$ = herd structure after harvest in the j^{th} year ($j = 0, 1, 2, 3, \ldots$)

In our earlier notation, the first of these six-dimensional vectors are:

$$\vec{G}_0 = \begin{pmatrix} AM \\ AF \\ YM \\ YF \\ CM \\ CF \end{pmatrix} \qquad \vec{G}_1 = \begin{pmatrix} AM' \\ AF' \\ YM' \\ YF' \\ CM' \\ CF' \end{pmatrix}$$

We must gather the harvest quotas into vectors, too. Put

$$\vec{Q}_1 = \begin{pmatrix} QM' \\ QF' \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

as an example of

$\vec{Q}_j$ = harvest in j^{th} year (last four entries are always zero).

With this notation established, it is time to rewrite (2) as

$$\begin{pmatrix} AM' \\ AF' \\ YM' \\ YF' \\ CM' \\ CF' \end{pmatrix} = \begin{pmatrix} .95 & 0 & .75 & 0 & 0 & 0 \\ 0 & .95 & 0 & .75 & 0 & 0 \\ 0 & 0 & 0 & 0 & .6 & 0 \\ 0 & 0 & 0 & 0 & 0 & .6 \\ 0 & .48 & 0 & 0 & 0 & 0 \\ 0 & .42 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} AM \\ AF \\ YM \\ YF \\ CM \\ CF \end{pmatrix} - \begin{pmatrix} QM'' \\ QF'' \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

or

$$(3)\quad \vec{G}_1 = M\,\vec{G}_0 - \vec{Q}_1$$

where M is the 6 x 6 matrix just above. The year-to-year process is given more generally as

$$(4)\quad \vec{G}_{j+1} = M\,\vec{G}_j - Q_{j+1}, \quad j = 0, 1, 2, 3, \ldots$$

This is the *after harvest model*, because it involves herd counts $\vec{G}_j$ taken just after the harvest is completed.

We may call M the *transformation matrix* of our model, for it transforms its input, the herd structure just after harvest, into the herd structure that birth, aging and death will produce just before harvest in the following year.

2.4 The Before-Harvest Model

The model just discussed is useful if we have a herd and want to examine what next year's harvest will give us as a new herd. But suppose we are trying to select this year's harvest so that next year's herd can be studied. Then we want *before*-harvest herd counts to which we can apply the harvest. They deserve a notation of their own:

$\vec{H}_j$ = herd *before* harvest in the j^{th} year, $j = 0, 1, 2, \ldots$

Thus $\vec{G}_j = \vec{H}_j - \vec{Q}_j$ and the last paragraph of Section 2.3 says that $\vec{H}_{j+1} = M\vec{G}_j$, $j = 0, 1, 2, \ldots$.

The *before-harvest-model* relates H_j to H_{j+1}. Clearly, H_j is diminished by Q_j at harvest and the new herd $H_j - Q_j$ undergoes the breeding transformation. Thus

(5) $H_{j+1} = M(H_j - Q_j)$, $j = 0, 1, 2, 3, \ldots$

2.5 Survival Rates Would Be Larger on a Ranch

One more comment. The birth and death rates were given for a herd living in the wild, subject to its natural predators. (The effects of man as a predator are reflected in QM and QF, not the given percentages.) Most of our effort, however, will be with questions that relate to ranching, where herds are fenced and natural predators almost absent. We would expect much larger fractions of each category to survive the year. However, there are no accepted numbers to use in M, and, rather than arbitrarily pick some, we will use the same matrix M for both wilderness and ranch applications. The results will be *qualitatively* the same for higher survival rates (as the author has checked in some detail).

2.6 Exercises and Computer Projects

1. In the week before harvest last year your ranch had a buffalo herd with this structure:

AM = 200	YM = 300	CM = 520
AF = 1000	YF = 300	CF = 500

Your harvesting policy each year is to take 100 adult males and 200 adult females. Calculate the structure of the herd

a. after last year's harvest
b. before this year's harvest
c. after this year's harvest
d. before next year's harvest
e. after next year's harvest

f. Is the herd likely to grow or shrink if you continue this policy into the future, or can't you tell? Justify your answer.

2. Use equation (4) repeatedly to show that, over several years involving different harvests, an initial herd G_0 will transform into

$$\vec{G}_1 = M\vec{G}_0 - \vec{Q}_1$$
$$\vec{G}_2 = M^2\vec{G}_0 - \vec{Q}_2 - M\vec{Q}_1$$
$$\vec{G}_3 = M^3\vec{G}_0 - \vec{Q}_3 - M\vec{Q}_2 - M\vec{Q}_1, \text{ etc.}$$

Now, provide a biological meaning for each term in the equations. For example, $M^2\vec{G}_0$ is the herd that results after two years if no harvests are taken. The other terms in the second equation correct this to account for the harvests. What does the $M\vec{Q}_1$ term mean?

The remaining problems in this section call for the use of a computer.

3. Write a computer program that will calculate next year's herd size from this year's, using the after-harvest model. It should receive as inputs: (1) the initial herd structure $\vec{G}_0$; (2) the constant harvest; (3) the number of years the herd is to be studied. The program should loop to calculate the herd size year by year for the number of years requested. It should print out the successive years and the herd structure that would result for that year, using our model.[4]

4. *The U.S. Buffalo Herd in 1830.* The authors of the BUFLO computer program (from which our model is taken; see the references) state that the total buffalo herd in the United States in 1830 consisted of 60 million animals distributed as follows:

---

[4] Beware of this trap as you work your program: if you compute the components of next year's herd in their usual order, a *new* value of AF will be computed before the *old* value can be used to calculate CM and CF. The old value of AF must be saved before it is replaced with the new one.

ERIC Full Text Provided by ERIC

30% male adults 27% female adults
9% male yearlings 8% female yearlings
14% male calves 12% female calves

(These figures should be taken as good historic guesses; estimates of the total herd size vary widely above and below 60 million.) Let this data give your initial herd. Take a constant harvest of 4 million animals annually for a period of ten years. Distribute that harvest in various ways among males and females, trying to find a harvest that leaves the herd approximately unchanged after ten years. That is, split the harvest among male and female adults in a specific way and trace the herd for ten years using your computer program for Exercise 3. Then try other splittings of the harvest in the same way. Several computer runs can be used or you can loop. A convenient way to get the number 60 million into the machine is 60.E6 in Fortran or Basic.

5. Start with you computer program from Exercise 3 and the initial herd given in Exercise 4. Take a 4 million animal harvest annually for twenty years, using these strategies:

   a. harvest 100% adult males
   b. harvest 75% adult males, 25% females
   c. harvest 50% adult males, 50% females
   d. harvest 25% adult males, 75% females
   e. harvest 100% adult females.

   The results are strikingly different. Discuss the biological reasons.

6. Repeat Exercise 5, taking a much larger harvest (say 12 million animals) annually. Compare to other results you have.

7. Let's examine the effects of a natural catastrophe (flood, range fire, etc.) on a herd. Take the initial herd from Exercise 4 again, and set the constant annual harvest to zero. Drastically reduce the birth and survival rates in the matrix M and transform the herd forward for one year, to simulate a catastrophe. Now put our usual numbers back in M and trace the herd forward for nine more years, still taking no harvest. What are the long-term effects of the catastrophe?

8. Repeat Exercise 7, but this time take constant annual harvests (in the catastrophic year and the others) of 1 million or 4 million animals, splitting the harvest among males and females in the ways listed in Exercise 5. Comment on the combined effects of catastrophe and harvest. Which harvests worsen the effects of the catastrophe? Which overcome it?

3. ASSUMPTIONS, STRENGTHS, AND WEAKNESSES OF THE MODEL

3.1 The Model's Basic Strengths

The examples in Sections 4 and 5 will show that we can really calculate with this model; it is a workable management tool. It does reflect the basic processes of birth, aging, and death among buffalo. The equations in Sections 2, 4 and 5 all have reasonable biological or economic interpretations.

The actual numbers used as birth and survival rates are reasonably close to correct figures. One piece of evidence for this is that, among adult buffalo, a life span of approximately 25 years was the rule[5] at the time when great wild herds roamed the plains. Our model predicts an average life span of 21.5 years (where we count buffalo that die between their 2nd and 3rd birthdays as age 2½, etc.)

In Example 3, Section 4.6, we will show that no more than about 14% of a herd may be harvested annually without eventually depleting the herd. This value would vary in nature, but the model is qualitatively correct enough to convince me that a steady harvest of (say) 20% of the herd would destroy the herd in time. Exercises 5 and 6 provide strong evidence of this.

[5] See E.D. Branch, The Hunting of the Buffalo, University of Nebraska Press, 1962, p. 11. Branch's figure of 25 is presumably drawn from journals of the 1800's and may well be high.

ERIC

3.2 Limitations of the Model

Whether one should blindly accept advice from the model is another matter. The model is built on a number of assumptions that do not correspond to nature. The most important of these is that the birth and survival rates used in M are assumed to be constant year after year; this would not be true in nature. We can regard the survival rates in M as averages for "normal" years that provide generally favorable weather and feeding conditions. Abnormal conditions like severe storms, range fires, drought, floods, and disease might temporarily cause much lower birth and survival rates. Our model does not provide for such catastrophes,[6] although they might not be rare in the wild or on a ranch.

The constant birth and survival rates do not permit the study of overcrowding or overpopulation. Instead, the model assumes that unlimited land, food and water are available for the herd. In the wild, an overpopulated herd would eat poorly and its birth and survival rates would decrease. It would be more subject to disease. On a well-run ranch we would not expect overpopulation. We will see in Sections 4 and 5 that the herd size can be related to the harvest in ways that make overpopulation manageable. In any case, the model is one of unlimited exponential growth for the herd, tempered by the harvesting process.

Another weakness of the model is that harvesting is done only once a year, rather than steadily or several times yearly. Reality was different: Plains Indian tribes held lengthy summer and winter buffalo hunts. White men slaughtered the buffalo continually in the 1800's. On a ranch today, the herd would be thinned as meat prices and the availability of rangeland and water dictate.

[6] We have considered an obvious way to simulate a catastrophe with the model in Exercises 7 and 8, Section 2.6.

Yet another weakness is that no economics is included in the model. The actual quotas harvested would surely be related to the price of meat and the cost of feeding the herd on any ranch. The manager of a game preserve might not be troubled by such questions (if his grazing lands are sufficient for the herd so that no feed is to be purchased). There is no single obvious way to extend the model so that economics is effectively included.

The breeding mechanisms of the model are not ideal. In fact, buffalo begin to reproduce at ages two or three; we have assumed that all two-year-olds are full adults. And the number of calves born has been made a simple fraction of the number of adult females. This is roughly true in a polygamously mating herd if reasonable numbers of adult bulls are in the herd. In our model, a value AM = 0 would not interrupt the mating process, as it would in nature. In fact, the actual herd would be in danger of extinction if any of the six categories grew too small. This can not be included in a linear model. In using the model, one could declare the herd "extinct" if any category were to grow too small.

Finally, we have lumped all adult buffalo into two categories and declared them all equal in their ability to survive and breed, ignoring the obvious variations with age.

Despite all these defects, and others that I've undoubtedly missed, the model as presented offers a useful simplification of the herd. Let's put it to use.

4. APPLICATIONS: ESTABLISHING A HERD

4.1 A Herd and Harvest That Continue Year After Year

Example 1. What size and structure of herd $\vec{G}_0$ must we have (or put together) this year so that next year we may take a pre-chosen harvest $\vec{Q}_1$ and then have a herd $\vec{G}_1$ such that $\vec{G}_1 = \vec{G}_0$?

A businessman planning to create a ranch might ask this question. He chooses his annual "product" $\vec{Q}_1$ and wants to know what "capital investment" $\vec{G}_0$ he should make so that it will maintain itself from year to year ($\vec{G}_1 = \vec{G}_0$) and yield product $\vec{Q}_1$. Since we end up with $\vec{G}_1 = \vec{G}_0$, the process of harvesting $\vec{Q}_1$ and maintaining a herd of the same size and structure can continue year after year: we call the herd and harvest vectors *steady state*.

As the chosen notation indicates, it is natural to use the *after-harvest-count* $\vec{G}_0$, $\vec{G}_1$ for the herd, because the year-long study-period for the herd progresses from initial herd through the breeding process to the pre-set harvest at the end of the period.

Thus we know $\vec{Q}_1$ and want to solve for $\vec{G}_0$ in

$$(6)\qquad \begin{aligned} \vec{G}_1 &= \vec{G}_0 \\ \vec{G}_1 &= M\vec{G}_0 - \vec{Q}_1 \quad \text{[compare (3)]}. \end{aligned}$$

We can replace $\vec{G}_1$ with $\vec{G}_0$ in the second equation of (6), getting

$$\vec{G}_0 = M\vec{G}_0 - \vec{Q}_1$$

and rearrange to read (I is the 6 x 6 identity matrix)

$$(7)\qquad (M-I)\vec{G}_0 = \vec{Q}_1.$$

This is a set of six linear equations for the unknowns $\vec{G}_0$; M-I and $\vec{Q}_1$ are known. In fact we are asked to solve

$$(M-I)\vec{G}_0 = \begin{bmatrix} -.05 & 0 & .75 & 0 & 0 & 0 \\ 0 & -.05 & 0 & .75 & 0 & 0 \\ 0 & 0 & -1 & 0 & .6 & 0 \\ 0 & 0 & 0 & -1 & 0 & .6 \\ 0 & .48 & 0 & 0 & -1 & 0 \\ 0 & .42 & 0 & 0 & 0 & -1 \end{bmatrix} \vec{G}_0 = \vec{Q}_1 .$$

There is a unique solution because M-I is non-singular. We will calculate $(M-I)^{-1}$ in Section 4.2 below. In terms of it we can write our solution to

to the problem posed in Example 1 as

$$(9)\qquad \vec{G}_0 = (M-I)^{-1}\,\vec{Q}_1$$

Notice that we've completely solved the problem at *matrix level*: we can write the solution in (9) without actually looking at any of the specific numerical entries of M; we use M as a single item, not a collection of 36 numbers. However, we do have to use the entries of M to establish that $(M-I)^{-1}$ exists and to actually *calculate* the solution in (9): that work is at *entry level*, not matrix level.

4.2 Calculation of $(M-I)^{-1}$

We would need $(M-I)^{-1}$ to proceed further with (9), so we have the opportunity to carry through an unpleasant matrix pivoting Gaussian-elimination calculation by hand, in detail.

The reader who would benefit from such an example is invited to follow along, *electronic calculator in hand*, verifying each step. The reader who prefers to see how the answer is used in the rest of this section is welcome to do so: skip to the paragraph containing equation (10) at the end of this section.

Recall that, to find the inverse[7], we list M-I and adjoin to it a six-by-six identity matrix to create a 6 x 12 matrix:

$$\left[\begin{array}{cccccc|cccccc} -.05 & 0 & .75 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -.05 & 0 & .75 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & .48 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & .42 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Now we reduce the left side to a six-by-six identity matrix using only *elementary row operations*: we may (1) multiply a row (all 12 columns) by a non-zero

[7] There are other, less efficient methods.

scalar, or (2) add a scalar multiple of one row to another row, or (3) interchange any two rows. To work now: multiply the top two rows by -20 each (to convert the -.05's into 1's for the 6 x 6 I). Get

check these rows (rows 1 and 2)

$$\left(\begin{array}{cccccc|cccccc}
1 & 0 & -15 & 0 & 0 & 0 & -20 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & -15 & 0 & 0 & 0 & -20 & 0 & 0 & 0 & 0\\
0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0\\
0 & .48 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & .42 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right).$$

The first column on the left is fine. Make the second column fit the goal of a 6 x 6 I by subtracting .48 times row 2 from row 5, and .42 of row 2 from row 6. These two elementary row operations give us

check these rows (rows 5 and 6)

$$\left(\begin{array}{cccccc|cccccc}
1 & 0 & -15 & 0 & 0 & 0 & -20 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & -15 & 0 & 0 & 0 & -20 & 0 & 0 & 0 & 0\\
0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 7.2 & -1 & 0 & 0 & 9.6 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 6.3 & 0 & -1 & 0 & 8.4 & 0 & 0 & 0 & 1
\end{array}\right).$$

Multiply row 3 by -1 and use that *new* row 3 to kill the -15 (in the 1,3 slot) by adding 15 of the new row 3 to row 1:

check these rows (rows 1 and 3)

$$\left(\begin{array}{cccccc|cccccc}
1 & 0 & 0 & 0 & -9 & 0 & -20 & 0 & -15 & 0 & 0 & 0\\
0 & 1 & 0 & -15 & 0 & 0 & 0 & -20 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & -.6 & 0 & 0 & 0 & -1 & 0 & 0 & 0\\
0 & 0 & 0 & -1 & 0 & .6 & 0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 7.2 & -1 & 0 & 0 & 9.6 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 6.3 & 0 & -1 & 0 & 8.4 & 0 & 0 & 0 & 1
\end{array}\right).$$

The first three columns now match a 6 x 6 I. Please notice that what we are about to do in column 4 does not disturb the first three columns. We gain this because we work from left to right, leaving friendly zeros behind. Multiply row 4 by -1 to get a new row 4. Add appropriate multiples of this new row 4 to rows 2, 5, and 6 so that the rest of column 4 is zeroed. Reach

new row 4 gotten first

$$\left(\begin{array}{cccccc|cccccc}
1 & 0 & 0 & 0 & -9 & 0 & -20 & 0 & -15 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & -9 & 0 & -20 & 0 & -15 & 0 & 0\\
0 & 0 & 1 & 0 & -.6 & 0 & 0 & 0 & -1 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & -.6 & 0 & 0 & 0 & -1 & 0 & 0\\
0 & 0 & 0 & 0 & -1 & 4.32 & 0 & 9.6 & 0 & 7.2 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 2.78 & 0 & 8.4 & 0 & 6.3 & 0 & 1
\end{array}\right)$$

You should be able to decide how we get to the next matrix. The result is:

$$\left(\begin{array}{cccccc|cccccc}
0 & 0 & 0 & 0 & 0 & -38.88 & -20 & -86.4 & -15 & -64.8 & -9 & 0\\
0 & 1 & 0 & 0 & 0 & -9 & 0 & -20 & 0 & -15 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & -2.592 & 0 & -5.76 & -1 & -4.32 & -.6 & 0\\
0 & 0 & 0 & 1 & 0 & -.6 & 0 & 0 & 0 & -1 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & -4.32 & 0 & -9.6 & 0 & -7.2 & -1 & 0\\
0 & 0 & 0 & 0 & 0 & 2.78 & 0 & 8.4 & 0 & 6.3 & 0 & 1
\end{array}\right).$$

Finally we multiply the 6th row by $\frac{1}{2.78}$ and clear the sixth column to reach

$$\left(\begin{array}{c|cccccc}
 & -20 & 31.080 & -15 & 23.310 & -9 & 13.986\\
 & 0 & 7.1944 & 0 & 5.3958 & 0 & 3.2374\\
 & 0 & 2.0720 & -1 & 1.5540 & -.6 & .93237\\
I & 0 & 1.8130 & 0 & .35972 & 0 & .21583\\
 & 0 & 3.4533 & 0 & 2.5900 & -1 & 1.5539\\
 & 0 & 3.0216 & 0 & 2.2662 & 0 & .35971
\end{array}\right)$$

The matrix that has appeared on the right is $(M-I)^{-1}$. The first, third and fifth columns are exact and the others are correctly rounded to five significant digits, which is more than we can make good use of below. Keeping four significant digits, our final result for the inverse is:

$$(10)\qquad (M-I)^{-1} = \begin{pmatrix}
-20 & 31.08 & -15 & 23.31 & -9 & 13.99\\
0 & 7.194 & 0 & 5.396 & 0 & 3.237\\
0 & 2.072 & -1 & 1.554 & -.6 & .9324\\
0 & 1.813 & 0 & .3597 & 0 & .2158\\
0 & 3.453 & 0 & 2.590 & -1 & 1.554\\
0 & 3.022 & 0 & 2.266 & 0 & .3597
\end{pmatrix}.$$

4.3 A Steady Harvest Plus Controlled Growth of the Herd

Example 2. Our ranch-planning businessman now wants a herd that yields harvest $\vec{Q}$ next year (and every year thereafter) while it grows by 40% during the first two years. The larger herd is to have exactly the same proportional structure as the original one.

Again we regard $\vec{Q}$ as known and use the after-harvest model. After a year's growth and next year's harvest, initial herd $\vec{G}_0$ (which we will calculate) will become

(11a) $$\vec{G}_1 = M\vec{G}_0 - \vec{Q}.$$

The next year's growth and eventual harvest yields

(11b) $$\begin{aligned}\vec{G}_2 &= M\vec{G}_1 - \vec{Q} \quad \text{(same Q each year)}\\ &= M(M\vec{G}_0 - \vec{Q}) - \vec{Q}\\ &= M^2\vec{G}_0 - M\vec{Q} - \vec{Q}\end{aligned}$$

and we want 40% growth (plus the harvests) after two years:

(11c) $$\vec{G}_2 = (1.4)\vec{G}_0.$$

From (11 b,c) we conclude

$$M^2\vec{G}_0 - M\vec{Q} - \vec{Q} = 1.4\,\vec{G}_0$$

and we rearrange this to

(12) $$\underbrace{(M^2 - 1.4I)}_{\text{known 6x6 matrix}}\ \underbrace{\vec{G}_0}_{\text{unknown}} = \underbrace{(M + I)\,\vec{Q}}_{\text{a known vector}}$$

In (12) we have a set of 6 linear equations that have a unique solution. (We won't prove that $M^2 - 1.4I$ has an inverse, but it's true.) Our problem has this solution, written at matrix level:

(13) $$\vec{G}_0 = (M^2 - 1.4I)^{-1}\,(M + I)\,\vec{Q}$$

4.4 Exercises

9. a. Write equations comparable to (6) or (11a,b,c) for this situation: We are given an annual harvest Q. We want to choose the herd $\vec{G}_0$ so that, after growth and a harvest next year, we will have a herd that is 12% larger. It is to have the same structure as $\vec{G}_0$ (i.e. be $1.12\,\vec{G}_0$).

   b. Solve your equations from (a) at matrix level for $\vec{G}_0$.

10. A buffalo herd $\vec{G}_0$ will be allowed to grow until next year, when harvest $\vec{Q}$ will be taken. The resulting herd $\vec{G}_1$ will be allowed to grow another year, when a larger harvest $1.1\vec{Q}$ will be taken. Calculate $\vec{G}_0$ so that this process leads to a final resulting herd $\vec{G}_2$ such that $\vec{G}_2 = \vec{G}_0$.

11. From this year's herd $\vec{G}_0$ a harvest $\vec{Q}$ will be taken next year. After another year's growth, a harvest $1.2\,\vec{Q}$ will be taken. The final resulting herd $\vec{G}_2$ is to be 25% larger than $\vec{G}_0$ (i.e. $\vec{G}_2 = 1.25\,\vec{G}_0$).

    a. Write equations for this situation comparable to (11a,b,c).

    b. Solve for $\vec{G}_0$.

12. A herd $\vec{G}_0$ grows for five years with no harvest being taken. In the fifth year, harvest $\vec{Q}$ is subtracted. The resulting herd $\vec{G}_5$ is exactly double $\vec{G}_0$. Find $\vec{G}_0$.

13. Find $\vec{G}_0$ if, after 5 years during which the same known harvest $\vec{Q}$ is taken at the end of each year, the herd is to double: $\vec{G}_5 = 2\,\vec{G}_0$.

14. Find $\vec{G}_0$ if the herd is to double in six years ($\vec{G}_6 = 2\,\vec{G}_0$). Assume that the same known harvest Q is taken after the second, fourth, and sixth years of growth.

4.5 Mathematical Insights

The example of Section 4.1 and 4.3 and the exercises of 4.4 should have provided you with experience that makes these comments believable:

a. When calculating with matrices, we find that algebra arises that is much like the algebra we learned long ago for numbers. Most of what we can do with numbers is also correct for matrices. (Key exception: matrix multiplication is not commutative.) We can even sum geometric series -- see Section 5.2 below. *It pays* *to think of a matrix as the analog of a single number.*

b. We may naturally need to calculate high powers (like M^{10}) of matrices. An easier way to do this would be very welcome. There is one: when you learn about "eigenvalues and eigenvectors" you will see that technique.

c. Expressions like $M^2 - 1.4I$, $M + I$, $M^9 + M^8 + M^7 + \dots + M^2 + M + I$ (see Section 5.2), called polynomials in the (square) matrix M, enter our work in a natural way and are worth study. They are polynomials in M in the same sense that $4x^2 - 3x + 5$ is a polynomial in x, i.e., they are sums of integer powers of M, or, equivalently, linear combinations of $I = M^0, M, M^2, M^3$, etc.

d. All our calculations in the example were at matrix level, and at that level we got a lot done. But further progress with expressions like (9) or (13) requires that we go to entry level (equation level). Matrix algebra is a powerful tool, but by dealing with the matrix as a whole we are out of touch with the individual entries, and their information may be critical.

4.6 An Efficiently Small Herd

Example 3. For any specified harvest quotas QM and QF, we have found an appropriate steady-state herd (which will yield those quotas) in Example 1. But perhaps our real goal is simply to harvest T animals, with T = QM + QF. Naturally, we wish to do this with the smallest possible herd (which would require the least land, feed, fencing, handling by employees, paperwork, etc.) Is there some way to split up T into QM and QF so that the herd is smallest?

We set up the algebra in this way: QM will be some fraction of T, say $QM = p \cdot T$ where $0 \le p \le 1$. Similarly, $QF = q \cdot T$ with $0 \le q \le 1$. Since T = QM + QF, $p + q = 1$.

(For example, if we end up selecting a harvest of 75% males and 25% females, p = .75, q = .25.) We wish to choose p and q.

Thus, in (9), using new scalars p, q and T, the harvest is

$$(14) \qquad \vec{Q}_1 = \begin{bmatrix} p \\ q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} T \quad \text{and} \quad \vec{G}_0 = (M - I)^{-1} \begin{bmatrix} p \\ q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} T.$$

The herd $\vec{G}_0$ is now a multiple of the total harvest T. We can think of

$$(15) \qquad (M - I)^{-1} \begin{bmatrix} p \\ q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

as the "herd structure per animal harvested" or the mini-herd needed to produce one harvested animal, because when multiplied by T, it becomes the total herd $\vec{G}_0$.

The *herd size* (the total number of animals in the herd) for a herd $\vec{G}_0$ will be

$$(16) \qquad (1,1,1,1,1,1) \cdot \vec{G}_0 = (1,1,1,1,1,1) \cdot (M-I)^{-1} \begin{bmatrix} p \\ q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

because multiplying by this vector (1,1,1,1,1,1) simply adds up the entries in $\vec{G}_0$. Since this is a multiple of T, we simplify by studying

HS = "herd size per animal harvested"

= "herd size"/T

$$= (1,1,1,1,1,1)(M-I)^{-1} \begin{bmatrix} p \\ q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Again: our goal is to select p and q to make HS as small as possible.

To this point, we have dealt at matrix level, aside from setting up $\vec{Q}_1$ with scalars p, q, and T. From here we must work at entry level, calculating the individual equations. We plug in $(M-I)^{-1}$ from (10), Section 4.2, and calculate

$$\text{(17)} \quad HS = (1,1,1,1,1,1)\begin{pmatrix} -20\,p + 31.08q \\ 7.194q \\ 2.072q \\ 1.813q \\ 3.453q \\ 3.022q \end{pmatrix} \doteq -20\,p + 48.63q.$$

Here we have rounded to two decimal places.

The goal was to select p and q such that $0 \le p \le 1$, $0 \le q \le 1$, $p + q = 1$, and HS is minimal. That's easy: as p increases, q must decrease and HS grows steadily smaller; thus, p = 1, q = 0 is the "right answer," and the correct herd size per animal harvested is HS = -20! Clearly nonsense!

We have ignored two biological restraints that will correct this nonsense. First, the herd size per animal harvested must be positive: HS>0. This imposes another condition on p,q:

$$HS = -20p + 48.63\,q > 0$$
$$\Leftrightarrow \quad p < \frac{48.63}{20}\,q = 2.4315\,q.$$

Since p + q = 1 we have $1 - q < 2.4315\,q \Leftrightarrow q > 1/3.4315 = .2914$ and $p < .7086$. Thus our nonsense value p = 1 is ruled out.

Secondly, all six components of the mini-herd that produces one animal for harvest [see (15)] must be positive. Once we plug in p and q, these components are given by the column vector shown in (17). (Trace the calculations until you see this.) All six will be positive if we insist that

$$-20\,p + 31.08q > 0$$
$$\Leftrightarrow \quad p < \frac{31.08}{20}\,q = 1.554\,q$$
$$\Leftrightarrow \quad 1 - q < 1.554\,q$$
$$\Leftrightarrow \quad q > 1/2.554 = .39154$$
$$\Leftrightarrow \quad p < .60846.$$

Conclusions: by taking p < .60846 but close to that value, and q = 1 - p, the herd size may be taken close to minimal.

In Table I, various values of p and q are used. The resulting values of HS and the resulting herds are shown. The *percentage* breakdown of the herd into its six components is given (or equivalently, an actual breakdown for a herd of 100 animals is given). Recall that HS is the size of the mini-herd that yields one animal for harvest; thus 1/HS is the fraction of the initial herd G_0 (investment) that is harvested after a year. Example: in the first column, each 6.90 animals breed to become 7.90 animals and yield a 1/6.90 or 14.5% "output." These figures are given as "% harvest."

Table 1. Structures of Seven Herds of Various Efficiencies

		①	②	③	④	⑤	⑥	⑦
	p	.608	.606	.605	.600	.580	.550	.500
	q	.392	.394	.395	.400	.420	.450	.500
	HS	6.90	7.04	7.11	7.45	8.82	10.88	14.3
	% harvest	14.5%	14.2%	14.1%	13.4%	11.3%	9.2%	7.0%
Initial herd % breakdown	AM	.34%	1.8%	2.5%	5.8%	16.5%	27.4%	38.7%
	AF	40.9	40.3	39.9	38.6	34.2	29.7	25.1
	YM	11.8	11.6	11.5	11.1	9.9	8.6	7.2
	YF	10.3	10.1	10.1	9.7	8.6	7.5	6.3
	CM	19.6	19.3	19.2	18.5	16.4	14.3	12.1
	CF	17.2	16.9	16.8	16.2	14.4	12.5	10.5

Here p = fraction of adults that are males; q = 1 - p = fraction of adults that are females. Herds of smaller size HS (animals per animal harvested) result as p is taken closer to .60846, which it cannot equal or exceed.

The structure and size of a herd that will yield a harvest of T animals annually varies considerably as we apportion the harvest differently among adult male and female animals. In a polygamous herd, there is no need to have anywhere near one bull per cow to achieve the birth rates for calves we have assumed. In this regard, it is common in cattle ranching to run 1 bull with 20-30 cows. The first three herds in the table above have cow-to-bull ratios of 120 (= 40.9/.34), 22, and 16; the other herds have much lower ratios. Thus herd (2) appears to be practical and is fairly close to minimal size.

4.7 Exercises

15. a. In Example 3, show that a 25:1 ratio of adult cows to bulls arises when p = .60624 is used.

    b. What value of p leads to a 30:1 ratio?

16. Check our work in Example 3 as follows: take a herd of one million animals structured like herd (2) in Table 1. (Thus there are 18,000 adult males, etc.) Use the after-harvest model as programmed in Exercise 3, and take a 14.2% harvest, using the values of p and q given in the table for herd 2 to calculate the constant annual harvest. On the computer, trace this initial herd for 20 years. It should remain roughly constant in size and structure.

5. APPLICATIONS: CALCULATING THE HARVEST

We will now ask what harvest should be taken from a herd already in our possession, if it is to be preserved in size for the future. We also will discuss harvests that provide for controlled growth of the herd. This is in contrast to Section 4, where we "designed" herds to provide specified harvests. Entirely different difficulties will appear.

5.1 Steady Annual Harvests and Herd

Example 4. Given "this year's" herd $\vec{H}_0$, what harvest Q_0 should be taken from it so that next year's herd H_1 will have the same size and structure as this year's herd, i.e. $\vec{H}_1 = \vec{H}_0$? (The process can then go on for many years, yielding *steady-state* harvests and herds.)

This question arises before we harvest, of course; *thus we use the count-before-harvest model*. Then we must solve

$$\vec{H}_1 = \vec{H}_0$$
$$\vec{H}_1 = M(\vec{H}_0 - \vec{Q}_0) \qquad \text{[compare (5)]}$$

for $\vec{Q}_0$, when $\vec{H}_0$ is known. Simplify the notation to $\vec{Q} = \vec{Q}_0$ and $\vec{H} = \vec{H}_0 = \vec{H}_1$ and use algebra to reach

$$MQ = (M - I)\,H. \tag{18}$$

(Here I is the 6 x 6 identity matrix.) The "obvious" next step is to multiply through by M^{-1} and get the "right answer" $\vec{Q} = M^{-1}(M - I)\vec{H}$. Unfortunately, M^{-1} does not exist!

So far we have worked at *matrix level*, i.e., we have used matrix algebra to calculate with the matrices as a whole, not their individual entries. To make more progress we must go down to *entry level* and look at the individual equations that make up the matrix level full system.

Let's examine (18) in detail. We appear to have six linear equations for the six unknowns in $\vec{Q}$. (The right side is known.) However, four of the six entries in $\vec{Q}$ were set as zero from the beginning. (We harvest only adult buffalo.) Thus, in (18) we have six equations in two unknowns, QM and QF. The equations are *overdetermined*. Usually, two conditions (equations) suffice to determine two unknowns. Only if we are lucky, by having the extra four conditions here add no contradictory requirements for QM and QF, will we have any solutions at all.

When are we lucky? The six equations say in detail:[8]

[8] Up to now we have used AM, AF, etc., as components of the herd after harvest, QM and QF as the number of buffalo just harvested. In Section 5 these variables are components of the herd *before* harvest and quotas of buffalo about to be harvested.

(19a) $\begin{Bmatrix} .95\ QM = -.05\ AM + .75\ YM \\ .95\ QF = -.05\ AF + .75\ YF \end{Bmatrix} \Longleftrightarrow \begin{Bmatrix} AM = .95\ (AM-QM) + .75\ YM \\ AF = .95\ (AF-QF) + .75\ YF \end{Bmatrix}$

(19b) $\begin{Bmatrix} 0 = -YM + .6\ CM \\ 0 = -YF + .6\ CF \end{Bmatrix} \Longleftrightarrow \begin{Bmatrix} YM = .6\ CM \\ YF = .6\ CF \end{Bmatrix}$

(19c) $\begin{Bmatrix} .48\ QF = .48\ AF - CM \\ .42\ QF = .42\ AF - CF \end{Bmatrix} \Longleftrightarrow \begin{Bmatrix} CM = .48\ (AF-QF) \\ CF = .42\ (AF-QF) \end{Bmatrix}$

Now, the values of AM, AF, YM, YF, CM, CF are assumed to be known, so we could solve for our unknowns, QM and QF, using equations (19a) alone. Then equations (19b,c) lead to a contradiction unless the values of AM, AF, YM, YF, CM, CF, QM and QF already known *happen* to satisfy (19b,c). Any herd for which these four equations (19b,c) are not satisfied *cannot* duplicate itself from this year to next *no matter what harvest is taken.* (Recall that we are requiring $\vec{H}_1 = \vec{H}_0$ with the strict mathematical meaning of equality for vectors.)

This makes sense if we read equations (19b,c) biologically. Consider (19b): to have $\vec{H}_1 = \vec{H}_0$, this year's yearlings (which, if they survive, are adults in H_1) must be exactly replaced in H_1 by the survivors of this year's calves. Equations (19b) say that YM and CM, YF and CF in our herd $\vec{H} = \vec{H}_0 = \vec{H}_1$ must be in the natural balance of six yearlings per ten calves for each sex *so that* the survival rate of .6 will cause this year's calves to exactly replace the yearling population as the year passes.

Now interpret (19c): This year's calves must also be precisely replaced by newborn calves if $\vec{H}_1 = \vec{H}_0$ is to be true. After the harvest, there will be AF - QF adult females and they will give birth to .48(AF - QF) new calf males and .42(AF - QF) new calf females by next year. Equations (19c) simply say that these births, forming the calf populations of $\vec{H}_1$, must exactly replace CM and CF in $\vec{H}_0$.

Thus, the four extra conditions in the overdetermined system (19) simply require that the herd have a natural age balance so that, considering the survival rates, it will replenish itself despite the harvest.

5.2 Constant Harvests From a Growing Herd

Example 5. We want to select a harvest $\vec{Q}$ so that, taking the same harvest every year, the herd will double in ten years while retaining the same proportional structure. That is, if $\vec{H}_0$ is our initial herd before harvest this year, then at the end of ten years we want to have $2\vec{H}_0$ as the herd structure.

We use the before-harvest-count because, again, that is when the question of selecting a quota arises. Let $\vec{H}_j$ be the herd before harvest in the j^{th} year, $j = 0, 1, 2, \ldots, 10$. Then

$$\begin{aligned} \vec{H}_1 &= M(\vec{H}_0 - \vec{Q}) \\ \vec{H}_2 &= M(\vec{H}_1 - \vec{Q}) = M\vec{H}_1 - M\vec{Q} \\ &= M^2(\vec{H}_0 - \vec{Q}) - M\vec{Q} \\ &= M^2\vec{H}_0 - M^2\vec{Q} - M\vec{Q} \\ \vec{H}_3 &= M(\vec{H}_2 - \vec{Q}) \\ &= M^3\vec{H}_0 - M^3\vec{Q} - M^2\vec{Q} - M\vec{Q}, \text{ etc.} \end{aligned}$$

Thus:

$$\begin{aligned} 2\vec{H}_0 = \vec{H}_{10} &= M^{10}\vec{H}_0 - M^{10}\vec{Q} - M^9\vec{Q} - \ldots - M\vec{Q} \\ &= M^{10}\vec{H}_0 - (I + M + M^2 + \ldots + M^9)M\vec{Q}. \end{aligned}$$

In this equation, we know $\vec{H}_0$ and want $\vec{Q}$. Therefore, write it as the set of linear equations

(20) $$\underbrace{(I + M + M^2 + \ldots + M^9)M}_{\text{known 6x6 matrix}}\ \underbrace{\vec{Q}}_{\text{unknown}} = \underbrace{(M^{10} - 2I)\vec{H}_0}_{\text{all known}}$$

All of this has been at matrix level. We push ahead in that spirit.

Have you noticed that $I + M + M^2 + \ldots + M^9$ looks like a geometric series? When numbers are involved, we know how to add up such expressions:

$$1 + a + a^2 + \ldots + a^{n-1} = \frac{1 - a^n}{1 - a} \text{ if } a \neq 1.$$

ERIC

Can we do something similar here, when M and I are square matrices?

Indeed we can. Put $S = I + M + M^2 + \dots + M^9$. Thus S is a 6 x 6 matrix, and MS makes sense: $MS = M + M^2 + \dots M^{10}$. Subtraction leads to the familiar massive cancellation:

$$(I - M)S = S - MS = I - M^{10}.$$

In fact, $(I - M)^{-1}$ does exist for our 6 x 6 matrix M. We calculated $(M - I)^{-1}$ in Section 4.2; of course $(I - M)^{-1} = -(M - I)^{-1}$. Thus

(21) $$S = I + M + M^2 + \dots + M^9 = (I - M)^{-1}(I - M^{10}).$$

The analogy to the numerical geometric series formula is striking. It might tempt us to believe the infinite geometric series formula:

$$I + M + M^2 + M^3 + \dots = (I - M)^{-1} \text{ analogous to } \text{"}\frac{I}{I - M}\text{"}.$$

Indeed, this formula is valid for certain families of matrices M and infinite series of matrices is a fascinating subject in its own right. We will not explore in that direction now, but one thing is clear: a sensible definition of "convergence" for such series would be our first task.

We were interested in solving the linear equations (20) for $\vec{Q}$. We have made progress: using (21) in (20) we obtain:

$$(I - M)^{-1}(I - M^{10})M\vec{Q} = (M^{10} - 2I)\vec{H}_0.$$

We can multiply through by I - M, and by $(I - M^{10})^{-1}$, which does exist (proof omitted):

(22) $$M\vec{Q} = (I - M^{10})^{-1}(I - M)(M^{10} - 2I)\vec{H}_0.$$

That is as far as we can go at matrix level in this example, because M^{-1} does not exist. The right side of (22) is known (although unpleasant to calculate). The system is overdetermined. Some herds can be doubled in ten years in the way we suggested, but most cannot.

Of course, we can *approximately* double the herd, and (22) will help us see how. We have examined whether we can precisely double it.

5.3 Exercises

17. a. Is the initial herd given in Exercise 1 a "natural" one which, if a proper harvest QM and QF were taken, could exactly reproduce itself next year? Explain your answer.

    b. Repeat a. for the initial herd of Exercise 4.

18. a. Show that M^{-1} does not exist. In how many ways can you do this?

    b. If we replace the ≠ 0 entries in M with arbitrary numbers a,b,c,d,e,f,g,h, we get

$$\hat{M} = \begin{bmatrix} a & 0 & b & 0 & 0 & 0 \\ 0 & c & 0 & d & 0 & 0 \\ 0 & 0 & 0 & 0 & e & 0 \\ 0 & 0 & 0 & 0 & 0 & f \\ 0 & g & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 & 0 \end{bmatrix}.$$

    Show that $(\hat{M})^{-1}$ does not exist, either. Thus the overdetermined nature of Examples 4 and 5 does not depend on specific birth and survival rates. (The reader who knows about determinants will have an advantage in this problem.)

19. a. Revise Example 5 so that the herd will grow by 50% in ten years. That is, set $H_{10} = 1.5\ H_0$ and carry through the algebra of Example 5 for this new case. Reach equations analogous to (22).

    b. Repeat a. with 50% growth over eight years.

20. Check our geometric series result in (21) by carefully multiplying out the left side of $(I-M)(I+M+M^2+\dots+M^9) = I-M^{10}$, to get the right side. (Why does this confirm equation (21)?) Identify all the algebraic properties of matrix multiplication and addition that you use, such as the associative law of multiplication, left distributive law, etc.

ERIC

6. REFERENCES

I first met this model when Karl Zinn of the Center for Research on Learning and Teaching at the University of Michigan introduced me to a computer program named BUFLO, written by L. Braun and R. L. Siegel of the Polytechnic Institute of Brooklyn and distributed nationally by the Program Library, Digital Equipment Corporation, Maynard, Massachusetts 01745. The program and its documentation are part of project EXTEND and the Huntington Two Computer Project. Program BUFLO interactively permits one to follow a buffalo herd through many years while applying a variety of management policies.

While equations (2) are taken directly from BUFLO, I am solely responsible for the mathematics that follows in this paper.

An alternative discussion of exactly the same model with different survival rates based on an actual modern buffalo herd may be found in:

Watt, Kenneth E. F., Ecology and Resource Management, McGraw Hill, 1968, p. 358 ff. This is an excellent book for all readers in applications of undergraduate-level math to biology.

The buffalo model discussed there is drawn from:

Fuller, W. A., "Biology and Management of the Bison of Wood Buffalo National Park," Canadian Department of Northern Affairs Natural Resources and Wildlife Management Bulletin, Series 1, No. 16, 1962.

As I read Watt, the survival coefficients matrix used by Fuller and Watt is:

$$M = \begin{pmatrix} .9 & 0 & .75 & 0 & 0 & 0 \\ 0 & .9 & 0 & .75 & 0 & 0 \\ 0 & 0 & 0 & 0 & .4 & 0 \\ 0 & 0 & 0 & 0 & 0 & .4 \\ 0 & .36 & 0 & 0 & 0 & 0 \\ 0 & .36 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and their "guesstimated" 1830 herd of 40 million buffalo is structured as:

$AM = 16.8$ million. $YM = 1.2$ $CM = 2.0$

$AF = 16.8$ $YF = 1.2$ $CF = 2.0$

Our model is a simplified variant of the more important Leslie models for populations with age structure. The original papers are:

Leslie, P.H., "The uses of matrices in certain population mathematics," Biometrika 33 (1945), pp. 183-212.

Leslie, P.H., "Some further notes on the use of matrices in population mathematics," Biometrika 35 (1948), pp. 213-245.

Much research by Leslie and others has followed, with the goal of overcoming the limitations of Leslie's original models. These limitations are much the same as the ones we have discussed for our simpler model: use of constant coefficients from year to year and linearity of the model. In addition, the Leslie approach has been applied to much more than buffalo herds. The interested reader might start with:

Pielou, E.C., An Introduction to Mathematical Ecology, Wiley-Interscience, New York, 1969. Chapter III covers the Leslie model. Pielou is a leading mathematical biologist; her books are among the basic advanced works in the field.

Usher, M.B., "A matrix approach to the management of renewable resources, with special reference to selection forests," Journal of Applied Ecology 3 (1966), pp. 355-367.

Usher, M.B., "A matrix approach to the management of renewable resources, with special reference to selection forests -- two extensions," Journal of Applied Ecology 6 (1969), pp. 347-8.

Usher, M.B., "A matrix model for forest management," Biometrics 25 (1969), pp. 309-315.

Fowler, Charles W. and Smith, Tim, "A matrix method for determining stable densities and age distributions and its application to African elephant populations." University of Washington Quantitative Science Paper No. 31, Seattle, January 1972. (Write Fowler or Smith at U. Washington, Seattle, 98195 for more information.)

A well-written discussion of the Leslie model with application to harvesting of herds (including data for sheep ranching) is

Anton, Howard, and Chris Rorres, Applications of Linear Algebra, John Wiley & Sons, 1977, Chapters 9 and 10.

7. ACKNOWLEDGEMENTS

I want to thank a number of people who have contributed as this paper has evolved into this third edition. Thanks to:

-- Karl Zinn of the Center for Research on Learning and Teaching, University of Michigan, for introducing me to the model and encouraging me to develop its mathematical content.

-- Tom Hern and Fred Rickey of Bowling Green State University, and David Staley of Ohio Wesleyan University for class-testing earlier editions in their linear algebra courses.

-- Bill Cannon of the U.S. Department of Agriculture Laboratory, Delaware, Ohio, for putting me in contact with much literature in this field and critically reading the first edition.

-- Edward Kelly, California State University at Hayward, for a careful review of the second edition.

-- Sol Garfunkel for administrative and editorial work at Project UMAP.

8. ANSWERS TO EXERCISES

1. I'll write vectors horizontally to save space. We are given $\vec{H}_0$ = (200, 1000, 300, 300, 520, 500) and $\vec{Q}$ = (100, 200, 0, 0, 0, 0).

   a. $\vec{G}_0 = \vec{H}_0 - \vec{Q}$
   = (100, 800, 300, 300, 520, 500)

   b. $\vec{H}_1 = M\vec{G}_0$ = (320, 985, 312, 300, 384, 336)

   c. $\vec{G}_1 = \vec{H}_1 - \vec{Q}$ = (220, 785, 312, 300, 384, 336)

   d. $\vec{H}_2 = M\vec{G}_1$ = (443, 971, 230, 202, 377, 330)
   (Decimal results have been rounded.)

   e. $\vec{G}_2 = \vec{H}_2 - \vec{Q}$ = (343, 771, 230, 202, 377, 330)

   f. The herd is shrinking slowly in the key category of adult females. This will continue for a while, causing the whole herd to shrink slowly.

2. Over two years $\vec{G}_0$, if left unharvested, would become M^2G_0. The harvest Q_2 is subtracted, of course. We also subtract, not Q_1, but the *descendants* of the harvested sub-herd Q_1 at the end of the two year period, $M\vec{Q}_1$. The linearity of the model assures that these sub-herds can all be superimposed.

3. A FORTRAN program is listed in Table 2, pages 34 and 35.

4. This may have been a frustrating problem -- it has no solution. The herd is inherently unstable because, in 1830, it was growing exponentially (or would have been, had not white man interfered). A harvest of 1.4 million males, 2.6 million females will convert the initial herd of 60 million into a herd of 59.984 million in ten years, but the herd structure is drastically changed. The new herd has many fewer calves than the original, and the herd is in fact headed for extinction. Other harvests of 4 million lead to herds that grow rapidly or decline rapidly, but this herd is inherently unstable. And that's the whole point.

5,6. Computer printouts are displayed in Tables 3 and 4 (pp.36-42). The point is that, by slaughtering females we also slaughter their potential progeny. The effect of harvesting a lot of females is to destroy the herd. Also, all of the herds that involve 20% harvest (Exercise 6) meet a fast extinction.

7. One example is shown in Table 5 (page 43), with commentary. You should try others.

9. a. $\vec{G}_1 = M\vec{G}_0 - \vec{Q}$
   $\vec{G}_1 = 1.12\ \vec{G}_0$

b. $\vec{G}_0 = (M - 1.12I)^{-1}\vec{Q}$

10. Equations $\vec{G}_1 = M\vec{G}_0 - \vec{Q}$

$\vec{G}_2 = M\vec{G}_1 - 1.1\vec{Q}$

$\vec{G}_2 = \vec{G}_0$

lead to solution

$\vec{G}_0 = (M^2 - I)^{-1}(M + 1.1I)\vec{Q}.$

11. Equations $\vec{G}_1 = M\vec{G}_0 - \vec{Q}$

$\vec{G}_2 = M\vec{G}_1 - 1.2\vec{Q}$

$\vec{G}_2 = 1.25\,\vec{G}_0$

lead to solution

$\vec{G}_0 = (M^2 - 1.25I)^{-1}(M + 1.2I)\vec{Q}.$

12. Equations $\vec{G}_1 = M\vec{G}_0$ (harvest is $\vec{Q} = \vec{0}$)

$\vec{G}_2 = M\vec{G}_1$

$\vec{G}_3 = M\vec{G}_2$

$\vec{G}_4 = M\vec{G}_3$

$\vec{G}_5 = M\vec{G}_4 - \vec{Q}$ (final harvest)

$\vec{G}_5 = 2\vec{G}_0$

have solution

$\vec{G}_0 = (M^5 - 2I)^{-1}\vec{Q}.$

13. Equations $\vec{G}_1 = M\vec{G}_0 - \vec{Q}$

$\vec{G}_2 = M\vec{G}_1 - \vec{Q}$

$\vec{G}_3 = M\vec{G}_2 - \vec{Q}$

$\vec{G}_4 = M\vec{G}_3 - \vec{Q}$

$\vec{G}_5 = M\vec{G}_4 - \vec{Q}$

$\vec{G}_5 = 2\vec{G}_0$

condense to

$2\vec{G}_0 = \vec{G}_5 = M^5\vec{G}_0 - (M^4 + M^3 + M^2 + M + I)\vec{Q}.$

Thus,

$\vec{G}_0 = (M^5 - 2I)^{-1}(M^4 + M^3 + M^2 + M + I)\vec{Q}.$

14. Equations $\vec{G}_1 = M\vec{G}_0$

$\vec{G}_2 = M\vec{G}_1 - \vec{Q}$

$\vec{G}_3 = M\vec{G}_2$

$\vec{G}_4 = M\vec{G}_3 - \vec{Q}$

$\vec{G}_5 = M\vec{G}_4$

$\vec{G}_6 = M\vec{G}_5 - \vec{Q}$

$\vec{G}_6 = 2\vec{G}_0.$

condense to

$2\vec{G}_0 = \vec{G}_6 = M^6\vec{G}_0 - (M^4 + M^2 + I)\vec{Q}.$

The solution is

$\vec{G}_0 = (M^6 - 2I)^{-1}(M^4 + M^2 + I)\vec{Q}.$

15. b. $r = .60661$

16. A computer printout appears as Table 6 (page 44). Results are right on target.

17. a. No; to give just one reason among many, equations (19b) are not satisfied by CM = 520, YM = 300.

b. No; again, equations (19b) are not satisfied by a herd with 12% female calves and 8% female yearlings.

19. a. Change the equations to

$\vec{H}_{10} = M^{10}\vec{H}_0 - M(I + M + M^2 + \dots + M^9)\vec{Q}$

$\vec{H}_{10} = 1.5\,\vec{H}_0.$

Then (22) is replaced by the overdetermined system

$M\vec{Q} = (I - M^{10})^{-1}(I - M)(M^{10} - 1.5I)\vec{H}_0.$

b. Now

$\vec{H}_8 = M^8\vec{H}_0 - M(I + M + M^2 + \dots + M^7)\vec{Q}$

$\vec{H}_8 = 1.5\vec{H}_0$

lead to this replacement for (22):

$M\vec{Q} = (I - M^8)^{-1}(I - M)(M^8 - 1.5I)\vec{H}_0.$

TABLE 2

A listing of my FORTRAN program, used to create all the printouts that follow, is given below. It does more than Problem 6 asks, because it gives the results in percentages and in actual millions of buffalo. The program was run on an IBM 1130 computer but should easily adapt to any standard FORTRAN.

```
C THIS PROGRAM ACCOMPANIES THE APPLICATION PAPER
C
C  'MANAGEMENT OF A BUFFALO HERD'
C
C  AND DOES THE CALCULATIONS REQUESTED IN PROBLEM 3 OF THAT PAPER.
C  IT RECEIVES PAIRS OF DATA CARDS AS INPUT. THE FIRST CARD SHOULD
C  LIST THE SIZE OF THE INITIAL HERD IN THE USUAL  SIX CATEGORIES, IN
C  6 F 10.4 FORMAT. IN MILLIONS THIS TELLS THE PROGRAM THE INITIAL
C  NUMBER OF AM,AF,YM,YF,CM,CF.  THE SECOND DATA CARD LISTS THE CONSTANT
C  HARVEST OF MALES, THEN FEMALES, IN 2 F 10.4 FORMAT (GIVE THESE IN
C  MILLIONS, TOO) AND THE NUMBER OF YEARS THAT THE HERD IS TO BE
C  TRACED, IN I2 FORMAT IN COLUMNS 21,22.
C
C  PLACE PAIRS OF DATA CARDS BEHIND ONE ANOTHER. PROGRAM TERMINATES
C  WHEN A FAKE DATA-CARD-PAIR IS FOUND WITH A NEGATIVE ENTRY IN THE AM SPOT.
C  THUS MANY HERDS MAY BE STUDIED WITH ONE COMPUTER RUN.
C
C  OUTPUT IS GIVEN IN MILLIONS OF ANIMALS AND ALSO IN A PERCENTAGE
C  BREAKDOWN OF THE HERD. YEAR BY YEAR.
C
      DIMENSION H(6),Q(6),SAVE(50,7)
    1 READ(2,100)  H,Q(1),Q(2),LONG
  100 FORMAT(     6F10.4/2F10.4,I2)
      IF(H(1).LT.0)      CALL EXIT
      WRITE(5,101)
  101 FORMAT(    1H1,24X,'MILLIONS OF BUFFALO')
      WRITE (5,102)
  102 FORMAT (  1X,'YEAR      TOTAL          AM          AF          YM          Y
     1F          CM          CF')
      NYEAR=0
      TOTAL=0
      DO 5 L=1,6
    5 TOTAL=TOTAL+H(L)
      WRITE(5,103)    NYEAR,TOTAL,H
  103 FORMAT(    2X,I2,4X,F7.3,6(3X,F7.3))
C CONVERT TO PERCENTS AND SAVE FOR LATER PRINTING.
      SAVE(1,1)=0.
      DO 10 L=2,7
      LL=L-1
      SAVE(1,L)=H(LL)/TOTAL*100
   10 SAVE(1,1)=SAVE(1,1)+SAVE(1,L)
```

Note All of this code simply sets up the initial herd properly

```
C MAIN LOOP
      DO 25 K=1,LONG
      NYEAR=K
      TEMPO=H(2)
      H(1) = .95*H(1) + .75*H(3) - Q(1)
      H(2) = .95*H(2) + .75*H(4) - Q(2)
      H(3) = .6*H(5)
      H(4) = .6*H(6)
      H(5) = .48*TEMPO
      H(6) = .42*TEMPO
      TOTAL =0
      DO 15 L=1,6
   15 TOTAL=TOTAL+H(L)
      WRITE(5,103)  NYEAR,TOTAL,H
      KK=K+1
      SAVE(KK,1)=0
      DO 20 L=2,7
      LL=L-1
      SAVE(KK,L) =H(LL)/TOTAL*100
   20 SAVE(KK,1) = SAVE(KK,1)+SAVE(KK,L)
   25 CONTINUE
```

Note TEMPO is used to avoid a key trap in the program if I fail to save the old value of H(2), I will not have it to use in the correct calculation of H(5) and H(6)

...continued next page......


```
C PRINT TABLE OF PERCENTS
      WRITE (5,104)
  104 FORMAT(  1H0,18X,'PERCENTAGE DISTRIBUTION OF HERD')
      WRITE(5,102)
      LL=LONG+1
      DO 30 K=1,LL
      NYEAR= K-1
      WRITE(5,105)  NYEAR,(SAVE(K,J),J=1,7)
  105 FORMAT(    2X,I2,5X,F5.1,6X,6(F4.1,6X))
   30 CONTINUE
      WRITE(5,106)    Q(1),Q(2)
  106 FORMAT(1H0,'CONSTANT ANNUAL HARVEST IS 'F6.2' MALES, ' F6.2' FEMAL
     1ES (MILLIONS)')
      GO TO 1
      END
```

The data cards that produce the printout of Table 3, page 36, are these, given as samples. Many pairs of data cards can precede the fake pair.

18.	16.2	5.4	4.8	8.4	7.2	(initial herd)	COMMENTS
4.	0.	20	(harvest of males, females; years traced)				
-100.			(fake data-card-pair to terminate program)				

blank card

TABLE 3

Twenty-year printouts for the five cases called for in Exercise 5 follow.

Case a)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	50.999	18.000	16.200	5.400	4.800	8.400	7.200
1	60.079	17.150	18.990	5.040	4.320	7.775	6.803
2	63.191	16.072	21.280	4.665	4.082	9.115	7.975
3	67.453	14.768	23.278	5.469	4.785	10.214	8.937
4	72.276	14.131	25.703	6.128	5.362	11.173	9.776
5	78.165	14.021	28.440	6.704	5.866	12.337	10.795
6	85.242	14.348	31.417	7.402	6.477	13.651	11.944
7	93.521	15.183	34.704	8.190	7.166	15.080	13.195
8	103.112	16.567	38.344	9.048	7.917	16.658	14.576
9	114.141	18.524	42.365	9.995	8.745	18.405	16.104
10	126.736	21.094	46.806	11.043	9.662	20.335	17.793
11	141.039	24.322	51.713	12.201	10.676	22.467	19.658
12	157.210	28.257	57.134	13.480	11.795	24.822	21.719
13	175.426	32.954	63.124	14.893	13.031	27.424	23.996
14	195.884	38.477	69.742	16.454	14.397	30.299	26.512
15	218.803	44.894	77.053	18.179	15.907	33.476	29.291
16	244.425	52.284	85.131	20.085	17.575	36.985	32.362
17	273.018	60.734	94.056	22.191	19.417	40.863	35.755
18	304.879	70.341	103.916	24.517	21.453	45.146	39.503
19	340.338	81.212	114.810	27.088	23.702	49.879	43.644
20	379.759	93.468	126.846	29.927	26.186	55.109	48.220

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	28.5	31.6	8.3	7.1	12.9	11.3
2	100.0	25.4	33.6	7.3	6.4	14.4	12.6
3	100.0	21.8	34.5	8.1	7.0	15.1	13.2
4	100.0	19.5	35.5	8.4	7.4	15.4	13.5
5	100.0	17.9	36.3	8.5	7.5	15.7	13.8
6	100.0	16.8	36.8	8.6	7.5	16.0	14.0
7	100.0	16.2	37.1	8.7	7.6	16.1	14.1
8	100.0	16.0	37.1	8.7	7.6	16.1	14.1
9	100.0	16.2	37.1	8.7	7.6	16.1	14.1
10	100.0	16.6	36.9	8.7	7.6	16.0	14.0
11	99.9	17.2	36.6	8.6	7.5	15.9	13.9
12	100.0	17.9	36.3	8.5	7.5	15.7	13.8
13	100.0	18.7	35.9	8.4	7.4	15.6	13.6
14	100.0	19.6	35.6	8.4	7.3	15.4	13.5
15	100.0	20.5	35.2	8.3	7.2	15.2	13.3
16	100.0	21.3	34.8	8.2	7.1	15.1	13.2
17	99.9	22.2	34.4	8.1	7.1	14.9	13.0
18	100.0	23.0	34.0	8.0	7.0	14.8	12.9
19	100.0	23.8	33.7	7.9	6.9	14.6	12.8
20	100.0	24.6	33.4	7.8	6.8	14.5	12.6

CONSTANT ANNUAL HARVEST IS 4.00 MALES, 0.00 FEMALES (MILLIONS)

TABLE 3 (Continued)

Case b)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	60.079	18.150	17.990	5.040	4.320	7.775	6.803
2	62.291	18.022	19.330	4.665	4.082	8.635	7.555
3	65.158	17.620	20.425	5.181	4.533	9.278	8.118
4	68.251	17.625	21.804	5.567	4.871	9.804	8.578
5	71.941	17.919	23.367	5.882	5.147	10.466	9.157
6	76.300	18.435	25.059	6.279	5.494	11.216	9.814
7	81.324	19.223	26.927	6.729	5.888	12.028	10.525
8	87.075	20.309	28.998	7.217	6.315	12.925	11.309
9	93.631	21.707	31.284	7.755	6.785	13.919	12.179
10	101.063	23.438	33.809	8.351	7.307	15.016	13.139
11	109.452	25.529	36.599	9.009	7.883	16.228	14.200
12	118.890	28.010	39.682	9.737	8.520	17.567	15.371
13	129.480	30.913	43.088	10.540	9.223	19.047	16.666
14	141.332	34.273	46.851	11.428	10.000	20.682	18.097
15	154.574	38.130	51.008	12.409	10.858	22.488	19.677
16	169.341	42.531	55.602	13.493	11.806	24.484	21.423
17	185.788	47.524	60.677	14.690	12.854	26.689	23.352
18	204.084	53.166	66.283	16.013	14.011	29.124	25.484
19	224.417	59.518	72.478	17.474	15.290	31.816	27.839
20	246.995	66.648	79.322	19.089	16.703	34.789	30.440

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	30.2	29.9	8.3	7.1	12.9	11.3
2	100.0	28.9	31.0	7.4	6.5	13.8	12.1
3	100.0	27.0	31.3	7.9	6.9	14.2	12.4
4	99.9	25.8	31.9	8.1	7.1	14.3	12.5
5	100.0	24.9	32.4	8.1	7.1	14.5	12.7
6	100.0	24.1	32.8	8.2	7.2	14.7	12.8
7	99.9	23.6	33.1	8.2	7.2	14.7	12.9
8	100.0	23.3	33.3	8.2	7.2	14.8	12.9
9	100.0	23.1	33.4	8.2	7.2	14.8	13.0
10	100.0	23.1	33.4	8.2	7.2	14.8	13.0
11	100.0	23.3	33.4	8.2	7.2	14.8	12.9
12	100.0	23.5	33.3	8.1	7.1	14.7	12.9
13	100.0	23.8	33.2	8.1	7.1	14.7	12.8
14	100.0	24.2	33.1	8.0	7.0	14.6	12.8
15	100.0	24.6	32.9	8.0	7.0	14.5	12.7
16	100.0	25.1	32.8	7.9	6.9	14.4	12.6
17	100.0	25.5	32.6	7.9	6.9	14.3	12.5
18	100.0	26.0	32.4	7.8	6.8	14.2	12.4
19	100.0	26.5	32.2	7.7	6.8	14.1	12.4
20	100.0	26.9	32.1	7.[illegible]	6.7	14.0	12.3

CONSTANT ANNUAL HARVEST IS 3.00 MALES, 1.00 FEMALES (MILLIONS)

ERIC

TABLE 3 (Continued)

Case c)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	60.079	19.150	16.990	5.040	4.320	7.775	6.803
2	61.391	19.972	17.380	4.665	4.082	8.155	7.135
3	62.863	20.473	17.573	4.893	4.281	8.342	7.299
4	64.226	21.119	17.905	5.005	4.379	8.435	7.380
5	65.717	21.817	18.295	5.061	4.428	8.594	7.520
6	67.359	22.522	18.701	5.156	4.512	8.781	7.684
7	69.126	23.263	19.151	5.269	4.610	8.976	7.854
8	71.038	24.052	19.651	5.386	4.712	9.192	8.043
9	73.120	24.889	20.203	5.515	4.826	9.432	8.253
10	75.389	25.781	20.812	5.659	4.952	9.697	8.485
11	77.864	26.737	21.486	5.818	5.091	9.990	8.741
12	80.571	27.764	22.230	5.994	5.244	10.313	9.024
13	83.533	28.871	23.052	6.188	5.414	10.670	9.336
14	86.781	30.069	23.960	6.402	5.602	11.065	9.682
15	90.344	31.367	24.964	6.639	5.809	11.501	10.063
16	94.257	32.778	26.072	6.900	6.038	11.982	10.484
17	98.559	34.314	27.297	7.189	6.290	12.514	10.950
18	103.290	35.991	28.651	7.508	6.570	13.102	11.465
19	180.496	37.823	30.146	7.861	6.879	13.752	12.033
20	114.230	39.828	31.798	8.251	7.220	14.470	12.661

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	31.8	28.2	8.3	7.1	12.9	11.3
2	100.0	32.5	28.3	7.5	6.6	13.2	11.6
3	100.0	32.5	27.9	7.7	6.8	13.2	11.6
4	99.9	32.8	27.8	7.7	6.8	13.1	11.4
5	99.9	33.1	27.8	7.7	6.7	13.0	11.4
6	100.0	33.4	27.7	7.6	6.6	13.0	11.4
7	100.0	33.6	27.7	7.6	6.6	12.9	11.3
8	99.9	33.8	27.6	7.5	6.6	12.9	11.3
9	100.0	34.0	27.6	7.5	6.6	12.9	11.2
10	100.0	34.1	27.6	7.5	6.5	12.8	11.2
11	99.9	34.3	27.5	7.4	6.5	12.8	11.2
12	100.0	34.4	27.5	7.4	6.5	12.8	11.2
13	100.0	34.5	27.5	7.4	6.4	12.7	11.1
14	100.0	34.6	27.6	7.3	6.4	12.7	11.1
15	100.0	34.7	27.6	7.3	6.4	12.7	11.1
16	100.0	34.7	27.6	7.3	6.4	12.7	11.1
17	100.0	34.8	27.6	7.2	6.3	12.6	11.1
18	100.0	34.8	27.7	7.2	6.3	12.6	11.0
19	100.0	34.8	27.7	7.2	6.3	12.6	11.0
20	100.0	34.8	27.8	7.2	6.3	12.6	11.0

CONSTANT ANNUAL HARVEST IS 2.00 MALES, 2.00 FEMALES (MILLIONS)

TABLE 3 (Continued)

Case d)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	60.079	20.150	15.989	5.040	4.320	7.775	6.803
2	60.491	21.922	15.430	4.665	4.082	7.675	6.715
3	60.568	23.325	14.720	4.605	4.029	7.406	6.480
4	60.201	24.613	14.006	4.443	3.888	7.065	6.182
5	59.493	25.715	13.222	4.239	3.709	6.723	5.882
6	58.417	26.609	12.343	4.033	3.529	6.346	5.553
7	56.928	27.304	11.374	3.808	3.332	5.925	5.184
8	55.002	27.795	10.304	3.555	3.110	5.459	4.777
9	52.610	28.071	9.122	3.275	2.866	4.946	4.327
10	49.715	28.125	7.815	2.967	2.596	4.378	3.831
11	46.277	27.944	6.372	2.627	2.298	3.751	3.282
12	42.251	27.517	4.778	2.250	1.969	3.058	2.676
13	37.587	26.830	3.016	1.835	1.605	2.293	2.006
14	32.229	25.865	1.069	1.376	1.204	1.447	1.266
15	26.115	24.603	-1.080	0.868	0.760	0.513	0.449
16	19.174	23.025	-3.456	0.308	0.269	-0.518	-0.453
17	11.329	21.105	-6.081	-0.311	-0.272	-1.659	-1.451
18	2.495	18.816	-8.981	-0.995	-0.871	-2.919	-2.554
19	-7.423	16.129	-12.185	-1.751	-1.532	-4.311	-3.772
20	-18.533	13.008	-15.725	-2.586	-2.263	-5.849	-5.117

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	33.5	26.6	8.3	7.1	12.9	11.3
2	100.0	36.2	25.5	7.7	6.7	12.6	11.1
3	100.0	38.5	24.3	7.6	6.6	12.2	10.6
4	100.0	40.8	23.2	7.3	6.4	11.7	10.2
5	100.0	43.2	22.2	7.1	6.2	11.3	9.8
6	100.0	45.5	21.1	6.9	6.0	10.8	9.5
7	100.0	47.9	19.9	6.6	5.8	10.4	9.1
8	100.0	50.5	18.7	6.4	5.6	9.9	8.6
9	100.0	53.3	17.3	6.2	5.4	9.4	8.2
10	100.0	56.5	15.7	5.9	5.2	8.8	7.7
11	100.0	60.3	13.7	5.6	4.9	8.1	7.0
12	100.0	65.1	11.3	5.3	4.6	7.2	6.3
13	100.0	71.3	8.0	4.8	4.2	6.1	5.3
14	100.0	80.2	3.3	4.2	3.7	4.4	3.9

CONSTANT ANNUAL HARVEST IS 1.00 MALES, 3.00 FEMALES (MILLIONS)
The date for years 15-20 is nonsensical, and means that the herd is extinct: after the 14th year, the required harvest of adult females is not available.

TABLE 3 (Continued)

Case e)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	60.079	21.150	14.989	5.040	4.320	7.775	6.803
2	59.591	23.872	13.480	4.665	4.082	7.195	6.295
3	58.273	26.178	11.868	4.317	3.777	6.470	5.661
4	56.175	28.107	10.107	3.882	3.397	5.696	4.984
5	53.269	29.613	8.150	3.418	2.990	4.851	4.245
6	49.475	30.696	5.985	2.911	2.547	3.912	3.423
7	44.730	31.344	3.597	2.347	2.053	2.873	2.514
8	38.965	31.538	0.957	1.723	1.508	1.726	1.510
9	32.099	31.254	-1.958	1.035	0.906	0.459	0.402

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	35.2	24.9	8.3	7.1	12.9	11.3
2	100.0	40.0	22.6	7.8	6.8	12.0	10.5
3	100.0	44.9	20.3	7.4	6.4	11.1	9.7
4	100.0	50.0	17.9	6.9	6.0	10.1	8.8
5	100.0	55.5	15.3	6.4	5.6	9.1	7.9
6	100.0	62.0	12.0	5.8	5.1	7.9	6.9
7	100.0	70.0	8.0	5.2	4.5	6.4	5.6
8	100.0	80.9	2.4	4.4	3.8	4.4	3.8

CONSTANT ANNUAL HARVEST IS 0.00 MALES, 4.00 FEMALES (MILLIONS)
Extinction occurs as a result of the harvest following the eighth year.

TABLE 4

Twenty percent harvests lead to early extinction in all five cases requested in Exercise 6:

Case a)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	52.079	9.149	18.990	5.040	4.320	7.775	6.803
2	47.591	0.472	21.280	4.665	4.082	9.115	7.975
3	44.633	-8.051	23.278	5.469	4.785	10.214	8.937

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	17.5	36.4	9.6	8.2	14.9	13.0
2	100.0	0.9	44.7	9.8	8.5	19.1	16.7

CONSTANT ANNUAL HARVEST IS 12.00 MALES, 0.00 FEMALES (MILLIONS)

Case b)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	52.079	12.149	15.989	5.040	4.320	7.775	6.803
2	44.891	6.322	15.430	4.665	4.082	7.675	6.715
3	37.748	0.505	14.720	4.605	4.029	7.406	6.480
4	30.522	-5.065	14.006	4.443	3.888	7.065	6.182

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	100.0	23.3	30.7	9.6	8.2	14.9	13.0
2	100.0	14.0	34.3	10.3	9.0	17.0	14.9
3	100.0	1.3	38.9	12.1	10.6	19.6	17.1

CONSTANT ANNUAL HARVEST IS 9.00 MALES, 3.00 FEMALES (MILLIONS)

Case c)

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	52.079	15.149	12.989	5.040	4.320	7.775	6.803
2	42.191	12.172	9.580	4.665	4.082	6.235	5.455
3	30.863	9.063	6.163	3.741	3.273	4.598	4.023
4	18.446	5.415	2.310	2.759	2.414	2.958	2.588
5	4.627	1.214	-1.994	1.775	1.553	1.108	0.970

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.00	30.0	27.0	9.0	8.0	14.0	12.0
1	100.00	29.0	24.9	9.6	8.2	14.9	13.0
2	100.00	28.8	22.7	11.0	9.6	14.7	12.9
3	100.00	29.3	19.9	12.1	10.6	14.8	13.0
4	100.00	29.3	12.5	14.9	13.0	16.0	14.0

CONSTANT ANNUAL HARVEST IS 6.00 MALES, 6.00 FEMALES (MILLIONS)

Table 4 (continued)

Case d)

			MILLIONS OF BUFFALO				
YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	52.079	18.150	9.989	5.040	4.320	7.775	6.803
2	39.491	18.022	3.730	4.665	4.082	4.795	4.195
3	23.978	17.620	-2.394	2.877	2.517	1.790	1.566

			PERCENTAGE DISTRIBUTION OF HERD				
YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.00	30.0	27.0	9.0	8.0	14.0	12.0
1	100.00	34.8	19.1	9.6	8.2	14.9	13.0
2	100.00	45.6	9.4	11.8	10.3	12.1	10.6

CONSTANT ANNUAL HARVEST IS 3.00 MALES, 9.00 FEMALES (MILLIONS)

Case e)

			MILLIONS OF BUFFALO				
YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	52.079	21.150	6.989	5.040	4.320	7.775	6.803
2	36.791	23.872	-2.119	4.665	4.082	3.355	2.935

			PERCENTAGE DISTRIBUTION OF HERD				
YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.00	30.0	27.0	9.0	8.0	14.0	12.0
1	100.00	40.6	13.4	9.6	8.2	14.9	13.0

CONSTANT ANNUAL HARVEST IS 0.00 MALES, 12.00 FEMALES (MILLIONS)

TABLE 5

Data for Exercise 7. As one example, the initial herd was transformed for one year in this catastrophic way:

$$M = \begin{pmatrix} .60 & 0 & .40 & 0 & 0 & 0 \\ 0 & .60 & 0 & .40 & 0 & 0 \\ 0 & 0 & 0 & 0 & .15 & 0 \\ 0 & 0 & 0 & 0 & 0 & .15 \\ 0 & .25 & 0 & 0 & 0 & 0 \\ 0 & .20 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and then transformed further for 19 more years using the usual matrix M. The results:

			MILLIONS OF BUFFALO				
YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	59.999	18.000	16.200	5.400	4.800	8.400	7.200
1	34.229	12.960	11.639	1.260	1.080	4.050	3.240
2	39.973	13.257	11.867	2.429	1.943	5.586	4.888
3	44.113	14.416	12.731	3.352	2.933	5.696	4.984
4	48.371	16.209	14.294	3.417	2.990	6.111	5.347
5	53.526	17.962	15.823	3.666	3.208	6.861	6.003
6	59.212	19.814	17.438	4.116	3.602	7.595	6.645
7	65.418	21.911	19.267	4.557	3.987	8.370	7.324
8	72.286	24.233	21.295	5.022	4.394	9.248	8.092
9	79.885	26.788	23.526	5.549	4.855	10.221	8.943
10	88.275	29.611	25.991	6.132	5.366	11.292	9.880
11	97.543	32.730	28.716	6.775	5.928	12.475	10.916
12	107.783	36.175	31.727	7.485	6.549	13.784	12.061
13	119.095	39.980	35.053	8.270	7.236	15.229	13.325
14	131.593	44.184	38.728	9.137	7.995	16.825	14.722
15	145.400	48.828	42.788	10.095	8.833	18.589	16.265
16	160.655	53.958	47.273	11.153	9.759	20.538	17.971
17	177.507	59.625	52.229	12.323	10.782	22.691	19.855
18	196.126	65.886	57.705	13.614	11.913	25.070	21.936
19	216.697	72.803	63.754	15.042	13.161	27.698	24.236
20	239.423	80.445	70.438	16.619	14.541	30.602	26.777

			PERCENTAGE DISTRIBUTION OF HERD				
YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	30.0	27.0	9.0	8.0	14.0	12.0
1	99.9	37.8	34.0	3.6	3.1	11.8	9.4
2	100.0	33.1	29.6	6.0	4.8	13.9	12.2
3	99.9	32.6	28.8	7.5	6.6	12.9	11.[illegible]
4	100.0	33.5	29.5	7.0	6.1	12.6	11.0
5	100.0	33.5	29.5	6.8	5.9	12.8	11.2
6	100.0	33.4	29.4	6.9	6.0	12.8	11.2
7	100.0	33.4	29.4	6.9	6.0	12.7	11.1
8	100.0	33.5	29.4	6.9	6.0	12.7	11.1
9	99.9	33.5	29.4	6.9	6.0	12.7	11.1
10	100.0	33.5	29.4	6.9	6.0	12.7	11.1
11	100.0	33.5	29.4	6.9	6.0	12.7	11.1
12	100.0	33.5	29.4	6.9	6.0	12.7	11.1
13	100.0	33.5	29.4	6.9	6.0	12.7	11.1
14	100.0	33.5	29.4	6.9	6.0	12.7	11.1
15	100.0	33.5	29.4	6.9	6.0	12.7	11.1
16	100.0	33.5	29.4	6.9	6.0	12.7	11.1
17	100.0	33.5	29.4	6.9	6.0	12.7	11.1
18	100.0	33.5	29.4	6.9	6.0	12.7	11.1
19	100.0	33.5	29.4	6.9	6.0	12.7	11.1
20	100.0	33.5	29.4	6.9	6.0	12.7	11.1

ANNUAL HARVEST IS 0.00 MALES, 0.00 FEMALES (MILLIONS)

ERIC

TABLE 6

Data for Exercise 16. The herd does indeed remain very stable. There is some roundoff error: the harvests taken were .086 and .056 annually (males, females, in millions), rather than the .086052 and .055948 that the table's data for herd 2 indicates.

MILLIONS OF BUFFALO

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	0.999	0.018	0.403	0.116	0.101	0.193	0.169
1	1.000	0.018	0.402	0.115	0.101	0.193	0.169
2	1.000	0.018	0.402	0.116	0.101	0.193	0.169
3	1.000	0.018	0.402	0.115	0.101	0.193	0.169
4	1.000	0.018	0.402	0.115	0.101	0.193	0.169
5	1.000	0.018	0.402	0.115	0.101	0.193	0.169
6	1.000	0.018	0.402	0.115	0.101	0.193	0.169
7	1.000	0.018	0.402	0.115	0.101	0.193	0.169
8	1.000	0.018	0.402	0.115	0.101	0.193	0.169
9	1.000	0.018	0.402	0.115	0.101	0.193	0.168
10	0.999	0.018	0.402	0.115	0.101	0.193	0.168
11	0.999	0.018	0.402	0.115	0.101	0.193	0.168
12	0.999	0.018	0.402	0.115	0.101	0.193	0.168
13	0.999	0.018	0.401	0.115	0.101	0.192	0.168
14	0.999	0.018	0.401	0.115	0.101	0.192	0.168
15	0.998	0.018	0.401	0.115	0.101	0.192	0.168
16	0.998	0.018	0.401	0.115	0.101	0.192	0.168
17	0.998	0.018	0.401	0.115	0.101	0.192	0.168
18	0.997	0.017	0.401	0.115	0.101	0.192	0.168
19	0.997	0.017	0.401	0.115	0.101	0.192	0.168
20	0.996	0.017	0.401	0.115	0.101	0.192	0.168

PERCENTAGE DISTRIBUTION OF HERD

YEAR	TOTAL	AM	AF	YM	YF	CM	CF
0	100.0	1.8	40.3	11.6	10.1	19.3	16.9
1	100.0	1.8	40.2	11.5	10.1	19.3	16.9
2	100.0	1.8	40.2	11.6	10.1	19.3	16.9
3	100.0	1.8	40.2	11.5	10.1	19.3	16.8
4	100.0	1.8	40.2	11.5	10.1	19.3	16.9
5	100.0	1.8	40.2	11.5	10.1	19.3	16.8
6	100.0	1.8	40.2	11.5	10.1	19.3	16.8
7	100.0	1.8	40.2	11.5	10.1	19.3	16.8
8	100.0	1.8	40.2	11.5	10.1	19.3	16.8
9	100.0	1.8	40.2	11.5	10.1	19.3	16.8
10	100.0	1.8	40.2	11.5	10.1	19.3	16.8
11	100.0	1.8	40.2	11.5	10.1	19.3	16.8
12	100.0	1.8	40.2	11.5	10.1	19.3	16.8
13	100.0	1.8	40.2	11.5	10.1	19.3	16.8
14	100.0	1.8	40.2	11.5	10.1	19.3	16.8
15	100.0	1.8	40.2	11.5	10.1	19.3	16.9
16	100.0	1.8	40.2	11.5	10.1	19.3	16.9
17	100.0	1.8	40.2	11.5	10.1	19.3	16.9
18	100.0	1.7	40.2	11.5	10.1	19.3	16.9
19	100.0	1.7	40.2	11.5	10.1	19.3	16.9
20	100.0	1.7	40.2	11.6	10.1	19.3	16.9

ERIC Full Text Provided by ERIC

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.

STUDENT FORM 2

Unit Questionnaire

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

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.)

ERIC

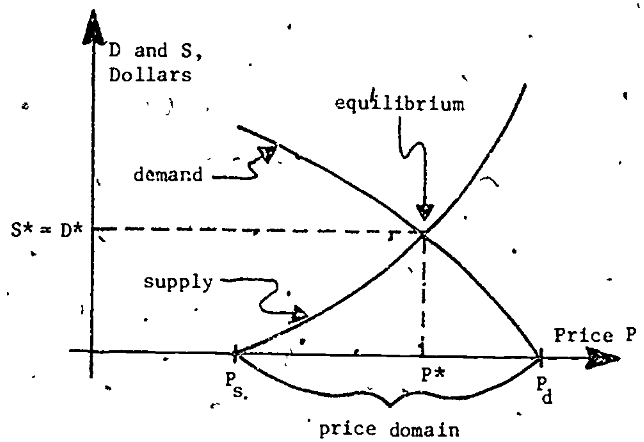
umap

UNIT 208

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

ECONOMIC EQUILIBRIUM: SIMPLE LINEAR MODELS

by Philip M. Tuchinsky



APPLICATIONS OF LINEAR ALGEBRA TO ECONOMICS

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

SE 036 478

ERIC

ECONOMIC EQUILIBRIUM: SIMPLE LINEAR MODELS

by

Philip M. Tuchinsky
7623 Charlesworth
Dearborn Hts., Michigan 4812-

TABLE OF CONTENTS

Intermodular Description Sheet: *UMAP Unit 208*

Title: ECONOMIC EQUILIBRIUM: SIMPLE LINEAR MODELS

Author: Philip M. Tuchinsky
7623 Charlesworth
Dearborn Hts., Michigan 48127

Dr. Tuchinsky is a computer scientist and applied mathematician at Ford Motor Company's Research and Engineering Center. He formerly taught in the Mathematical Sciences Department of Ohio Wesleyan University (where earlier editions of this paper were written).

Review Stage/Date: III 9/20/79

Classification: APPL LIN ALG/ECON

Approximate Class Time: Two 50-minute classes.

Intended Audience: Linear algebra students who have just learned about calculation of matrix inverses. To read Part VI, a student should have had some contact with differential calculus. The paper is also suitable for independent reading or seminar presentation by more advanced students.

Prerequisite Skills:

1. Elementary high-school algebra.
2. Graphing of straight lines.
3. Familiarity with functions and function notation.
4. Knowledge of the domain of a function.
5. Interval notation [a,b].
6. Matrix and vector notation.
7. Elementary matrix algebra including multiplication of matrices.
8. Matrix inverses as a concept, with algebraic laws and notation.

(For Part VI only:)

9. The derivative and its notation.
10. Continuity and general smoothness concepts.
11. Taylor's Theorem (the equation of the tangent line).
12. Newton's Method (one variable) is mentioned in Exercise 17. (Part VI is a nice vehicle for motivating Newton's Method.)

Output Skills:

1. Discuss the movement of prices due to shifts in supply and demand, and price equilibrium.
2. Define total demand and free supply and describe the effect on prices of an increase in either.
3. Describe an application leading to a set of linear equations.
4. Tell whether calculations are at matrix level or entry level in linear algebra.
5. (Optimistically) Ability to generalize a simple model from one variable to two and then many.
6. (Part VI) Describe an application of the tangent line.

Other Related Units:

Unit 209: *General Equilibrium: A Leontief Economic Model.*

© 1979 EDC/Project UMAP.
All rights reserved.

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
Paula M. Santillo	Administrative Assistant
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 Kenneth R. Rebman of California State University at Hayward and K.L. Huehn of California Polytechnic 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.

ECONOMIC EQUILIBRIUM: SIMPLE LINEAR MODELS

PART I: SUPPLY AND DEMAND FOR A SINGLE PRODUCT

1. Price Equilibrium

A product is "in equilibrium" or "at its equilibrium price" when supply equals demand for it. This means the amount of the product available from sellers equals the amount that purchasers want to buy. (We include any commodity, service or manufactured product under the general umbrella of "products" here.)

Of course, supply and demand are seldom exactly equal for any product and even if achieved, equilibrium is momentary. If supply exceeds demand, sellers lower their prices to attract buyers; i.e., prices tend to decrease. If demand exceeds supply, the buyers who most want the product bid up its price, and prices rise in response. It is exactly when supply equals demand that these two opposite economic forces are balanced, leaving the price at a standstill. That balanced state of opposing forces is exactly the usual meaning of "equilibrium."

2. The Purpose of This Paper

We will study several versions of a very elementary mathematical model of price equilibrium in this paper. Hopefully, the economic content is clear and interesting, but our main goal is mathematical. We will discover that mathematical economists inevitably find themselves using linear algebra to express their ideas. If we went beyond our simple model to some of the multitude of economic models proposed in recent decades we would find more advanced mathematical tools in use: queueing theory, differential/difference equations, time series forecasting, linear programming, etc. All of these use linear algebra and linearizing methods to achieve practical results—so

a very simple linear algebraic introduction to mathematical economics is appropriate.

Our work here can serve as one instance of an important phenomenon: linear algebra is a basic tool used in virtually all areas of applied mathematics.

3. Assumptions about the Economy

We will assume an economy that is grossly simplified from reality, a classic, competitive, capitalistic economy of the Adam Smith variety. Prices are not controlled by government, buyers or sellers in this economy—they fluctuate freely in response to supply and demand. There are no monopolies, no cartels, no collusion among buyers and sellers. Inflation is not modeled; the entire discussion is in terms of "1967 dollars" or some other standard monetary unit of purchasing power.

Buyers and sellers in our economy have "perfect information." This means that they all know the current supply, demand and price, as if all buying and selling were done in one large auction room with all potential buyers and sellers participating.

4. Supply and Demand depend on Price

Let's analyze supply and demand for one product. Let

(1)
D = current demand for the product (in dollars)
S = current supply of the product (in dollars)
P = current price of the product (dollars/item)

We might have expressed D and S as the amounts demanded and supplied in production-units (boxcar-loads, dozens of eggs, etc.). However, we will want to compare one product to another later, so we'll express D and S in dollars from the start. Once the current price P is known we convert the amounts demanded and supplied into dollars to calculate D and S. (If 4 million dozen eggs are demanded

ERIC

at a wholesale price of 0.5 dollars/dozen, we have a $2 million demand D for eggs.)

In fact, it is natural to regard D and S as *functions of the price P*. This goes hand-in-hand with our assumption of a purely capitalistic economy of value-conscious buyers and profit-conscious sellers. (In reality supply and demand depend on price as well as such emotional elements as style, fads, and the effects of fantasy-oriented advertising.)

5. The One-Product Model

The simplest way to make D and S functions of P is to use straight lines. That is, let's take as our mathematical model

$$(2)\qquad \begin{aligned} D &= a + bP \\ S &= c + dP \end{aligned}$$

where a, b, c, and d are real constants. What can we say about a, b, c, and d on qualitative grounds? As the price P grows, we expect demand to drop (at a higher price there are fewer buyers), so slope $b < 0$. Since $D \geq 0$, we know $a > 0$. And as P grows, the supply S will grow because more companies find it profitable to make the product, hence slope $d > 0$. Figure 1 sketches this situation and shows $c < 0$; let's see why. There will be some price-of-first-supply P_s (namely, the cost of manufacturing) such that no supplier will make the product if $P < P_s$. Thus our straight line must cross the price axis at positive P_s and c, its intercept on the vertical axis, must be negative.

Figure 1 also shows the price-of-last-demand P_d at which the demand line reaches zero: at prices $P > P_d$ no one is interested in buying the product. Only non-negative values of D and S make economic sense, of course. Thus we'll consider P only in the domain $[P_s, P_d]$, as shown in Figures 1 and 2.

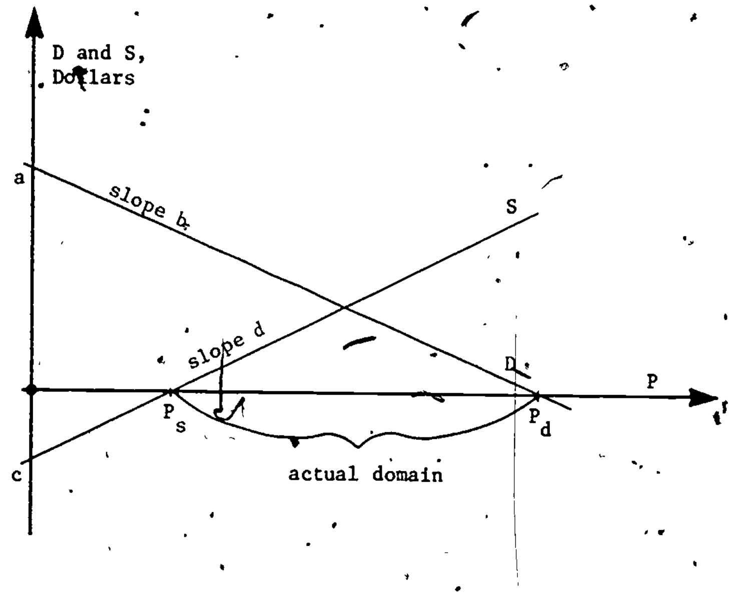


Figure 1. Supply and demand lines for one product.

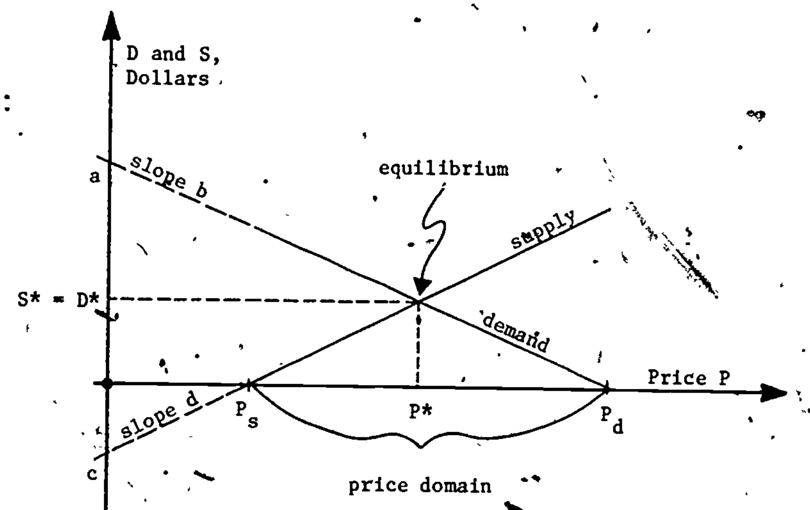


Figure 2. The one-product model.

ERIC

6. Equilibrium in the One-Product Model

The price equilibrium occurs when S = D (supply equals demand). As the sketch shows (Figure 2), there is one price P* for which our model predicts equilibrium (see Exercise 1). The corresponding dollar amounts S* and D* are also sketched. We can calculate P* by setting S = D in (2) and solving to get

$$(3a) \qquad P^* = \frac{a - c}{d - b}.$$

By plugging P* back in for P in either equation of (2) we also find the equilibrium demand/supply level:

$$(3b) \qquad S^* = D^* = \frac{da - bc}{d - b}.$$

What have we achieved with this bit of high-school algebra? Under the crude assumption that simple equations like (2) hold, we can *predict* the price P* a product should sell at and the amounts S* = D* that people should make and will buy! Our next goal is to extend this model to more complicated cases of *general equilibrium* where many competing products are in equilibrium simultaneously.

Exercise 1. Find P*, D*, S* if the formulas for supply and demand are

$D = 22 - 1.5P$

$S = -5 + 5.25P.$

Use these three methods:

a. algebraically solve for P*, D*, S*.

b. identify a, b, c, and d and substitute them in (3a,b)

c. graph the lines and read the equilibrium point off the graph.

Exercise 2. Repeat Exercise 1 for

$D = 30 - 4P$

$S = 6P - 2.$

Exercise 3. Equations (2) with $a > 0$, $b < 0$, $c < 0$, $d > 0$ give linear functions of real P.

a. How do you know that these lines meet exactly once somewhere in the plane?

b. How do you know that the point of intersection $(P^*,S^*) = (P^*,D^*)$ satisfies $P_s \leq P^* \leq P_d$ and $S^* \geq 0$, $D^* \geq 0$? (This requires an economic argument. Show that it is not true based on the mathematical facts alone.)

PART II: THE ANALOGOUS TWO-PRODUCT MODEL

7. Modeling Two Interrelated Products

In a real economy, the supply and demand for a product depends on its price and on the prices of other related products (and on other factors). When products can substitute for each other, this is especially clear. For example, as large cars have become expensive to buy and operate, people have substituted smaller cars. It is logical to think of the supply and demand amounts for large cars as functions of *both* large car and small car prices. The demand function for cars of any size might also depend on the price of labor for having the car serviced and repaired, the price of gasoline, the price of auto parts, etc. It does not depend on most other prices (like that of perfume) but there are important interrelationships amoung products and, to make our mathematical model more realistic, we should include many of those relationships.

As a first step, let's study a two-product market. Put

(4) D_1, S_1, P_1 = demand, supply, price for the first product (large cars, say)

D_2, S_2, P_2 = demand, supply, price for the second product (small cars, say).

We assume D_1, S_1, D_2, S_2 all to be functions of the two prices P_1, P_2 and all are expressed in dollars. Again we assume the simplest functions [compare notation with (2)]:

$$(5)\quad \begin{cases} D_1 = a_1 + b_{11}P_1 + b_{12}P_2 \\ D_2 = a_2 + b_{21}P_1 + b_{22}P_2 \end{cases} \qquad \begin{cases} S_1 = c_1 + d_{11}P_1 + d_{12}P_2 \\ S_2 = c_2 + d_{21}P_1 + d_{22}P_2. \end{cases}$$

The a's, b's, c's and d's are all known real constants. In the context of our large cars—small cars example, we can predict the signs of these constants. The demand for large cars, D_1, should be positive, should decrease as P_1 increases and should increase as P_2 increases (i.e., as small cars become more expensive and hence less attractive to buyers). Thus $a_1 > 0$, $b_{11} < 0$, $b_{12} > 0$. Similarly, $a_2 > 0$, $b_{21} > 0$, $b_{22} < 0$. The supply S_1 of large cars should grow as P_1 increases and also grow as P_2 increases (because higher prices for small cars should shift demand to their competitive large cars and hence stimulate production of large cars). Thus $c_1 < 0$ (for the same threshold-of-manufacturing-costs reasons as before), $d_{11} > 0$ and $d_{12} > 0$. Similarly, $c_2 < 0$, $d_{21} > 0$, $d_{22} > 0$.

8. The Two-Product Model in Vector and Matrix Notation

Of course we will set $S_1 = D_1$ and $S_2 = D_2$ (supply equals demand) and try to calculate the equilibrium prices P_1^*, P_2^*. But that will be easier to do after we arrange (5) as

$$(6)\quad \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}, \qquad \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$$

and shift to the obvious matrix notation. Define

$$(7)\quad \vec{D} = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix},\ \vec{S} = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix},\ \vec{P} = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}, \qquad \vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},\ b = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},\ \vec{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},\ d = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

and rewrite (6) as

$$(8)\quad \vec{D} = \vec{a} + b\vec{P}, \qquad \vec{S} = \vec{c} + d\vec{P}.$$

Compare (2). Notice how naturally (2) has been generalized through the use of linear algebra. The "supply equals demand" equations are now $S_1 = D_1$ and $S_2 = D_2$, i.e.,

$$(9)\quad \vec{S} = \vec{D}.$$

The equilibrium price vector $\vec{P}^* = \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix}$ is the value of $\vec{P}$ we get by substituting (8) into (9):

$$\vec{c} + d\vec{P}^* = \vec{a} + b\vec{P}^*.$$

Elementary matrix algebra leads to

$$(10)\quad (d-b)\vec{P}^* = \overrightarrow{a-c},$$

ERIC Full Text Provided by ERIC

which is a set of linear equations for $\vec{P}^*$. We'll assume that the 2×2 matrix $d - b$ has an inverse and we'll multiply through by $(d-b)^{-1}$ from the left:

(11a) $$\vec{P}^* = (d-b)^{-1}(\overrightarrow{a-c}).$$

Compare this to (3a): multiplication by the matrix inverse of d-b here very naturally replaces multiplication by the reciprocal of scalar d-b there.

Exercise 4. Find the equilibrium prices if

$$D_1 = 12 - 1.5P_1 + P_2$$

$$D_2 = 20 + 2P_1 - P_2$$

$$S_1 = -6 + 1.6P_1 + 2P_2$$

$$S_2 = -5 + 4P_1 + 5P_2$$

by

a. direct calculation from $S_1 = D_1$ and $S_2 = D_2$

b. identification of $\vec{a}$, b, $\vec{c}$, d and substitution in (11a).

---

9. Equilibrium Supply and Demand in the Two-Product Model

Seeking a complete analogy between (3a,b) and the two-product model, we next substitute $\vec{P}^*$ from (11a) into the equations for $\vec{D}$ and $\vec{S}$ to find $\vec{D}^* = \vec{S}^*$. We get:

(11b) $$\vec{D}^* = \vec{a} + b(d-b)^{-1}(\overrightarrow{a-c})$$
$$\vec{S}^* = \vec{c} + d(d-b)^{-1}(\overrightarrow{a-c}).$$

Hmm . . . that doesn't look much like (3b) . . . in fact, it's not so obvious that $\vec{D}^* = \vec{S}^*$ at all. Has our analogy died?

Exercise 5. Substitute your $\vec{P}^*$ solution from Exercise 4 into the equations to calculate $\vec{D}^* = \vec{S}^*$.

Of course, $\vec{D}^*$ and $\vec{S}^*$ in (11b) *are* equal, as we should expect from the way we calculated $\vec{P}^*$, $\vec{D}^*$ and $\vec{S}^*$. A little matrix algebra will show this:

$$\vec{D}^* = \vec{a} + b(d-b)^{-1}(\overrightarrow{a-c})$$

$$= (d-b)(d-b)^{-1}\vec{a} + b(d-b)^{-1}(\overrightarrow{a-c})$$

(because $\vec{a} = I\vec{a} = (d-b)(d-b)^{-1}\vec{a}$)

$$= d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{a}$$

$$+ b(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}.$$

After the cancellation:

(11c) $$\vec{D}^* = d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}.$$

Exercise 6. With this start

$$\vec{S}^* = \vec{c} + d(d-b)^{-1}(\overrightarrow{a-c})$$

$$= (d-b)(d-b)^{-1}\vec{c} + d(d-b)^{-1}(\overrightarrow{a-c}),$$

show that

$$\vec{S}^* = d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}$$

also.

By writing (3a) rather clumsily as

$$D^* = S^* = \frac{da - bc}{d - b} = \frac{da}{d - b} - \frac{bc}{d - b}$$

(12) $$= d\left(\frac{1}{d - b}\right)a - b\left(\frac{1}{d - b}\right)c$$

$$= d(d-b)^{-1}a - b(d-b)^{-1}c,$$

ERIC

we discover that the analogy between (3b) and (11b) is not dead at all, but who would ever write $(da-bc)/(d-b)$ in so complicated a way?! Unfortunately, the liberties we enjoy with scalar arithmetic—we could use any of

$$\frac{da-bc}{d-b} = \frac{ad-cb}{d-b} = (d-b)^{-1}(ad-bc) = (da-bc)(d-b)^{-1}$$

among other forms—are simply not available when b and d are matrices and a, c are vectors. The main problem is that matrix multiplication is not commutative.

Exercise 7. Prove that

$$d(d-b)^{-1} = (d-b)^{-1}d$$

if and only if db = bd.

Exercise 8. Prove that

$$b(d-b)^{-1} = (d-b)^{-1}b$$

if and only if db = bd.

Ordinarily we must expect that matrices b and d will not commute—commutative matrices are the exception and not the rule in mathematics.

If b and d *happen* to commute, we would have

$$\vec{D}^* = \vec{S}^* = d(d-b)^{-1}a - b(d-b)^{-1}c$$

$$= (d-b)^{-1}(da-bc)$$

as in (3b), but we would be wiser to consider the "natural" form of (3a) to be (from (12))

$$D^* = S^* = d\left(\frac{1}{d-b}\right)a - b\left(\frac{1}{d-b}\right)c.$$

Only the commutativity of multiplication of real numbers allows a simpler form like (3a).

10. Matrix Level vs. Entry Level Calculations

We used *exactly* the same steps to calculate P* (see (3a) and (11a)) from the one- and two-product models. In the two-product case, all calculations were at *matrix level:* we thought of $\vec{a}$, $\vec{c}$, $\vec{P}$, $\vec{D}$, $\vec{S}$, b, d, b-d, $(b-d)^{-1}$ as vector and matrix entities, single objects, without thinking about the individual numbers a_j, c_j, ..., b_{ij}, d_{ij}, etc., that make them up. All the calculations in Sections 8 and 9 above were at matrix level.

To actually calculate the components P_1^* and P_2^* of $\vec{P}^*$ in (11a), however, we must calculate the 2 × 2 matrix inverse of d-b and multiply it by the vector $\overrightarrow{a-c}$. Such calculations are at *entry level* (they use the entries, the numbers that form the vectors and matrices). This calculation is quite a bit more complicated than the single division needed to compute P* in (3a). The great beauty and wonder of linear algebra is the extent to which we can do useful calculations at matrix level, as if we had single "numbers" (the matrices and vectors themselves) to work with. Eventually, we must complete our work with grubby arithmetic at entry level, however.

It's to our advantage to seek (at matrix level) a form of our expression that is least painful to work with at entry level. For example, we used

$$\vec{D}^* = \vec{S}^* = d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}$$

to show that $\vec{D}^* = \vec{S}^*$. If we actually use this to calculate D* = S* we will compute one matrix inverse and four matrix multiplications. We don't have to work that hard:

$$\vec{D}^* = \vec{S}^* = \vec{a} + b(d-b)^{-1}(\overrightarrow{a-c})$$

$$= \vec{c} + d(d-b)^{-1}(\overrightarrow{a-c})$$

each involve one matrix inversion followed by only two matrix multiplications.

ERIC

PART III: GENERALIZATION TO n-PRODUCTS

11. The Model with n-Products

Why stop with supply and demand functions that interrelate two products? Suppose an economy is made up of n products, commodities, services, etc., and let

(13) D_j, S_j, P_j = demand, supply, price for the jth product, for $j = 1, 2, \ldots, n$.

We still assume that D_j and S_j depend linearly on the prices but we permit any and all possible interrelationships by using all the prices in each demand or supply function:

(14)
$$\begin{cases} D_1 = a_1 + b_{11}P_1 + b_{12}P_2 + \cdots + b_{1n}P_n \\ D_2 = a_2 + b_{21}P_1 + b_{22}P_2 + \cdots + b_{2n}P_n \\ \vdots \\ D_n = a_n + b_{n1}P_1 + b_{n2}P_2 + \cdots + b_{nn}P_n \end{cases}$$

$$\begin{cases} S_1 = c_1 + d_{11}P_1 + d_{12}P_2 + \cdots + d_{1n}P_n \\ \vdots \\ S_n = c_n + d_{n1}P_1 + d_{n2}B_2 + \cdots + d_{nn}P_n \end{cases}$$

Please compare this to (5) and (2), which are simply the special cases n = 2 and n = 1.

For the reasons discussed in Part I, all $a_i > 0$ and all $c_i < 0$. Most of the b_{ij} and d_{ij} will be zero; they will be nonzero only when i and j are competing products. Consider product 1, which might be large cars, for example. Naturally $b_{11} < 0$ and $d_{11} > 0$: as their prices rise, demand for large cars decreases and supply increases.

Now suppose that products 2, 425 and 7514 (small cars, motorcycles and rapid transit fares, perhaps) compete with product 1. As *their* prices rise, product 1 looks more attractive to buyers, so $b_{1,2}$, $b_{1,425}$ and $b_{1,7514}$ are all positive while the other b_{1j} are all zero. A rising price for a competing product tends to increase the supply of product 1, as explained in Part II. Thus $d_{1,2}$, $d_{1,425}$ and $d_{1,7514}$ will all be positive while the other d_{ij} are zero.

Following the pathway from Equations (5) to (6), we rewrite (14) using matrix products:

(15)
$$\begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_n \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}$$

$$\begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & & & \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}$$

Naturally we introduce these vectors and matrices:

$$\vec{D} = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_n \end{bmatrix}, \quad \vec{S} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix}, \quad \vec{P} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}$$

$$(16)\qquad \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix}, \quad b = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

$$\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}, \quad d = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}$$

and write (15) compactly:

$$(17)\qquad \begin{aligned} \vec{D} &= \vec{a} + b\vec{P} \\ \vec{S} &= \vec{c} + d\vec{P}. \end{aligned}$$

This is an exact copy of (8)!

12. Solution of the n-Product Model

The matrix-level calculations that led us from the two-product model (8) to its solutions (11a,c) are not limited to 2-vectors and 2×2 matrices. One of the great advantages of matrix level work is that it applies to n-vectors and $n \times n$ matrices for any n. Exactly the same reasoning and algebraic operations that led us from (8) to (11a,c) work on (17) to give us its equilibrium solution:

$$(18)\qquad \begin{aligned} \vec{P}^* &= (d-b)^{-1}(\overrightarrow{a-c}) \\ \vec{D}^* = \vec{S}^* &= d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}. \end{aligned}$$

To really calculate P_1^*, P_2^*, ..., P_n^* and the components of D^* and S^* does depend on the dimension n: the entry

level effort needed to calculate $(d-b)^{-1}$ increases rapidly as n increases. We would need a computer to deal with the large n we would want to use in a genuine economic study.

PART IV: HOW DID WE GET THIS FAR?

13. Making a Start

Let's take on the role of the applied mathematician who first developed this model. How do we start? What brainstorms along the way lead to progress and why do they occur? What have we learned from earlier modeling work that we put to use here?

So, we must now imagine that we do not know about this model. An economist comes to us with a question: "Supply, demand and price have these clear intuitive relationships. Can mathematics help us understand the relationship more accurately? Can we predict the price and supply/demand at which a product will/should sell?" We do some preliminary reading and thinking and talk with the economist until we understand the main mechanism: when supply/demand is in excess, this causes a shift in the price downwards/upwards towards a "fair market value price" where the forces of supply and demand are in balance. In that wording, it seems that price is influenced by supply and demand:

$$\text{price} = f(\text{supply}, \text{demand}).$$

We also turn around the language, however: as the price increases/decreases, the supply should increase/decrease while the demand decreases/increases. This wording suggests that supply and demand are influenced by price:

$$\text{supply} = g(\text{price}) \text{ and demand} = h(\text{price}).$$

As experienced applied mathematicians, we prefer to work with the latter approach: we have more equations and can easily express supply = demand. Thus we make the basic decisions that lead to the model of Part I: We'll think about the simplest conceivable economy (one product) by expressing supply and demand as functions of the price. We hope to write down concrete functions:

$$S = g(P) \text{ and } D = h(P)$$

and to solve S = D, a single equation in the one valuable P:

$$g(P) - h(P) = 0$$

for the equilibrium price P*.

The details of Part I now follow when we decide to make g and h very simple (Equations (2)) as a first effort. And we are successful: we predict P*, D*, S* in (3a,b).

14. Improving on Our First Effort

The answer to one question leads to the asking of many more! Here are two reasonable ones:

A. Can we choose functions g and h more realistically? How can we know and measure that we achieve better realism?

B. Can we include more of the complexity of a real, interrelated economy in the model?

Both questions have received lots of attention from applied mathematicians.

Since we can't do many things at once and want to proceed by small steps, we choose arbitrarily to attack (B): What factors of a complex economy should we include? The emotional elements like fads look difficult to get a handle on. We decide to consider two competing products. Copying as much of our successful model in Part I as we can, we decide to make supply and demand

for both products depend on the two prices and we specialize to the easiest concrete functions, in (5).

Aha! A mathematical brainstorm—we can write (5) using matrices as in (6). Our skills with linear algebra take over—we introduce the vectors and matrices of (7), reach the "same" model in (8) that we had in (2), set supply = demand and use matrix algebra to reach P*, D*, S* in (11a,c). Almost nothing is new here: based on our skills with linear algebra we have transformed the success of Part I into results for a more complex economy in Part II.

Now the jump to n-products is easy—we follow the path that linear algebra points out to us, expanding two-vectors and 2 × 2 matrices to n-vectors and n × n matrices. It works again!

15. Hindsight is Perfect

Now that we have the model of Part III and see that the models of Parts I and II are just the special cases n = 1 and n = 2, we know that the n-product model (17) and its solutions (18) are what we were after when we began! We didn't know then that matrix inverses would be involved or that we would find 200 interrelated products just as easy to handle (at matrix level, anyway) as 20 or 2,000, but now that all seems clear, natural and inevitable!

PART V: TWO ECONOMIC INSIGHTS FROM THE MODEL

16. Total Demand

In the one-product model (2):

$$D = a + bP \qquad S = c + dP$$

we might call a the *total demand* because it is the amount of demand if the product were free ($P = 0$) and thus the largest conceivable demand.

Suppose the total demand shifts in our economy from a to $a+\Delta a$, i.e., the economy grows and is able to absorb more of our product. The shift in total demand causes a shift in the equilibrium price from [see (3a)] $P^* = \frac{a-c}{d-b}$ to $P^* + \Delta P^* = \frac{(a+\Delta a)-c}{d-b}$. Thus the resulting change in equilibrium price is

$$\Delta P^* = \frac{(a+\Delta a)-c}{d-b} - \frac{a-c}{d-b} = \frac{\Delta a}{d-b}.$$

This change is positive when $\Delta a > 0$, as we should expect: a larger total demand implies larger demand at any price level and thus upward pressure on prices. Our model agrees with economic common sense. But it lends quantitative *detail* to that common sense, too: we have predicted the amount of the price increase. Common sense alone does not do that.

Exercise 9. In the two-product model, let the total demand change from

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \text{ to } \vec{a} + \vec{\Delta a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \end{bmatrix},$$

causing an equilibrium price change from $\vec{P}^*$ to $\vec{P}^* + \vec{\Delta P}^*$. Calculate $\vec{\Delta P}^*$.

Exercise 10. Repeat Exercise 9 for the n-product model.

Exercise 11. In the one-product model: as the total demand a changes to $a + \Delta a$ there is a change in the equilibrium price, as we've analyzed above. There is also a change in the equilibrium supply = demand level from $S^* = D^*$ to $S^* + \Delta S^* = D^* + \Delta D^*$. Starting with (3b)

$$S^* = D^* = \frac{da - bc}{d - b},$$

we have

$$S^* + \Delta S^* = D^* + \Delta D^* = \frac{d(a+\Delta a)-bc}{d-b}.$$

Calculate $\Delta S^* = \Delta D^*$. Explain why its sign is reasonable, based on economic good sense.

Exercise 12. Repeat Exercise 11 for

a. the two-product model

b. the n-product model.

---

17. Free Supply

Again, in (2),

$$D = a + bP, \quad S = c + dP$$

we can call c the *free supply* or *supply in nature* because it is the supply when $P = 0$. For most products or commodities, $c > 0$ makes no sense because no product can economically be given away for free. However, in many places on the American frontier in the 1800's, fresh water was a free commodity in plentiful supply; until recently, road maps were given away free by gas station owners.

Suppose a product has a free supply c and this supply changes to $c + \Delta c$. This causes a change in the equilibrium price of the product from $P^* = (a-c)/(d-b)$ to

$$P^* + \Delta P^* = \frac{a-(c+\Delta c)}{d-b}; \quad \text{thus } \Delta P^* = -\frac{\Delta c}{d-b}.$$

The sign of ΔP^* again corresponds to economic intuition: as the free supply increases ($\Delta c > 0$), the demand, the amount of the product people will buy, should decrease (since more of the product is supplied free) and thus its price should decrease: $\Delta P^* < 0$. As with total demand, we are able to predict the *amount* of the price drop.

Exercise 13. Repeat the free supply discussion for the two-product model: what change ΔP^* in equilibrium price occurs when the free supply changes from c to $c + \Delta c$?

Exercise 14. Repeat Exercise 13 for the n-product model.

Exercise 15. As the free supply c changes to $c + \Delta c$, the equilibrium price changes by ΔP^* (above) and the equilibrium amount changes from $S^* = D^*$ to $S^* + \Delta S^* = D^* + \Delta D^*$. Calculate $\Delta S^* = \Delta D^*$ for

a. the one-product model

b. the two-product model

c. the n-product model.

PART VI: ARE LINEAR FUNCTIONS CRUCIAL TO THE MODEL?

18. A Job for Taylor's Theorem

Of course, it is unrealistic to take supply and demand as linear functions of a product's price and the prices of its competitors. Yet all our use of linear algebra—our whole ability to calculate equilibrium prices—seems to depend on having such linear functions. How can we resolve this dilemma?

First of all, in the one-product model, how might more realistic functions $D(P)$ and $S(P)$ look? Since the supply increases and demand decreases as prices rise, we take curves with the appropriate monotonicity for $D(P)$ and $S(P)$. When we put such curves (choosing them, as a first example, to be continuous and differentiable) into Figure 2, we arrive at Figure 3. Both curves have $[P_s, P_d]$ as domain, as in Section 5. From Figure 3 it is clear that there is still a unique equilibrium price $P^* \epsilon [P_s, P_d]$.

If we needed to know S and D for all P in the full price domain $[P_s, P_d]$, we would be stuck with these

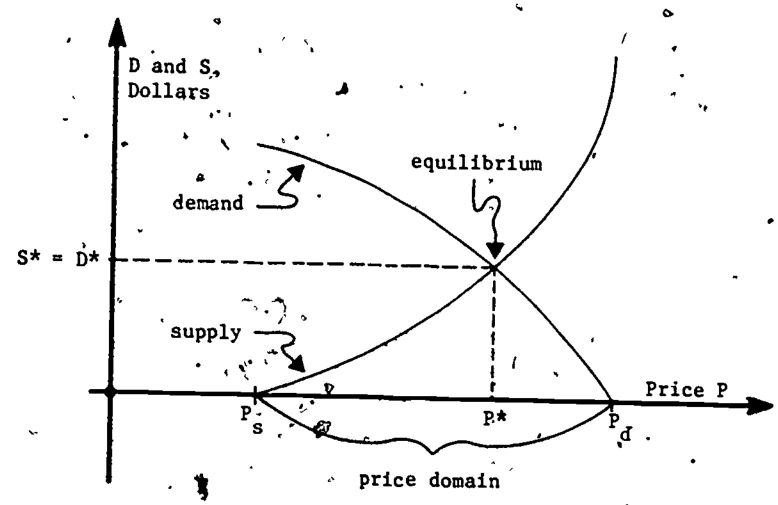


Figure 3. Smooth nonlinear supply and demand functions.

nonlinear curves $D(P)$ and $S(P)$. But recall our goal: we want to calculate P^*, so we only have to think about values of P close to P^*. Probably we know (or can guess on economic grounds) a price P_0 that is fairly close to P^*. We could replace $D(P)$ by the tangent line to $D(P)$ at P_0, getting

$$\hat{D}(P) = D(P_0) + D'(P_0)(P-P_0)$$
(19)
$$= [D(P_0) - P_0 D'(P_0)] + D'(P_0)P.$$

We have written $\hat{D}(P)$ as $a + bP$ above, with constants a and b that we can calculate once we know P_0 and $D(P)$. Recall that $\hat{D}(P)$ is a close approximation to $D(P)$ for P close to P_0.

We can similarly take the tangent line at P_0 to $S(P)$,

$$\hat{S}(P) = S(P_0) + S'(P_0)(P-P_0)$$
(20)
$$= [S(P_0) - P_0 S'(P_0)] + S'(P_0)P$$

as a close approximation to S(P) for P near P_0. Figure 4 shows these two tangent lines.

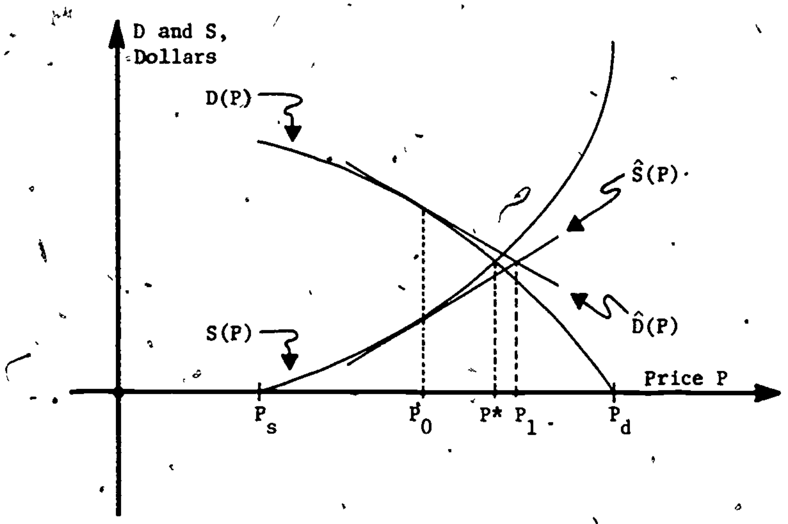


Figure 4. Tangent lines approximate the supply and demand curves.

Equations (19),(20) are a *linearized* version of the nonlinear one-product model. These equations are exactly (2), with

$$a = D(P_0) - P_0 D'(P_0) \qquad b = D'(P_0)$$

$$c = S(P_0) - P_0 S'(P_0) \qquad d = S'(P_0).$$

When we solve for the price equilibrium of the approximate linearized equations we get

$$(21) \qquad P_1 = P_0 - \frac{S(P_0)-D(P_0)}{S'(P_0)-D'(P_0)}.$$

This is of course the price where the tangent lines cross in Figure 4, and $P_1 \neq P^*$.

Exercise 16. Substitute a, b, c, d above into (3a) and thus derive (21). Also calculate, from (3b),

$$\hat{D}(P_1) = \hat{S}(P_1) = \frac{S'(P_0)D(P_0)-D'(P_0)S(P_0)}{S'(P_0)-D'(P_0)}.$$

Exercise 17. (For readers who know Newton's Method of approximately solving f(x) = 0 for a root x given an initial guess x_0 close to the root.)

Equation (21) clearly has a relationship to Newton's Method. What is that relationship? What function f is involved?

Figure 4 suggests that P_1 is a better approximation of P* than our initial approximation P_0 was. Theory (we omit it here) proves this true *if* P_0 is sufficiently close to P*. We can of course repeat the process: taking P_1 as our new guess, we write down equations of the tangent lines to D(P) and S(P) *at* P_1 and use them to calculate P_2. After a few rounds of this, we will get a very good approximation of P*. The method does generalize to multi-product cases.

So, when D(P) and S(P) are smooth functions of the price, with more work we can still approximate P* (and thus D* = S*) closely. The crucial assumption about D and S seems now to be that they change smoothly as P changes. It is not crucial that they be linear.

19. Discontinuous Supply and Demand Curves

And is it realistic to expect that supply and demand curves will be smooth? Unfortunately, no. The supply curve, especially, may have jump discontinuities, as shown in Figure 5.

For there will be threshold values of P (such as P_a and P_b shown) where it becomes economical to open a new factory or put another shift on an assembly line, causing

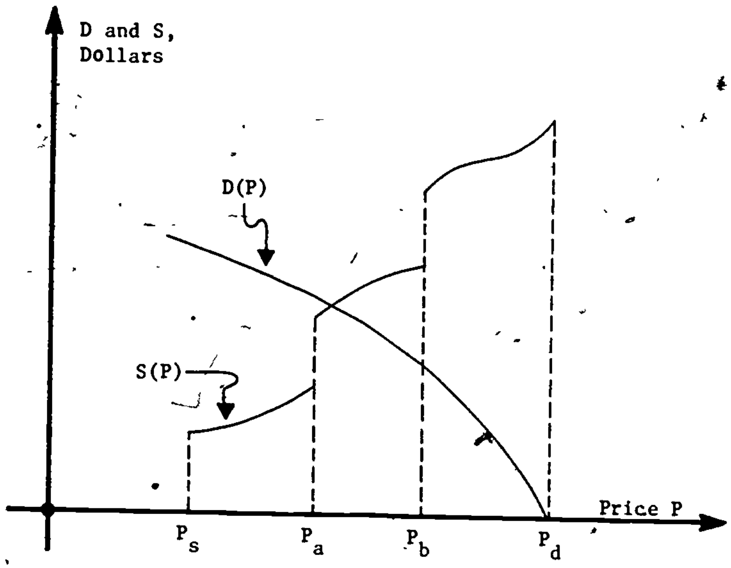


Figure 5. A discontinuous supply curve.

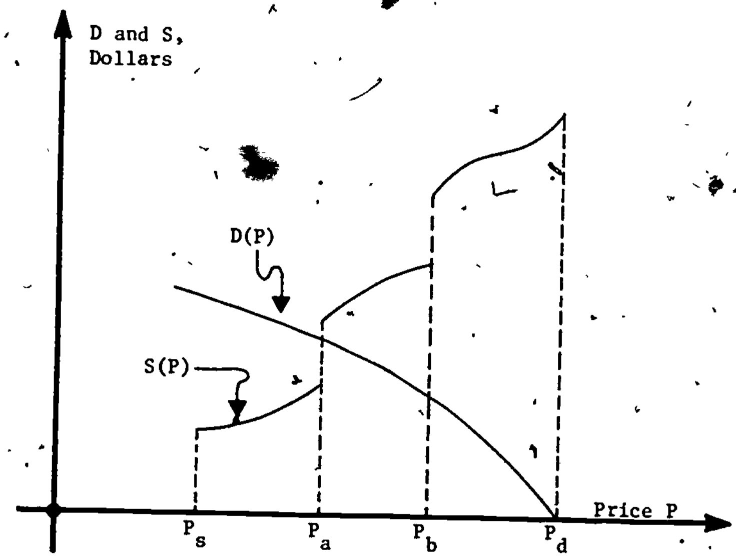


Figure 6. A more difficult case of nonsmooth supply and demand.

supply to jump dramatically. However, in this example, the equilibrium we seek happens to fall in a part of the price domain, namely $[P_a, P_b]$, where both S(P) and D(P) are smooth; we can apply our methods after restricting the price domain to $[P_a, P_b]$. We can draw other examples, like Figure 6, where the method does not apply.

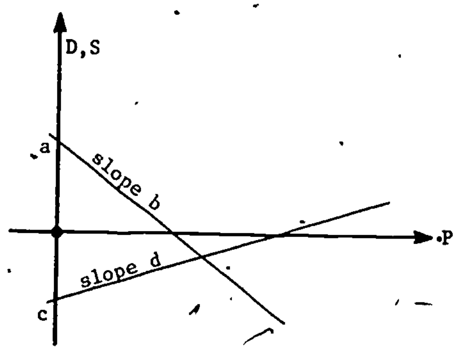
PART VII: SOLUTIONS TO EXERCISES

1. $P^* = 4$ dollars/item, $S^* = D^* = 16$ dollars.

2. $P^* = 3.2$ dollars/item, $S^* = D^* = 17.2$ dollars.

3. a. In a plane, two straight lines with unequal slopes always have exactly one intersection. The slopes here are $b < 0$ and $d > 0$.

   b. Two lines with $a > 0$, $b < 0$, $c < 0$, $d > 0$ yet $S^* = D^* < 0$ can be easily drawn:



On economic grounds, if there is any market for a product, its demand must be positive at $P = P_s$, the minimal price. In that case

$$D(P_s) = a + bP_s > 0$$

$$S(P_s) = c + dP_s = 0$$

$$\Rightarrow P_s = -\frac{c}{d}, \quad \Rightarrow a - \frac{bc}{d} > 0, \quad \Rightarrow ad - bc > 0.$$

Since $d-b > 0$ we have $S^* = D^* = \frac{ad-bc}{d-b} > 0$ from (3b). Since $D^* > 0$ we have $P^* \leq P_d$ and $S^* > 0$ implies $P^* \geq P_s$.

4,5. $\vec{P^*} = \begin{bmatrix} 5 \\ 2.5 \end{bmatrix}, \quad S^* = D^* = \begin{bmatrix} 7 \\ 27.5 \end{bmatrix}.$

6. $\vec{S^*} = \vec{c} + d(d-b)^{-1}\overrightarrow{(a-c)}$

$$= (d-b)(d-b)^{-1}\vec{c} + d(d-b)^{-1}\overrightarrow{(a-c)}$$

$$= d(d-b)^{-1}\vec{c} - b(d-b)^{-1}\vec{c} + d(d-b)^{-1}\vec{a} - d(d-b)^{-1}\vec{c}$$

$$= d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}.$$

7. $d(d-b)^{-1} = (d-b)^{-1}d$

$$\Leftrightarrow (d-b)[d(d-b)^{-1}](d-b) = (d-b)[(d-b)^{-1}d](d-b)$$

$$\Leftrightarrow (d-b)d = d(d-b)$$

$$\Leftrightarrow d^2 - bd = d^2 - db$$

$$\Leftrightarrow bd = db.$$

8. Handle as in solution to 7, above.

9,10. The supply equals demand equation is

$$\overrightarrow{a + \Delta a} + b\overrightarrow{(P^* + \Delta P^*)} = c + d\overrightarrow{(P + \Delta P^*)}$$

$$\Rightarrow \overrightarrow{P^* + \Delta P^*} = (d-b)^{-1}\overrightarrow{(a + \Delta a - c)}.$$

Since

$$\overrightarrow{P^*} = (d-b)^{-1}\overrightarrow{(a-c)},$$

$$\overrightarrow{\Delta P^*} = (d-b)^{-1}\overrightarrow{\Delta a}.$$

11. $\Delta S^* = \Delta D^* = (S^* + \Delta S^*) - S^*$

$$= \frac{d(a+\Delta a)-bc}{d-b} - \frac{da-bc}{d-b} = \frac{d\Delta a}{d-b}.$$

This has the same sign as Δa, as we should expect: when total demand goes up ($\Delta a > 0$), there is naturally an increased demand at all price levels, including price P^*.

12. From (11c) we have

$$\vec{S^*} + \vec{\Delta S^*} = (D^* + \Delta D^*)$$

$$= d(d-b)^{-1}\overrightarrow{(a+\Delta a)} - b(d-b)^{-1}\vec{c}$$

and $\vec{S^*} = \vec{D^*} = d(d-b)^{-1}\vec{a} - b(d-b)^{-1}c.$

Subtraction gives

$$\Delta S^* = \Delta D^* = d(d-b)^{-1}\overrightarrow{\Delta a}.$$

13,
14. Equating supply and demand we get

$$\vec{a} + b\overrightarrow{(P^* + \Delta P^*)} = \overrightarrow{c + \Delta c} + d\overrightarrow{(P^* + \Delta P^*)}$$

$$\Rightarrow \overrightarrow{P^* + \Delta P^*} = (d-b)^{-1}\overrightarrow{(a - c - \Delta c)}.$$

Since $P^* = (d-b)^{-1}\overrightarrow{(a-c)}$, subtraction gives

$$\Delta P^* = -(d-b)^{-1}\overrightarrow{\Delta c}.$$

15. From (11c),

$$S^* + \Delta S^* = D^* + \Delta P^*$$

$$= d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\overrightarrow{(c+\Delta c)}$$

while

$$S^* = D^* = d(d-b)^{-1}\vec{a} - b(d-b)^{-1}\vec{c}.$$

Thus

$$\overrightarrow{\Delta S^*} = \overrightarrow{\Delta D^*} = -b(d-b)^{-1}\overrightarrow{\Delta c}.$$

This may be applied for 1, 2 or n.

17. We really want to solve S(P) = D(P), i.e.,

$$S(P) - D(P) = 0,$$

for P. Thus f is the supply function minus the demand function and the usual Newton's Method formula

$$x_1 = x_0 - \frac{f(x)}{f'(x)}.$$

is exactly (21).

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.)

umap

UNIT 209

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

GENERAL EQUILIBRIUM: A LEONTIEF ECONOMIC MODEL

by Philip M. Tuchinsky

"Like Adam Smith and The Wealth of Nations, Marshall and Principles of Economics, and Keynes and The General Theory, Leontief and 'Input-Output' are becoming permanent words in the economics vocabulary."

— Walter Isard and Phyllis Kaniss in Science, Vol. 182, Nov. 9, 1973, p. 571.

APPLICATIONS OF LINEAR ALGEBRA TO ECONOMICS

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

GENERAL EQUILIBRIUM: A LEONTIEF ECONOMIC MODEL

by
Philip M. Tuchinsky
Department of Mathematical Sciences
Ohio Wesleyan University
Delaware, Ohio 43015

TABLE OF CONTENTS

ED 036 478

ERIC

Intermodular Description Sheet: UMAP Unit 209

Title: GENERAL EQUILIBRIUM: A LEONTIEF ECONOMIC MODEL

Author: Philip M. Tuchinsky
Department of Mathematical Sciences
Ohio Wesleyan University
Delaware, Ohio 43015

Review Stage/Date: III (peer reviewed & revised) 12/28/77

Classification: ECONOMICS/ELEM LIN ALG

Computer Projects are natural to this application but not essential. Realistic data for Yugoslavian economy is included.

Estimated Teaching Time: A total of 1 hour including discussion of computer project output.

Prerequisites: Solution of a nonsingular system of linear equations; matrix multiplication; matrix inverse. No economic preparation is assumed.

Output Skills:

1. Define/explain the one-product company as an input-output process unit.
2. Define meaning of entry a_{ij} of Leontief matrix.
3. Discuss an application leading to large sets of linear simultaneous equations.
4. Calculate equilibrium production levels for a multi-sector Leontief economy.
5. Calculate $(I-A)^{-1}$ approximately using geometric series for matrices.
6. Explain why $(I-A)^{-1}$ will have entries ≥ 0 if matrix A has entries ≥ 0.
7. Discuss money as simply one more product in an economy.
8. Discuss "economics of scale" vs. "constant returns to scale" and inevitability of latter in a linear model.
9. Convert real economic data (input-output flow by sectors) into Leontief matrix entries.

Structure of the Module: Sections 1, 2 and 5 form a satisfactory applied unit about the basic Leontief open model. Section 3 is optional; it discusses use of geometric series to approximate $(I-A)^{-1}$, which arose in Section 2. Section 4 can follow Section 2 or 3 and extends the model so that labor costs are included; it can be omitted also.

UMAP editor for this module: Solomon Garfunkel

Other Related Units:

© 1978 EDC/Project UMAP.
All Rights Reserved.

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 Weitzel	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Dianne Lally	Project Secretary
Paula M. Santillo	Financial Assistant/Secretary
Jack Alexander	Editorial Consultant
Edwina Michener	Editorial Consultant

NATIONAL STEERING COMMITTEE

W.T. Martin	MIT (Chairman)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
James D. Forman	Rochester Institute of Technology
Ernest J. Henley	University of Houston
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
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 Kenneth R. Rebman for his review 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.

1. INTRODUCTION

With his famous 1941 book [3], Professor Wassily Leontief began the study of the economy as an input-output system. For this work, he received the Nobel Prize in Economics in 1973.

His method has been applied in more than fifty nations and international agencies as a predictive tool for economic planning. In Section 5 we will discuss the uses and shortcomings of the method; first, let's examine it in some detail.

1.1 One-Product Companies as Input-Output Machines

Imagine an economy made up of companies that each make one product. (We can, at least in theory, mentally split up a multi-product company to satisfy this.) The manufacture of one unit of the product requires a known recipe of input products, commodities and services. For example, the manufacture of a ton of writing paper requires specific amounts of wood fiber, recyclable paper, water, capital investment in the form of machinery, labor, electricity, etc. A paper-making company can be thought of as an "input-output machine" that converts a recipe of inputs into a unit of output many times a day. We will think of all the companies in this way, each with its own recipe for making its one product.

1.2 Consumers and Companies Form a Complex Economy

The companies are elaborately interrelated. The output of a steel company is an input to a vast number of other companies which make autos, appliances, steel nuts and bolts, steel alloys, and the thousands of other steel-using products. The output of a textile manufacturer is an input to the manufacture of clothing, upholstered furniture, carpet, etc., and cloth is also sold directly to the public. We draw a distinction between companies (which convert a product into other products so that the input products are not really consumed but are "replaced" on the market by the output products) and the public, considered as the "final consumers" of finished goods. The amount of a product demanded by the public is the "final demand" for that product; suppliers (companies) must satisfy that final demand in addition to providing input materials to other companies. For some products (like steel), final demand is almost zero, while for others (like blenders), final demand is a very large fraction of the total demand.

1.3 The Problem: To Balance Supply and Demand

A sensible economic question: how much of each product should be produced to closely satisfy the total demand for the product by all users? That is, how can we match outputs to inputs throughout an elaborately interconnected economy? To answer this question is to find a "general equilibrium" (as economists say); that is, we seek the production amount for each good that will simultaneously make supply equal demand for them all.

Leontief applied linear algebra to this problem. We'll look at his simplest model in this paper. From a knowledge of the final demand for each product (that is, the market basket of all goods that the public is to buy), *we can calculate* the production amounts that will supply that final demand. Some restrictive economic assumptions are involved.

2. LEONTIEF'S MODEL

2.1 Notation for Production and Final Demand Levels

We will look at an economy made up of n companies, each creating one product, commodity or service from a fixed recipe of input "ingredients." We assume that prices are constant and known for each product; we will say that the i^{th} company makes x_i dollars worth of its product. Let d_i dollars of this be the "final demand" (sales to households) of product i while the rest of x_i is used as inputs by other companies. We take the d_i as known (thus assuming that there is some known mix of products that the public is ready to buy) and hope to calculate the x_i values needed to produce a "final demand vector" $(d_1, d_2, \ldots, d_n)$ containing the desired final amounts of all the products. This vector is really just the total public "market basket" of all products consumed, in dollar amounts.

2.2 Leontief's Input-Output Coefficients

Next, we need to express the recipe that the i^{th} company converts into *one* dollar's worth of its output. (That recipe, multiplied through by x_i, will yield x_i dollars of output for the i^{th} company. This is an assumption called "constant returns to scale;" more about it later.) Let a_{ij} be the dollar amount of product i that is used to make one dollar's worth of product j. Thus $a_{47} = .23$ would mean that, to make a dollar's worth of product 7, we use 23¢ worth of product 4. The full mix of products 1, 2, 3, ..., n used to make one dollar's worth of product j is a_{1j} dollars of product 1, a_{2j} of product 2, $a_{3j}, \ldots, a_{nj}$. These numbers form the j^{th} column of the matrix $A = (a_{ij})$. Since less than a dollar's worth of inputs are used in making a dollar's worth of the output (or company j would be out of business), we know that

$$\sum_{i=1}^{n} a_{ij} < 1 \text{ and of course } a_{ij} \geq 0.$$

Thus the columns of A have sums of less than one. (The rows of A have no comparable economic interpretation.)

2.3 Concise Summary of the Notation

For i and j, each running 1,2,3,...,n,

(1)
- x_i = total dollars produced of product i,
- d_i = total dollars worth of product i that is consumed by households = "final demand",
- a_{ij} = amount in dollars of product i used in making one dollar's worth of product j,

2.4 Equating Supply and Demand

The supply of product i will be x_i dollars. The demand for product i will be d_i dollars of final demand, plus $a_{i1}x_1$ dollars (of product i) used in making the x_1 dollars of product 1, plus $a_{i2}x_2$ dollars (of product i) used in making x_2 dollars of product 2, and so on. The "supply equals demand" equation for product i is

$$x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + d_i . \tag{2}$$

We have such an equation for each i = 1,2,...,n, thus n equations in all. Our goal, once again, is to calculate all the x_i from given d_i and known technological constants a_{ij}. The model provides for the use of each product as an input to every other product including itself; of course, many of the a_{ij} will be zero.

2.5 The Model in Matrix Notation

We are ready to switch to matrix notation: put

$$\vec{x} = \begin{pmatrix} x_1 \\ n_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}, \quad \vec{d} = \begin{pmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{3}$$

Then the n equations of (2) may be compactly written

$$\vec{x} = A\vec{x} + \vec{d} \tag{4}$$

or

$$(I - A)\vec{x} = \vec{d} . \tag{5}$$

The problem is now almost solved. In (5), we know the nxn matrix I - A and the n-vector $\vec{d}$. Then (5) is simply a set of n non-homogenous linear equations with the wanted x_i as the unknowns.

2.6 Solving for the Production Levels

You recall that a set of equations like (5) may have no solution, exactly one solution, or infinitely many solutions. In our case, although we will not prove it, there must be exactly one solution. In fact $(I - A)^{-1}$ must exist for our given matrix A. (This is true for any matrix where $a_{ij} \geq 0$ and the column sums satisfy $\Sigma_i a_{ij} < 1$.) We may use the inverse to solve for $\vec{x}$ in (5):

(6) $\vec{x} = (I - A)^{-1} \vec{d}$.

We have achieved our goal: to produce a market basket $\vec{d}$ of final consumer goods, we should produce the amounts $\vec{x}$ given in (6).

Two questions arise at once. One is economic: can the consumer afford to pay for the market basket $\vec{d}$? Consumers usually pay for goods and services by exchanging their labor. Can we fit the cost of labor into the model, where it has not been mentioned so far? We'll discuss this in Section 4.

The other question is mathematical: in (6) we are asked to calculate $(I - A)^{-1}$ for a matrix that may well be 500 x 500 or even 10,000 x 10,000: we must include many companies to treat the economy with any realism. Is there some way to calculate $(I - A)^{-1}$ easily? See Section 3.

2.7 Exercises

1. Although we have considered individual companies making specific products like stoves, the model can be applied to broadly-drawn sectors of an economy. This "two-company" fictional example is taken from [8], page 61: In hundreds of billions of dollars, let the flow be:

		CONSUMPTION			
	FROM / TO →	AGRICULT.	MFG.	HO.	TOTAL
PRODUCTION	Agriculture	4	6	10	20
	Manufacturing	8	18	4	30
	Households	8	6	6	20
	Total	20	30	20	70

This array should be read as follows: there is a total flow of 70 (hundred-billion dollars) among two "companies," agriculture and manufacturing, and one "open sector," households. Agriculture uses 4 units of its own production, 8 units of manufacturing production (fertilizer, machines, etc.) and 8 units of household production (labor), 20 units in all, to produce 20 units which are distributed as follows: 4 to agriculture, 6 to manufacturing, 10 to households. The input of 6 units of household production (labor) to household consumption is domestic labor -- the labor of housewives, for example.

The data above is not the Leontief input output array we have studied, but we can calculate the Leontief matrix from it easily. The recipe of inputs to agriculture is 4/20 from agriculture and 8/20 from manufacturing. The recipe of inputs to the manufacturing sector is 6/30 from agriculture and 18/30 from manufacturing. Thus the technical matrix and final demand vector are

$$A = \begin{pmatrix} 4/20 & 6/30 \\ 8/20 & 18/30 \end{pmatrix} \doteq \begin{pmatrix} .2 & .2 \\ .4 & .6 \end{pmatrix} \text{ and } \vec{d} = \begin{pmatrix} 10 \\ 4 \end{pmatrix}.$$

a. Using the A and $\vec{d}$ just above, hand-calculate the solution $\vec{x}$ of the set of linear equations

$$(I - A)\vec{x} = \vec{d}.$$

b. Calculate $(I - A)^{-1}$ and then find $\vec{x}$ again from $\vec{x} = (I - A)^{-1}\vec{d}$.

c. How could you have predicted your answer to a. and b. from the table in the exercise?

(Continued in Exercise 5.)

2. In this exercise, we alter Exercise 1 so that agriculture, manufacturing and households are the three sectors or "companies" involved, while savings is the open sector. Each "company" produces its product (which is still labor in the case of the households) so as to supply the other two companies and create a final product called *investment*, while invested funds, called *savings*, are inputted only to the household sector (say, to build houses. This example, again in hundreds of billions of dollars, is from [8], page 182:

CONSUMPTION

↓FROM / TO →	AGRICULT.	MFG.	HO.	INVESTMENT	TOTAL
Agriculture	4	6	3	7	20
Manufacturing	8	18	1	3	30
Households	8	6	4	2	20
Savings	0	0	12	0	12
Total	20	30	20	12	82

a. Convert this data into a 3-company Leontief model by finding A and $\vec{d}$ by the method explained in Exercise 1.

b. Predict $\vec{x}$ from the table above without any use of the Leontief model.

c. Solve $(I - A)\vec{x} = \vec{d}$ for $\vec{x}$. Show your calculations in detail.

d. Calculate $(I - A)^{-1}$ and then get $\vec{x}$ from $\vec{x} = (I - A)^{-1}\vec{d}$. Show your calculations.

Answers to b, c, d should all be the same. This exercise is continued in Exercise 6.

3. HOW TO CALCULATE $(I - A)^{-1}$ EASILY

3.1 An Old Acquaintance Returns

There is an elegant way to calculate $(I-A)^{-1}$. In the back of your mind, you should think of the matrix A as though it were a single number (say a) and of I as though it were 1. Then $(I-A)^{-1}$ becomes analogous to

$$\frac{1}{1-a} \quad \text{and} \quad \frac{1}{1-a}$$

should make you think of -- geometric series! You recall the geometric series formula

(7) $$1 + a + a^2 + a^3 + \ldots = \frac{1}{1-a} \quad (\text{if } |a| < 1).$$

For our matrix A, the conditions $a_{ij} \geq 0$ and "column sums $\Sigma_i a_{ij} < 1$" take the place of $|a| < 1$ and it is true that

(8) $$I + A + A^2 + A^3 + \ldots = (I - A)^{-1},$$

a complete analogy to (7).

3.2 How the Series Aids Calculation of $(I-A)^{-1}$

We'll consider a plausibility argument for (8) shortly (an ironclad proof is just a little beyond the intended level of this paper because it requires "matrix norms"), but first let's see the usefulness of (8). If A^4, A^5, A^6 and all the higher power terms are "negligibly small," then the 4-term partial sum $I + A + A^2 + A^3$ is a good approximation of the hard-to-compute matrix inverse $(I-A)^{-1}$ needed for (6). (The inverse *is* nasty to compute: think of the methods of matrix-inversion you know and consider applying them to a 30 x 30 or 2000 x 2000 matrix $I-A$.) In fact, a partial sum of quite a few terms from (8) is cheap and convenient to compute by comparison to direct computation of $(I-A)^{-1}$.

3.3 An Example

Just to see how the calculation goes, put

$$A = \begin{bmatrix} .1 & .2 & .1 \\ 0 & .2 & 0 \\ .2 & 0 & .1 \end{bmatrix} \text{ so that } I - A = \begin{bmatrix} .9 & -.2 & -.1 \\ 0 & .8 & 0 \\ -.2 & 0 & .9 \end{bmatrix}$$

and, to four decimal places,

$$(I-A)^{-1} = \begin{bmatrix} 1.1392 & .2848 & .1266 \\ 0 & 1.25 & 0 \\ .2532 & .0633 & 1.1392 \end{bmatrix}$$

You should check *all* the calculations here. Use an electronic calculator. Let's look at some partial sums of the geometric series:

$$I + A = \begin{bmatrix} 1.1 & .2 & .1 \\ 0 & 1.2 & 0 \\ .2 & 0 & 1.1 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} .03 & .06 & .02 \\ 0 & .04 & 0 \\ .04 & .04 & .03 \end{bmatrix}, \text{ thus } I+A+A^2 = \begin{bmatrix} 1.13 & .26 & .12 \\ 0 & 1.24 & 0 \\ .24 & .04 & 1.13 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} .007 & .018 & .005 \\ 0 & .008 & 0 \\ .010 & .016 & .007 \end{bmatrix}, \text{ thus } I+A+A^2+A^3 = \begin{bmatrix} 1.137 & .278 & .125 \\ 0 & 1.248 & 0 \\ .25 & .056 & 1.137 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} .0017 & .005 & .0012 \\ 0 & .0016 & 0 \\ .0024 & .0052 & .0017 \end{bmatrix}, \text{ thus}$$

$$I + A + A^2 + A^3 + A^4 = \begin{bmatrix} 1.1387 & .283 & .1262 \\ 0 & 1.2496 & 0 \\ .2524 & .0612 & 1.1387 \end{bmatrix}$$

This five-term partial sum is convincingly close to $(I-A)^{-1}$. This example was fabricated so that the infinite series would converge within a few terms; entries like .1 and .2 become rapidly smaller when multiplied by one another in matrix products. However, in a large matrix A the entries would mostly be small and many would be zero. Remember, all a_{ij} are ≥ 0 and the column sums are less than one. The geometric series is a practical way to approximate $(I-A)^{-1}$.

3.4 Why Geometric Series Extends to Matrix Cases

A plausibility argument for the truth of (8) was promised. This matrix calculation closely mimics the usual proof of the scalar case (7): notice that, for any finite partial sum,

$$(9) \qquad (I + A + A^2 + \ldots + A^{k-1})(I - A) = I - A^k.$$

All other terms cancel out. For matrices like our A with small positive entries, the powers A^k approach the zero n x n matrix 0 as k increases, because products of small positive numbers get smaller. To say

$$\lim_{k\to\infty} A^k = 0$$

means that all n^2 of the matrix entries approach zero as k increases, and this is true for the matrices we are studying. Now let $k\to\infty$ in (9):

$$(I + A + A^2 + A^3 + \ldots)(I - A) = I - 0 = I.$$

But this exactly says that $(I-A)^{-1} = I + A + A^2 + A^3 + \ldots$

3.5 $(I-A)^{-1}$ Will Have Nonnegative Entries

From (8) we can conclude that all the entries of $(I-A)^{-1}$ will be ≥ 0. (This means that negative production levels x_j cannot arise in (6), which is comforting: we would throw away a model that failed to yield all the $x_j \geq 0$.) To see that $(I-A)^{-1}$ cannot have negative entries, simply recall that $a_{ij} \geq 0$ for all i,j. Thus I, A, A^2, A^3, A^4 ... all contain entries that are ≥ 0 (think about

the multiplication $A \cdot A = A^2$, and so on.) Then their sum $I + A + A^2 + \ldots = (I-A)^{-1}$ also has non-negative entries.

3.6 Exercises

3. Using $A = \begin{pmatrix} .6 & .4 \\ .3 & .4 \end{pmatrix}$:

   a. show that $(I - A)^{-1} = \begin{pmatrix} 5 & 10/3 \\ 5/2 & 10/3 \end{pmatrix}$

   b. Write and run a short computer program that calculates and prints $I + A$, $I + A + A^2$, $I + A + A^2 + A^3$, $I + A + A^2 + A^3 + A^4$, etc. Print partial sums until you have $(I - A)^{-1}$ well approximated. This will take quite a few terms.

   c. How many terms must you include in the partial sum in b. before you have approximated $(I - A)^{-1}$ within .5 in each entry? Within .05? Within .005?

4. Verify the matrix calculation in (9). State each law of matrix algebra you use (e.g., the "left distributive law").

5. (Exercise 1, continued) For the matrix A and demand vector $\vec{d}$ of Exercise 1, calculate by computer successive approximate solutions

   $(I + A)\vec{d}$
   $(I + A + A^2)\vec{d}$
   $(I + A + A^2 + A^3)\vec{d}$
   etc.

   These will converge slowly to your solution $\vec{x}$ in Exercise 1.

6. (Exercise 2, continued) For the matrix A and demand vector $\vec{d}$ of Exercise 2, calculate successive approximations of $\vec{x}$ by using partial sums of the series for $(I - A)^{-1}$.

4. MODELING LABOR IN LEONTIEF'S ECONOMY

4.1 The Value of Labor

Now let's turn to the economic question we raised in Section 2.6: can the public contribute enough labor to the economy to pay for the final-demand market basket $\vec{d}$ it has ordered? It is easy to calculate the value of labor in our economic model. To make one dollar's worth of the j^{th} product, we recall, involves a_{1j} dollars of product 1, a_{2j} dollars of product 2, ..., and a_{nj} dollars of product n; in all the dollar's worth of product j contains

$$a_{1j} + a_{2j} + \dots + a_{nj} < 1$$

dollars worth of input materials made by the n companies. The maximum amount that can be paid for labor is

$$a_{0j} = 1 - \sum_{i=1}^{n} a_{ij}$$

dollars per dollar's worth of product j that is manufactured. The new constant a_{0j} (for $j = 1, 2, 3, \dots, n$) are labor's maximal slice of the pie. When x_j dollars of product j are made, labor receives $a_{0j}x_j$ dollars in pay. Thus the total economy-wide earnings of labor are at most

$$\sum_{j=1}^{n} a_{0j}\, x_j = (a_{01}, a_{02}, a_{03}, \dots, a_{0n}) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \vec{a}_0 \cdot \vec{x}.$$

Here $\vec{a}_0$ denotes the row-vector $(a_{01}, a_{02}, \dots, a_{0n})$.

4.2 Labor's Earnings and Consumption are Equal

The total worth of the final demand vector $\vec{d}$ is $d_1 + d_2 + \dots + d_n$ dollars. Thus the final demand vector $\vec{d}$ is *feasible* (can be paid for by the public) if

(12) $$\vec{a}_0 \cdot \vec{x} \geq d_1 + d_2 + \dots + d_n .$$

We will now prove that equality must hold in (12), $\vec{a}_0 \cdot \vec{x} = d_1 + d_2 + \dots + d_n$, if we use production levels $\vec{x}$ calculated from the Leontief model, from (6), and pay labor its maximal earnings, the a_{0j} from (10). We will be proving that labor's earnings exactly pay for the "market basket" that households consume. This turns out to be true because we have build "conservation of value" into the model: the value of output is equal to the value of input products and labor if we use (10).

The proof involves more matrix algebra. First, notice that, by introducing an n-vector containing all ones,

$$u = (1,1,1,\dots,1)$$

we can write the right side as a matrix product:

$$d_1 + d_2 + \dots d_n = (1,1,\dots,1) \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} = u \cdot \vec{d}.$$

Now we can write the n equations of (10) compactly as

(13) $$(a_{01}, a_{02}, \dots, a_{0n}) = (1,1,\dots,1) - (1,1,1,\dots,1)\, A = u \cdot (I-A).$$

The $(1,1,\dots,1)\,A$ term here gives the column sums that appear in (10). Now

(14) $$\begin{aligned} a_0 \cdot \vec{x} &= (a_{01}, \dots, a_{0n}) \cdot \vec{x} \\ &= u\,(I-A) \cdot \vec{x} && \text{from (13)} \\ &= u\, \underbrace{(I-A)(I-A)^{-1}}_{\text{cancels}}\, \vec{d} && \text{from (6)} \\ &= u \cdot \vec{d} = d_1 + d_2 + \dots + d_n . \end{aligned}$$

5. ABOUT THE MODEL AND ITS USES

5.1 Open and Closed Leontief Economies

The model we have looked at is known as Leontief's *open model* because of the separate treatment of companies and public. In a closed model, the public (or labor force or households) is treated as one more company to which the input recipe is the market basket $\vec{d}$, while the output is the labor ingredient in the input recipe of more traditional companies. As we have just seen, the dollar-worth of inputs to the household sector will equal the dollar-worth of its output (labor) in the same way that the inputs of goods and labor to a manufacturer equal the value of its output. The open and closed models are equivalent. The distinction between "final consumption goods" in our open model and inputs that the household sector "processes" into an output product called labor in a closed model is of economic interest, but makes no mathematical difference.

5.2 Profit and Savings Have Been Included

We have emphasized so strongly the equal value of the inputs and outputs of each company that you may wonder how a company can make any profit. In fact, profit is one of the input ingredients to each company. One of the products or commodities that flows through the economy we have modeled is money. The paper manufacturer mentioned at the beginning of this paper really receives a few pennies of money along with the physical inputs (like wood fiber) and labor-time in exchange for the dollar's worth of output (paper) made from these inputs. The public receives some money as part of its market basket -- this is savings. Money is simply one of the n products "manufactured" by n companies: one company in this economy is a commercial bank. Certainly the role played by money is unrealistically simplified -- we have not built an investment or credit structure into the model. That can be done, however.

This model is only concerned with the complex flow of goods among the companies and consumer/labor sector of the economy. No risk is modeled -- each company knows how much of its product it can sell to the public; prices do not change. We are modeling the *distribution* process of the economy, not its other aspects.

5.3 Using Linear Algebra in Economics -- Benefits and Difficulties

Leontief has chosen linear algebra as his mathematical tool. He benefits from that -- to find $\vec{x}$ in terms of $\vec{d}$ we simply solve a (large) set of linear equations, which we know how to do. The great contribution of Leontief's models is that they permit actual calculation of general equilibria in terms of input data (the technological constants a_{ij} and final demands d_j) which we can hope to actually know. Other models that attempt to equate supply and demand (i.e., to study general equilibrium) tend to be so theoretical that no useful numbers can be calculated from them; one can instead use them to prove that one or more general equilibria must exist! In fact, several Leontief models have been fully researched and are in use as planning devices.

But there is a price paid for the use of linear algebra; the models are subject to a key criticism. We have assumed "constant returns to scale," as economists say. This means that, if a specific recipe of inputs makes one dollar's worth of output for a given company, then N copies of that recipe will make exactly N dollar's worth of output. In reality, companies can reduce the cost-per-unit-produced by enlarging their production. For example, once an assembly line has been purchased and installed, it can be used for one, two or three eight-hour shifts daily. When used for three shifts, the capital investment in the machinery is spread over three times more output than is the case if one shift is used. The input of capital to any one unit of production is much less when the machines are used to capacity. (There are extra expenses involved in running machinery around the clock -- repair and maintenance expenses, extra pay for work done on night shift, etc. -- but these expenses are easily overcome by the three-to-one savings.) It is generally less costly (per unit of production) to mass-produce more of any product than less; that is, there are "economies of scale." This phenomenon is an important reason for the clear tendency toward large corporations in our economy.

Linear equations like (5) cannot deal with economies of scale. Indeed, doubling $\vec{d}$ in (5) leads to a new solution $\vec{x}$ that is double the old $\vec{x}$. "Constant returns to scale" is an inevitable assumption if linear algebra's calculation advantages are to be exploited.

The use of constant technological data, the a_{ij}, has also been widely criticized. The input-output process in each company is excessively rigid in the model. In reality, a furniture manufacturer might very casually switch from one upholstery cloth to another. However, that amounts to creating a whole new economy in our model! The recipe for the furniture maker must be altered (changing a column of A) and new production levels must be calculated for all companies. This is another price for the use of linear algebra -- all the companies are rigidly interconnected.

5.4 The Model is Widely Used as a Planning Aid

When a nation, a region or a city needs to know the impact that alternative development projects -- a steel mill, a cultural center, an auto assembly line, a food processing plant -- will have if built, input-output analysis is of great help. The model can predict the flow of goods and services, including transportation needs, new employment and pollution problems (such factors may be added to the model we have discussed) and point to serious shortfalls or oversupplies in the current economy. Its answers are only approximate, of course, but they give crucial insight into a very complex problem.

The United Nations and the World Bank use Leontief models. The Bureau of Labor Statistics of the U.S. federal government has been a major sponsor of Leontief's research and employs a massive model of the U.S. economy. Government agencies of more than fifty other countries, including the Scandinavian nations, Western Europe, Eastern Europe, the USSR and many developing nations use such models.

5.5 The Model's Great Impact on Economics

In *Science* magazine, Walter Isard and Phyllis Kaniss (10) reviewed Leontief's contributions at the time of his winning the Nobel Prize. They highlight the power of input-output analysis for planning, but concede that the model's predictions have contained large errors when

> WASSILY W. LEONTIEF was born in Leningrad in 1906. He fled Communist rule in Russia in the early 1920s with his family. At the age of 22 he completed a doctorate at the University of Berlin. From 1929 to 1931 he was economic advisor to the Chinese government; in 1931 he joined the National Bureau of Economic Research in New York. His main input-output methodology matured during the '30s. He was chief of the Russian Economic Subdivision of the Office of Strategic Services during World War II. Leontief has been a professor at Harvard since 1946. Sources (9) and (10).

the method has been used by inexperienced planners. Such errors can arise, they point out, in these key ways:

-- constant coefficients in the matrix A make the "recipes" of inputs used by companies inflexible;

-- the effects of inevitable changes in technology are not included;

-- the extensive and precise data needed for the model is often unavailable, "borrowed" from another region or nation, etc. This has been a problem in developing nations, especially.

-- one product can sometimes be substituted for another in our economy; Leontief does not include this possibility in his models.

Aside from planning and predictive uses, Isard and Kaniss report a major impact upon economics. Since the model requires complete, consistent data, it has forced many nations to take economic data gathering more seriously. Uniform definitions of products and sectors of an economy and uniform accounting procedures have been needed; thus planning and data collection agencies in many nations have coordinated their programs. Much easier comparative study of related national economies has resulted.

Writing in *Newsweek* (9), Paul Samuelson (himself a famous doctoral student of Leontief's at Harvard) mentioned these uses of input-output analysis:

-- As the Vietnam War wound down, Leontief predicted the results of the shift of a billion dollars in gross national product from war to peacetime production. He concluded that there would be an *expansion* in employment.

-- Leontief discovered that exports from the United States are more labor intensive than our imports, confounding those who decry the use of "cheap foreign labor" as a source of unemployment here. His conclusion is that the net result of importation and exportation is to increase use of U.S. labor.

-- The U.S. Congress discovered the great impact of steel-price raises on inflation in the United States.

6. REFERENCES

Advanced References:

1. R.G.D. Allen, Mathematical Economics, 2nd ed., St. Martin's Press, New York, 1959.
2. David Gale, Theory of Linear Economic Models, Prentice-Hall, Englewood Cliffs, N.J., 1961.
3. Wassily W. Leontief, The Structure of the American Economy, 1919-1929, Harvard University Press, Cambridge, Mass., 1941.
4. Wassily W. Leontief, The Structure of the American Economy, 1919-1929, 2nd ed., Oxford University Press, Fairlawn, N.J., 1951.
5. Wassily W. Leontief, Input-Output Economics, Oxford University Press, Fairlawn, N.J., 1966.
6. Ben Noble, "Application of Matrices to Economic Models and Social Science Relationships," a lecture in Proceedings, Summer Conference for College Teachers on Applied Mathematics, University of Missouri - Rolla, 1971, published by C.U.P.M., Berkeley, 1973, pp. 111-117.

Among many elementary presentations of Leontief models the author's favorite is:

7. A.C. Chiang, Fundamental Methods of Mathematical Economics, McGraw Hill, New York, 1967.

The author wishes to thank Holden-Day, Inc., for permission to draw exercises and data from this source:

8. Andrei Rogers, Matrix Methods in Urban and Regional Analysis, Holden-Day, San Francisco, 1971, pp. 59-77.

I found the three magazine articles listed below to be particularly understandable and worthwhile. (There are many articles by and about Leontief in periodicals that almost all college libraries will have. Look up "Leontief" in the *Reader's Guide to Periodical Literature*.)

9. "Nobel Laureate Leontief," Paul Samuelson, Newsweek, Vol. 82, Nov. 5, 1973, p. 94.
10. "The 1973 Nobel Prize for Economic Science," Walter Isard and Phyllis Kaniss, Science, Vol. 182, Nov. 9, 1973, pp. 568-591.
11. "Input-Output Economics," Wassily W. Leontief, Scientific American, October, 1951.

7. Exercises: The Yugoslavian Economy in 1962 and 1958

7. In [8] page 69 ff., there is given an eight-"company" model of the Yugoslavian economy as of 1962. The data is reproduced by permission of Holden-Day, Inc. The closed sectors or "companies" are given in rows/columns numbered 1 through 8. A variety of open sectors are given in columns 10-14; use the total in column 16 to represent a single open sector. The input-output matrix A is given also. You will have to construct $\vec{d}$ as in Exercise 1.

   Your assignment, should you choose to accept it:

   a. Use a standard linear-equations solving program, already available for your computer, to find the production vector $\vec{x}$ for this model.

   b. Write a linear-equations solving program that, say, uses Gauss-elimination, to solve the equations $(I-A)\vec{x} = \vec{d}$ for this model. (This is a fairly large project.)

   c. Have the computer print out successive approximate solutions

ERIC Full Text Provided by ERIC

for this model, as required in problem 6. Convergence will not be immediate but will occur by about the twentieth round.

8. Consolidate the data in the tables used in Exercise 7 so that the production and consumption "industries" are
   1. "manufacturing," made up of old manufacturing (1) and construction (4);
   2. "agriculture," made up of the old agriculture (2) and forestry (3);
   3. services, made up of the old sectors (5), (6), (7), (8).

   The open sector is the subtotal row/column 16 used in Exercise 7. Repeat Exercise 7 for this consolidated model. Compare to the results of Exercise 7.

9. Comparable data (to that of 1962 used in Exercises 7 and 8) for 1958 appear on p.21. You should regard rows/columns 1-8 as the "companies" and subtotals in column 16 as the single open sector, as in Exercise 7.
   a. Calculate the appropriate matrix A and final demand vector $\vec{d}$.
   b. Solve the linear equations $(I-A)\vec{x} = \vec{d}$.
   c. Approximate $\vec{x}$ by using successive partial sums of the series for $(I-A)^{-1}$, as required in Exercise 5.

Tables for Exercises 7 and 8, reproduced from [8], by permission of Holden-Day, Inc.

Input Output Table for the Yugoslavian Economy 1962 (In Millions of Dinars)

Origin \ Destination	Manufacturing 1	Agriculture 2	Forestry 3	Construction 4	Transport and Communications 5	Trade 6	Services and Crafts 7	Others 8	Subtotal (1-8) 9
1 Manufacturing	1,848,873	81,37[illegible]	4,584	253,527	118,369	37,904	43,704	9,326	2,397,665
2 Agriculture	230,180	523,[illegible]	3,566		78	3,897	44		760,834
3 Forestry	79,122	[illegible]466	550	6,656	220	1,299	370	76	88,759
4 Construction	16,086	1,322	1,235	137,391	26,189	3,113	408	703	186,447
5 Transport and Communications	106,351	11,314	2,453	39,900	32,946	12,299	1,253	859	207,375
6 Trade	21,643	14,292	746	20,508	6,407	5,579	10,714	894	130,783
7 Services and Crafts	31,624	9,028	958	8,939	8,561	7,069	1,613	614	68,406
8 Others	39,256	237	130	2,100	1,063	2,849	378	277	46,290
9 Subtotal (1-8)	2,423,135	641,106	14,222	469,021	193,833	74,009	58,484	12,749	3,886,559
10 Depreciation	149,666	42,677	11,458	21,300	54,785	16,112	2,096	2,106	300,200
11 Personal Income	402,748	525,599	60,257	173,067	94,313	120,266	38,621	16,090	1,430,961
12 Accumulation (savings)	1,060,709	203,281	30,673	202,960	134,187	340,735	34,764	37,870	2,045,179
13 Subtotal (9-12)	4,036,258	1,412,663	116,610	866,348	477,118	551,122	133,965	68,815	7,662,899
14 Decrease in Stocks	11,983	37,598	2,090						51,671
15 Imports	668,171	146,133	1,122		20,247			9,588	845,261
16 Total (13-15)	4,716,412	1,596,394	119,822	866,348	497,365	551,122	133,965	78,403	8,559,831

Input Output Table for the Yugoslavian Economy 1962 (Continued)

Origin \ Destination	Increase in Stocks 10	Gross Investment 11	Exports 12	Consumption: Personal Consumption 13	Consumption: General Consumption 14	Consumption: Total Consumption (13-14) 15	Subtotal (10-14) 16	Total Output (9+16) 17
1 Manufacturing	178,952	463,628	552,553	960,817	162,802	1,123,619	2,318,747	4,716,412
2 Agriculture	6,918		72,637	731,979	24,026	756,005	835,560	1,596,394
3 Forestry		681	8,835	20,423	1,124	21,547	31,063	119,822
4 Construction		655,773	380		23,748	23,748	679,901	866,348
5 Transport and Communications	3,501	6,733	109,758	150,851	19,1[illegible]	169,998	289,990	497,365
6 Trade	6,106	39,485	51,571	306,870	[illegible]	323,177	420,339	551,122
7 Services and Crafts	1,204	2,645		50,684	11,026	61,710	65,559	133,965
8. Others	264		1,008	10,190	20,651	30,841	32,113	78,403
9 Subtotal (1-8)	196,945	1,168,940	796,742	2,231,814	278,831	2,510,645	4,673,272	8,559,831

Source: *Savezni Zavod za Statistiku* (1966) "Medusobni Odnosi Privrednih Delatnosti Jugoslavije u 1962 Godini" ("Interindustry Relations of the Yugoslav Economy in 1962") Beograd

$$A = \begin{bmatrix} 0.3920 & 0.0510 & 0.0383 & 0.2926 & 0.2380 & 0.0688 & 0.3262 & 0.1190 \\ 0.0488 & 0.3277 & 0.0298 & 0 & 0.0002 & 0.0071 & 0.0003 & 0 \\ 0.0168 & 0.0003 & 0.0046 & 0.0077 & 0.0004 & 0.0024 & 0.0028 & 0.0010 \\ 0.0034 & 0.0008 & 0.0103 & 0.1586 & 0.0527 & 0.0056 & 0.0030 & 0.0090 \\ 0.0225 & 0.0071 & 0.0205 & 0.0461 & 0.0662 & 0.0223 & 0.0094 & 0.0110 \\ 0.0152 & 0.0090 & 0.0062 & 0.0237 & 0.0129 & 0.0101 & 0.0800 & 0.0114 \\ 0.0067 & 0.0057 & 0.0080 & 0.0103 & 0.0172 & 0.0128 & 0.0120 & 0.0078 \\ 0.0083 & 0.0001 & 0.0011 & 0.0024 & 0.0021 & 0.0052 & 0.0028 & 0.0035 \end{bmatrix}$$

Programmed by Ervin Bell

Technical Coefficient Matrix for the Yugoslavian Economy 1962

Data for Exercise 9, reproduced from [8], pp. 73-74 by permission of Holden-Day, Inc.

The 1958 Yugoslavian Economy

Input-Output Table for the Yugoslavian Economy 1958 (in Millions of Dinars)

Origin \ Destination	Manufacturing 1	Agriculture 2	Forestry 3	Construction 4	Transport and Communications 5	Trade 6	Services and Crafts 7	Others 8	Subtotal (1-8) 9
1. Manufacturing	1,081,250	49,133	1,485	104,452	85,882	21,919	86,083	5,166	1,435,370
2. Agriculture	123,354	268,853	5,377		50	2,316	868	10	400,828
3. Forestry	50,030	1,019	15,462	4,222	195	1,227	1,320	6	73,481
4. Construction	4,922	199	299		10,437	73		498	16,428
5. Transport and Communications	47,216	3,654	1,393	15,685	14,843	8,208	2,635	301	93,935
6. Trade	33,143	7,680	79	10,845	1,158	397	9,738		63,040
7. Services and Crafts	12,992	7,102	88	56,352	2,829	4,912			84,275
8. Others	26,496			414	370	1,849	662	40	29,831
9. Subtotal (1-8)	1,379,403	337,640	24,183	191,970	115,764	40,901	101,306	6,021	2,197,188
10. Depreciation	73,362	20,789	4,971	7,933	34,500	7,249	4,391	1,249	151,444
11. Personal Income	156,018	388,407	45,332	55,435	41,358	54,695	56,380	8,446	806,071
12. Accumulation (savings)	630,132	77,250	28,967	53,622	50,085	139,197	45,267	22,676	1,046,196
13. Subtotal (9-12)	2,238,915	824,086	100,453	307,960	241,707	242,042	207,344	38,392	4,200,899
14. Decrease in Stocks	1,851								1,851
15. Imports	443,691	91,566	683	30	8,086			2,558	546,614
16. Total (13-15)	2,684,457	915,652	101 136	307,990	249 793	242,042	207,344	40,950	4,749,364

Input-Output Table for the Yugoslavian Economy 1958 (Continued)

Origin \ Destination	Increase in Stocks 10	Gross Investment 11	Exports 12	Consumption: Personal Consumption 13	Consumption: General Consumption 14	Consumption: Total Consumption (13-14) 15	Subtotal (10-14) 16	Total Output (9+16) 17
1. Manufacturing	119,406	281,797	276,257	438,903	132,724	571,62[illegible]	1,249,087	2,684,457
2. Agriculture	10,549	5,953	70,250	419,447	8,625	428,0[illegible]	514,824	915,652
3. Forestry		1,413	7,544	17,830	1,168	18,998	27,655	101,136
4. Construction		272,000	426		19,136	19,136	291,562	307,990
5. Transport and Communications	1,010	4,946	58,688	74,599	16,615	91,214	185,858	249,793
6. Trade	1,986	18,959	25,628	120,135	12,294	132,429	179,002	242,042
7. Services and Crafts	2,301		2,584	102,374	15,810	118,184	123,069	207,344
8. Others	480		242	4,159	6,238	10,397	11,119	40,950
9. Subtotal (1-8)	135,732	584,768	441,619	1,177,447	212,610	1,390,057	2,552,176	4,749,364

Source: *Savezni Zavod za Statistiku* (1962). "Medusobni Odnosi Privrednih Delatnosti Jugoslavije u 1958 Godini" (Interindustry Relations of the Yugoslav Economy in 1958"). Beograd.

8. ANSWERS TO SOME EXERCISES

1\. a. $(I-A)\vec{x} = \begin{pmatrix} .8 & -.2 \\ -.4 & .4 \end{pmatrix} \vec{x} = \begin{pmatrix} 10 \\ 4 \end{pmatrix}$ has solution

$$\vec{x} = \begin{pmatrix} 20 \\ 30 \end{pmatrix}.$$

b. $(I-A^{-1}) = \begin{pmatrix} 5/3 & 5/6 \\ 5/3 & 10/3 \end{pmatrix}$ precisely. The same solution $\vec{x} = \begin{pmatrix} 20 \\ 30 \end{pmatrix}$ results.

c. The "production totals" in the table give us $\vec{x}$, as we should expect from the definition of $\vec{x}$.

2\. a. $$A = \begin{pmatrix} 4/20 & 6/30 & 3/20 \\ 8/20 & 18/30 & 1/20 \\ 8/20 & 6/30 & 4/20 \end{pmatrix} = \begin{pmatrix} .2 & .2 & .15 \\ .4 & .6 & .05 \\ .4 & .2 & .2 \end{pmatrix}$$

and $\vec{d} = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix}$.

b. "Production totals" predict $\vec{x} = \begin{pmatrix} 20 \\ 30 \\ 20 \end{pmatrix}$.

c. $(I-A)\vec{x} = \begin{pmatrix} .8 & -.2 & -.15 \\ -.4 & .4 & -.05 \\ -.4 & -.2 & .8 \end{pmatrix} \vec{x} = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} = \vec{d}$

has the solution $\vec{x} = \begin{pmatrix} 20 \\ 30 \\ 20 \end{pmatrix}$ as expected.

The arithmetic is nasty if a methodical Gaussian approach is used, but easy if one tiptoes through the equations using a little foresight.

3\. b. About 50 iterations are needed to get noticeable convergence.

Results:

THE 5 TERM APPROXIMATION IS

$$\begin{pmatrix} 2.838399 & 1.455999 \\ 1.091999 & 2.110399 \end{pmatrix}$$

THE 10 TERM APPROXIMATION IS

3.979849	2.447314
1.835486	2.756191

THE 15 TERM APPROXIMATION IS

4.518542	2.915179
2.186384	3.060952

THE 20 TERM APPROXIMATION IS

4.772777	3.135986
2.351990	3.204784

THE 25 TERM APPROXIMATION IS

4.892762	3.240196
2.430147	3.272664

THE 30 TERM APPROXIMATION IS

4.949389	3.289377
2.467033	3.304700

THE 35 TERM APPROXIMATION IS

4.976114	3.312588
2.484441	3.319820

THE 40 TERM APPROXIMATION IS

4.988727	3.323542
2.492657	3.326955

THE 45 TERM APPROXIMATION IS

4.994679	3.328712
2.496534	3.330323

THE 50 TERM APPROXIMATION IS

4.997489	3.331152
2.498364	3.331912

5\. Reproduction of computer results are just below, giving the matrix sums and results after multiplication by $\vec{d}$:

THE 5 TERM APPROXIMATION IS

$$\begin{pmatrix} 1.473599 & 0.569599 \\ 1.139199 & 2.612799 \end{pmatrix}$$

AND LEADS TO OUTPUTS

$$\vec{x} = \begin{pmatrix} 17.014399 \\ 21.843199 \end{pmatrix}$$

THE 10 TERM APPROXIMATION IS

1.621937	0.772231	
1.544463	3.166400	

AND LEADS TO OUTPUTS
19.308298
28.110235

THE 15 TERM APPROXIMATION IS

1.656303	0.819177
1.638354	3.294658

AND LEADS TO OUTPUTS
19.839746
29.562180

THE 20 TERM APPROXIMATION IS

1.664265	0.830053
1.660107	3.324373

AND LEADS TO OUTPUTS
19.962872
29.898565

THE 25 TERM APPROXIMATION IS

1.666110	0.832573
1.665146	3.331257

AND LEADS TO OUTPUTS
19.991398
29.976499

THE 30 TERM APPROXIMATION IS

1.666537	0.833157
1.666314	3.332852

AND LEADS TO OUTPUTS
19.998006
29.994555

THE 35 TERM APPROXIMATION IS

1.666636	0.833292
1.666585	3.333221

AND LEADS TO OUTPUTS
19.999538
29.998738

THE 40 TERM APPROXIMATION IS

1.666659	0.833323
1.666647	3.333307

AND LEADS TO OUTPUTS
19.999892
29.999707

THE 45 TERM APPROXIMATION IS

1.666665	0.833331
1.666662	3.333327

AND LEADS TO OUTPUTS
19.999974
29.999931

THE 50 TERM APPROXIMATION IS

$$\begin{pmatrix} 1.666666 & 0.833332 \\ 1.666665 & 3.333331 \end{pmatrix}$$

AND LEADS TO OUTPUTS

$$\vec{x} = \begin{pmatrix} 19.999993 \\ 29.999983 \end{pmatrix}$$

ERIC Full Text Provided by ERIC

6. Reproduction of computer printouts of successive matrix approximations and the $\vec{x}$ they yield from multiplication by $\vec{d}$:

THE 5 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

$$\begin{pmatrix} 1.633799 & 0.690499 & 0.309799 \\ 1.266199 & 2.699099 & 0.323299 \\ 0.979199 & 0.833999 & 1.432899 \end{pmatrix} \qquad \vec{x} = \begin{pmatrix} 14.127699 \\ 17.607299 \\ 12.222199 \end{pmatrix}$$

THE 10 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

1.959441	1.084958	0.420515	17.811995
1.952976	3.532773	0.555972	25.381104
1.410627	1.356132	1.579797	17.102386

THE 15 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.080723	1.232059	0.461666	19.184608
2.209013	3.843307	0.642840	28.278700
1.571250	1.550941	1.634293	18.920162

THE 20 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.125927	1.286879	0.477001	19.696132
2.304429	3.959032	0.675213	29.358532
1.631108	1.623539	1.654602	10.597582

THE 25 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.142771	1.307308	0.482716	19.886759
2.339987	4.002158	0.687277	29.760947
1.653415	1.650594	1.662170	19.850033

THE 30 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.149048	1.314921	0.484845	19.957799
2.353239	4.018230	0.691773	29.910913
1.661728	1.660677	1.664991	19.944112

THE 35 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.151388	1.317759	0.485639	19.984272
2.358177	4.024219	0.693449	29.966800
1.664826	1.664434	1.666042	19.979172

THE 40 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.152259	1.318816	0.485935	19.994138
2.360017	4.026451	0.694073	29.987627
1.665980	1.665834	1.666433	19.992238

THE 45 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

2.152584	1.319210	0.486045	19.997815
2.360703	4.027283	0.694306	29.995388
1.666411	1.666356	1.666579	19.997107

THE 50 TERM APPROXIMATION IS AND LEADS TO OUTPUTS

$$\begin{pmatrix} 2.152705 & 1.319357 & 0.486086 \\ 2.360959 & 4.027593 & 0.694392 \\ 1.666571 & 1.666551 & 1.666634 \end{pmatrix} \qquad \vec{x} = \begin{pmatrix} 19.999185 \\ 29.998281 \\ 19.998921 \end{pmatrix}$$

7. Rogers, in [8], page 72, gives these results which I have not confirmed. Only $\vec{d}$ and five $\vec{x}$ vectors are given.

Iterative Solution of the Input-Output Model Yugoslavia, 1962

Industrial Sector	$\vec{d}$ Final Demand	$\vec{x}$ = Total Output after				Total Output 20th Round $\vec{x}$
		1st Round	2nd Round	3rd Round	10th Round	
1 Manufacturing	2,318,747	3,593,595	4,203,567	4,483,704	4,715,489	4,716,412
2 Agriculture	835,560	1,226,465	1,418,674	1,512,349	1,596,030	1,596,394
3 Forestry	31,063	76,903	100,006	110,812	119,786	119,822
4. Construction	679,901	814,776	848,359	859,071	866,321	866,348
5. Transport and Communications	289,990	409,701	458,548	479,915	497,296	497,365
6 Trade	420,339	492,930	524,941	539,248	551,074	551,122
7. Services and Crafts	65,559	104,519	120,686	127,933	133,941	133,965
8. Others	32,113	56,310	68,182	73,751	78,385	78,403
Total	4,673,272	6,775,199	7,742,963	8,186,783	8,558,322	8,559,831

Programmed by Ervin Bell

(Reproduced by permission of Holden-Day, Inc.)

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.

ERIC

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.)

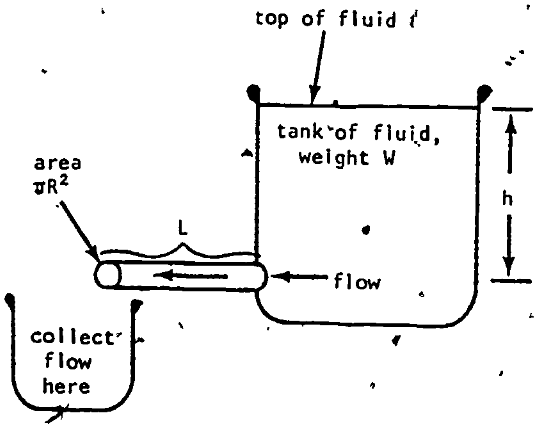
umap

UNIT 210

MODULES AND MONOGRAPHS IN UNDERGRADUATE
MATHEMATICS AND ITS APPLICATIONS PROJECT

VISCOUS FLUID FLOW AND THE INTEGRAL CALCULUS

by Philip Tuchinsky



APPLICATIONS OF CALCULUS TO ENGINEERING

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

ED 036 478

VISCOUS FLUID FLOW AND THE INTEGRAL CALCULUS

by

Philip Tuchinsky

TABLE OF CONTENTS

* This section may be omitted without affecting the readability of later sections.

Intermodular Description Sheet: UMAP Unit 210

Title: VISCOUS FLUID FLOW AND THE INTEGRAL CALCULUS

Author: Philip Tuchinsky
7623 Charlesworth
Dearborn Heights, MI 48127

Dr. Tuchinsky is a member of the Computer Science Department of Ford Motor Company's Research and Engineering Center. He formerly taught in the Mathematical Sciences Department at Ohio Wesleyan University (where earlier editions of this paper were written).

Review Stage/Date: III 9/1/78

Classification: APPL CALC/ENGINEERING

Suggested Support Materials: None are essential. A lab set up like that shown in Section 10 would make an interesting display. Exercise 4 calls for use of a computer or programmable calculator.

Approximate Class Time Needed: One 50 minute class.

Intended Audience: Calculus students learning how to integrate polynomials. The paper is suitable for independent reading and seminar presentation by more advanced students as well.

References: See Section 12 of the paper.

Prerequisite Skills:
1. Calculation of the integrals $\int x\,dx$ and $\int x^3\,dx$.
2. Knowledge that $\int c\,f(x)\,dx = c\int f(x)\,dx$.
3. Recognition of an integral as a limit of Riemann sums.
4. Comfort with summation results like $1 + 2 + 3 + \ldots + n = n(n+1)/2$.
5. Elementary computer programming (for Exercise 4 only).

Output Skills:
1. Replace a simple integral by a discrete sum, calculate both and compare results.
2. Average a function over an interval.
3. Reduce simple Riemann-Stieltjes integrals to Riemann integrals and calculate the latter (if the optional Section 7 is included).
4. Discuss how well Poiseuille's Law models a specified viscous fluid flow situation.
5. Describe a laboratory procedure for finding the coefficient of viscosity of a fluid.
6. Identify local vs. global information.

UMAP Editor for this module: Solomon Garfunkel

Other Related Units:
The Human Cough (forthcoming as UMAP Unit 211) Starts with the result of this paper that total flow is proportional to R^4 and goes on to discuss maximizing the speed of air flow during a cough. Differential calculus is its method.

© 1978 EDC/Project UMAP
All rights reserved.

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 Weitzel	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Dianne Lally	Project Secretary
Paula M. Santillo	Financial Assistant/Secretary

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
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 Melvin A. Nyman, Peter Signell and L.M. Larsen 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.

VISCOUS FLUID FLOW AND THE INTEGRAL CALCULUS

1. LAMINAR FLOW

When a thick, sticky (viscous) fluid flows through a pipe, it does not all flow at the same speed. Instead the fluid closest to the wall of the pipe suffers so much friction with the wall that it hardly moves at all, while fluid closer to the central axis of the pipe moves more rapidly. The fluid's speed increases steadily as the distance from the wall increases. Because of circular symmetry, the effect is that of concentric tubes of fluid sliding over one another (see Figure 1).

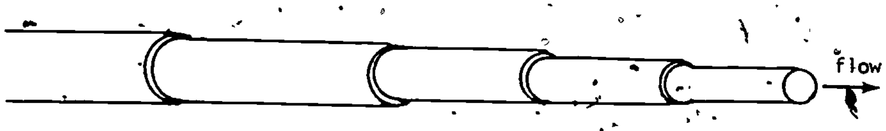


Figure 1. Laminar Flow in a Cylindrical Pipe.

We call this *laminar flow*: each lamina or layer of fluid moves at its own speed. Different laminae move at different speeds.

The exact way in which laminar flow happens was found by a French scientist named Poiseuille more than a century ago. He was studying blood pressure, which had just been accurately measured for the first time. He wanted to know how much blood flows through a blood vessel in a given time. From that information and analysis of blood samples one can say how much oxygen and nutrients are being delivered to the cells serviced by that blood vessel. Knowledge of blood flow is a basic part of understanding the body as a physical system.

Poiseuille's result about viscous fluid flow has many other applications. We can use it to study the flow of air in the windpipe, oil in a pipeline, water in a pipe system, grain flowing by pipe into the hold of a ship, etc. The assumptions involved in the result make it more applicable to some of these problems than others (see Section 3), but it provides a good first approximation to them all.

Another important use of Poiseuille's Law is to measure the relative viscosity of fluids. More about this later, in Section 10.

We will use Poiseuille's Law to calculate total flow through a pipe using a finite sum and the "continuous summation" process called integration. The two results will deserve comparison.

2. POISEUILLE'S LAW

Poiseuille discovered and others later deduced from theory (see Section 12) that the velocity of the particles of fluid at a distance r centimeters out from the center axis of the pipe is

$$(1) \qquad v(r) = \frac{P}{4kL}(R^2 - r^2) \qquad (\text{cm/sec.})^*$$

where (refer to Figure 2)

- R = radius of the pipe in cm. (Thus $0 \le r \le R$)
- L = length of the pipe (cm.)
- P = pressure change $P_1 - P_2$ down the length of the pipe, (dyne/cm²)
- k = coefficient of viscosity (poise)

---

* Variables will be given with their cm-gram-second (cgs) units to help us understand their physical meaning. Any system of units could be used, of course.

(Let me remind you that pressure is force per unit cross sectional area.) One can prove that the pressure will decrease steadily [as a straight line (linear) function] as the fluid moves through the pipe. It is the difference in final vs. initial pressure that enters the equation. The cgs unit of viscosity, the poise, is named after Poiseuille.

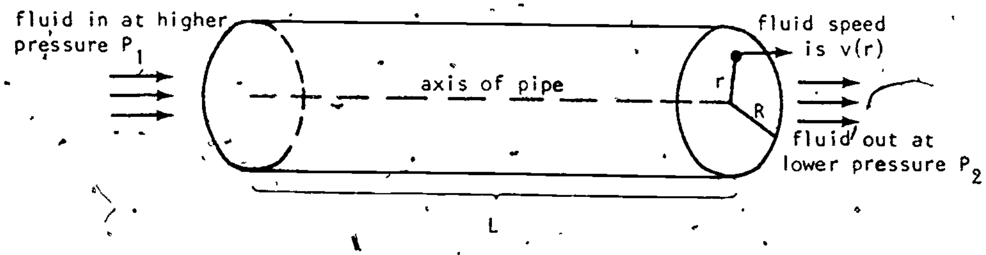


Figure 2.

3. WHEN DOES THIS LAW HOLD?

The major assumptions that must be true to have equation (1) valid are these:

a) There must be *no turbidity* in the fluid. This means that there is no swirling; particles of fluid move in straight lines down the pipe.

b) The speed of flow v is assumed to depend on r only. Thus v does not change as fluid moves down the length of the pipe; and it does not change with time; the flow is neither speeding up nor slowing down, it is *steady state*.

c) The fluid is *incompressable*, i.e., made up of particles that cannot be crushed or packed in closer together (by the forces present).

JEAN LEONARD MARIE POISEUILLE (1797-1869) was a well-known physiologist and physicist. He invented the mercury manometer to measure blood pressure, improving the pioneering work of Stephan Hales. The law considered here appeared in a paper of 1840 and was found through laboratory experiments with distilled water, ether and mercury. The mathematical derivation was first found in 1860 by F. Neumann and J. E. Hagenbach, who named the result Poiseuille's Law. But the name is disputed: G. H. L. Hagen found the same law independently in 1839; his work went unnoticed for decades. Reference: Dictionary of Scientific Biography, 1975 edition, vol.II, p. 62.

d) Fluid is *conserved*, i.e. neither created nor lost, in the pipe. Thus no fluid is leaking out through the pipe wall and no feeder-pipes are pouring fluid in or out.

e) The tube is horizontal and the (very slight) downward pulling effects of gravity are ignored. For a vertical tube this minor variation on (1) is true:

$$v(r) = \frac{P + g\rho L}{4kL}(R^2 - r^2)$$

where g = 980 cm/sec/sec is the gravitational constant and ρ is the density of the fluid, i.e., its mass per unit volume. For a slanted pipe, these horizontal and vertical velocities must be vector-added. For simplicity we will use (1).

f) The pipe is a right-circular cylinder with constant dimensions L and R.

g) There is so much friction at the wall that fluid there does not move at all. (Notice that $r = R$ leads to $v(R) = 0$.)

h) One assumption that is *not* present: in other classes you may study so-called "ideal fluids" in which particles slip frictionlessly by each other. We are assuming that each layer exerts a drag on the layer next-further-in. Ours is not an ideal fluid.

These assumptions are satisfied to various degrees by the applications mentioned earlier. Swirling, turbid effects are bound to occur in any large diameter pipe. This limits the usefulness of our law in studying water pipes, oil pipelines, grain chutes, etc. Blood vessels flex: their dimensions change a little. Blood surges because of the heart's pumping action; thus the flow is not steady-state. Oxygen and nutrients leave a blood vessel by osmosis through the pipe's wall and wastes are added to the blood flow, so that fluid is only approximately conserved.

Despite these and other practical short-comings, Poiseuille's Law is a valid simplification of viscous fluid flow. It is the right sort of law: $v(r)$ is zero at the pipe wall and increases steadily as r decreases and we approach the pipe's center. It has a solid, well-understood theoretical basis. We can really calculate with it, as we shall shortly see. And in the laboratory, the assumed conditions can be made almost true, giving a practical way to measure the viscosity coefficient k for any fluid. This coefficient is a fundamental property of the fluid, important in design and engineering work.

4. THE VELOCITY OF FLOW AND THE AMOUNT OF FLOW

We want to use Poiseuille's Law to calculate the total flow through a pipe of radius R. The *flow* F is the total volume of fluid passing through the pipe each second, in units of $(cm)^3/sec$.

First, we need a preliminary result. Consider, in Figure 3, any typical small piece of cross-sectional area of the pipe, consistiong of ΔA square centimeters, located r cm out from the center. How much fluid will leave the pipe through this bit of area *in one second*? The fluid moves $v(r)$ cm in the one second; thus a stack of fluid $v(r)$ cm long of constant cross section $\Delta A (cm)^2$ (shown) will flow out of the pipe through ΔA in the one second. This stack has volume $v(r) \cdot (\Delta A)$.

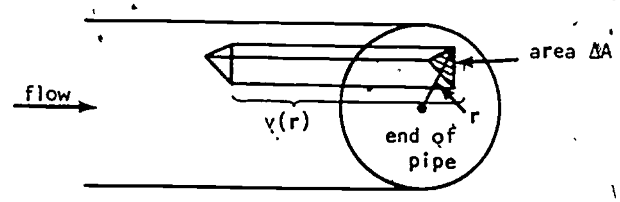


Figure 3.

Thus fluid leaves ΔA at a steady rate of $v(r) \cdot \Delta A (cm)^3/sec$.

(2) Summary: If ΔA is any area through which fluid flows at a *constant* velocity v, then $v \cdot \Delta A$ is the total flow through the area ΔA, per second.

5. THE TOTAL FLOW THROUGH A PIPE OF RADIUS R

In the pipe's cross-sectional circle of radius R, the velocity $v(r)$ given by Poiseuille's Law is the same at all points located r cm from the center. If we consider concentric rings of area (Figure 4), the fluid's velocity will be approximately constant in each ring. We can then use (2) to calculate the total flow through each ring; the sum of these ring-by-ring flows will be the total flow through the pipe, which we set out to find.

To clearly identify these rings, partition the interval

$$(0 \le r \le R)$$

into n pieces using partition points

$$0 = r_0 < r_1 < r_2 < \ldots < r_{n-1} < r_n = R$$

(perhaps not equally spaced).

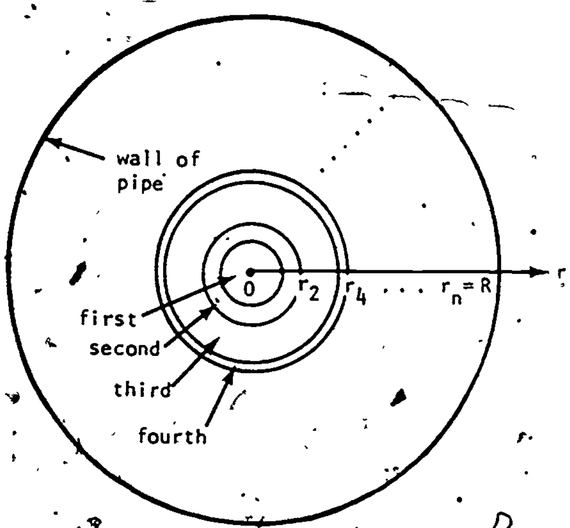


Figure 4.

The first, second, ... regions are then chosen as sketched. For $j = 1, 2, \ldots, n$, the j^{th} region is a ring with inner and outer radii r_{j-1} and r_j, and thus has area

$$\pi(r_j)^2 - \pi(r_{j-1})^2.$$

If we take n large and the r_j's close to each other, the velocity of fluid flowing through any one region will be almost constant, although different from ring to ring. What value will approximate the constant velocity in the j^{th} ring? Pick any point in that ring; say, pick a point that is t_j units out from the center, with $r_{j-1} \le t_j \le r_j$. Then $v(t_j)$ is a typical speed for the j^{th} ring and (2) says that

$$\text{the flow through the } j^{th} \text{ ring} \approx v(t_j)\cdot[\pi r_j^2 - \pi r_{j-1}^2]$$

We call t_j an *evaluation point* for the j^{th} subinterval $[r_{j-1}, r_j]$.

Thus the total flow through all n rings is

$$(3) \qquad F \approx \sum_{j=1}^{n} v(t_j)\,[\pi r_j^2 - \pi r_{j-1}^2].$$

We write "approximately" instead of equality because we have replaced all the various values of v(r) in the j^{th} ring by the single value $v(t_j)$. In fact, we have a vast family of approximations of F in Equation (3). For any choice of a partition $r_0, r_1, \ldots, r_n$ and any choice of evaluation points $t_1, t_2, \ldots, t_n$ (such that $r_{j-1} \le t_j \le r_j$ for each j) we get an approximation of F. As we take larger values of n and more closely spaced r_j's and t_j's, the theory of integration tells us that such sums approach a limiting value more and more closely, and that limit is an integral.

6. THE RIEMANN INTEGRAL

We must do a bit more work on Equation (3) before it is recognizable as a Riemann sum. Let the width of the j^{th} subinterval be

$$\Delta r_j = r_j - r_{j-1}$$

ERIC

Then

$$\pi r_j^2 - \pi r_{j-1}^2 = \pi(r_{j-1} + \Delta r_j)^2 - \pi r_{j-1}^2$$

$$= 2\pi r_{j-1}(\Delta r_j) + \pi(\Delta r_j)^2.$$

As n increases, r_j and r_{j-1} approach each other and Δr_j becomes small. The the $(\Delta r_j)^2$ term above is negligibly small by comparison to the first term, and becomes more negligible as n grows larger. Thus, from (3),

$$F \approx \sum_{j=1}^{n} v(t_j)[2\pi r_{j-1}\, \Delta r_j].$$

As $n \to \infty$, and all subinterval widths Δr_j shrink to zero, this Riemann sum becomes

$$F = \int_0^R v(r)(2\pi r)dr$$

$$= \int_0^R \frac{P}{4kL}(R^2 - r^2)\, 2\pi r\, dr = \frac{\pi R^4 P}{8kL}.$$

You are asked to calculate the integral in Exercise 1.

Another conversion of (3) into a Riemann sum: Since

$$\pi r_j^2 - \pi r_{j-1}^2 = \pi(r_j + r_{j-1})(r_j - r_{j-1}),$$

we have from (3)

(4) $$F \approx \sum_{j=1}^{n} v(t_j)\ \pi(r_j + r_{j-1})(r_j - r_{j-1}).$$

As $n \to \infty$, t_j, r_j, and r_{j-1} all approach each other and we get

$$F = \int_0^R v(r)\ \pi(r + r)\ dr$$

$$= \int_0^R v(r)(2\pi r)\ dr \quad \text{as before.}$$

7.* THE RIEMANN-STIELTJES INTEGRAL

The integral usually studied by calculus students is the Riemann integral,

$$\int_a^b f(x)\ dx.$$

An important generalization is the Riemann-Stieltjes integral where the "dx" representing change in x can be replaced by "d g(x)", the change in a function of x between one partition point and the next. That is, the Riemann sums and the limits they approach have the forms

$$\sum_{j=1}^{n} f(t_j)[x_j - x_{j-1}] \to \int_a^b f(x)\ dx$$

while the comparable Riemann-Stieltjes forms are

$$\sum_{j=1}^{n} f(t_j)[g(x_j) - g(x_{j-1})] \to \int_a^b f(x)\ d\,g(x).$$

In each case $a = x_0 < x_1 < \ldots < x_n = b$ is a partition of [a, b] and t_j is an evaluation point in the j^{th} subinterval: $x_{j-1} \le t_j \le x_j$.

We can now recognize (3) as a Riemann-Stieltjes sum with this integral as its limit

$$F = \int_0^R v(r)\ d(\pi r^2)$$

$$= \int_0^R \frac{P}{4kL}(R^2 - r^2)\ d(\pi r^2).$$

* This section can be omitted without affecting readability of later sections.

We can convert this integral to a Riemann integral by using this theorem:

> If f is continuous and g has a continuous first derivative for $a \le x \le b$, then
>
> $$\int_a^b f(x)\, d\, g(x) = \int_a^b f(x)\, g'(x)\, dx.$$

We get [since $g(r) = \pi r^2$ has derivative $g'(r) = 2\pi r$]

$$F = \int_0^R \frac{P}{4kL} (R^2 - r^2)\, 2\pi r\, dr,$$

the same Riemann integral as in Section 6.

Why should we be interested in the Riemann-Stieltjes integral if it simply leads us back to the Riemann integral we derived twice in Section 6? The Stieltjes case becomes interesting when g is not a smooth function, when $g'(x)$ does not exist. Then Riemann-Stieltjes theory must be used directly; we cannot escape to the easier Riemann case. There are important applications, especially in theoretical economics, where g must be taken as a step function, for example.

8. DISCRETE SUMMATION

Is it valid to let $n \to \infty$, taking rings of arbitrarily smaller and smaller width? That is, should we convert (3) into an integral? The fact that you are learning calculus is not sufficient to make the answer "yes"! In fact, we often should not take the limit. After all, blood is made up of red blood cells and other particles. They have a certain non-zero thickness Δr and no layer of blood can be thinner than that thickness. The same is true of all fluids, in fact.

To develop this idea, we should let all the rings have that fixed finite thickness Δr. Thus $r_0 = 0 = 0\cdot\Delta r$, $r_1 = 1\cdot\Delta r$, $r_2 = r_1 + \Delta r = 2\Delta r$, etc.; the $n+1$ partition points are $r_j = j\cdot\Delta r$, $j = 0, 1, 2, \ldots, n$. Let's simplify by taking the evaluation points to be $t_j = j\cdot\Delta r$ also. Then, from (4),

$$(5) \qquad F = \sum_{j=1}^{n} \frac{P}{4kL} (R^2 - (j\Delta r)^2)\pi[j\Delta r + (j-1)\Delta r][\Delta r]$$

$$= \frac{P\pi}{4kL} \sum_{j=1}^{n} (R^2 - j^2(\Delta r)^2)(2j - 1)(\Delta r)^2.$$

Plug in $R = n\cdot\Delta r$ and simplify to:

$$= \frac{P\pi(\Delta r)^4}{4kL} \sum_{j=1}^{n} (n^2 - j^2)(2j - 1)$$

$$= \frac{P\pi(\Delta r)^4}{4kL} [-2\Sigma(j^3) + \Sigma(j^2) + 2n^2 \Sigma(j) - n^2 \Sigma 1].$$

We can prove by mathematical induction that

$$\sum_{j=1}^{n} (j^3) = 1^3 + 2^3 + \ldots + n^3 = \frac{n^2(n+1)^2}{4}$$

$$\sum_{j=1}^{n} (j^2) = 1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{j=1}^{n} j = 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

$$\sum_{j=1}^{n} 1 = \underbrace{1 + 1 + \ldots + 1}_{n \text{ times}} = n.$$

Plug these in and do the algebra to reach

$$F = \frac{P\pi(\Delta r)^4}{8kL} n^2(n+1)(n-1)$$

(6)
$$= \frac{P\pi(n\cdot\Delta r)^4}{8kL} \cdot \frac{n+1}{n} \cdot \frac{n-1}{n}$$

$$= \frac{P\pi R^4}{8kL}\left[1 - \frac{1}{n^2}\right].$$

As $n \to \infty$, $\frac{1}{n^2} \to 0$ and this does approach the integral's value, as it should.

When we want to compute a sum, we often use the integral to approximate in a problem (like our current one) where $n \to \infty$ does not make sense. If n is in fact very large, only a small error is made. To do the actual sum for large n would be cumbersome; by letting $n \to \infty$ we gain all the calculational power of the integral calculus and save the algebra that led to (6).

There are other problems in which it is an integral we want but we are forced to use a sum. (Many integrals can't be calculated by anti-differentiation). By taking n sufficiently large, a high accuracy approximation of the integral can be gotten with the help of a computer.

Integration and discrete summation are associates. Each can help as a replacement for the other, in appropriate circumstances.

9. INTEGRATION: LOCAL DATA YIELDS GLOBAL RESULTS

Poiseuille's Law contains local information: the speed of fluid flow at a specific spot in the pipe is v(r). Our result (2) that $v\cdot\Delta A$ is the total flow through a bit of area ΔA where v is the (almost) constant speed of flow is still local information.

When we sum that local data over all parts of the pipe's cross-sectional circle, we gather the local data into a "global" result, referring to the pipe's total flow, to the pipe as an entity in itself. Integration[9] (or discrete summation, which is used less), converts locally varying information into the global. We are reasoning from the more detailed to the less detailed when we integrate.

Do we lose information through that process? Can we reason back to the local if we know the global result? You might immediately answer "no" or "sure, just differentiate." Can you justify either answer carefully? My question is

> Suppose we know that
>
> $$\int_0^R v(r)\, d(\pi r^2) = \frac{P\pi}{8kL} R^4 \text{ for any } R > 0.$$
>
> Can we deduce Poiseuille's Law, that
>
> $$v(r) = \frac{P\pi}{4kL}(R^2 - r^2)?$$

I leave it unanswered here.

10. CALCULATION OF VISCOSITY

To calculate k for a specific liquid, set up a tank and pipe in the laboratory as in Figure 5. Get a steady flow going, then collect (say) ten seconds flow in a beaker. Measure that volume of fluid.

According to our integration, in ten seconds the volume of fluid flowing out should be

$$10F = 10 \cdot \frac{\pi R^4 P}{8kL}.$$

In this equation we know every constant except k, which we calculate. We know R and L by measurement. To find P we take the difference between the pressures, P_1 and P_2, at the beginning and end of the flowpipe. The outlet pressure P_2 is simply atmospheric pressure.

ERIC

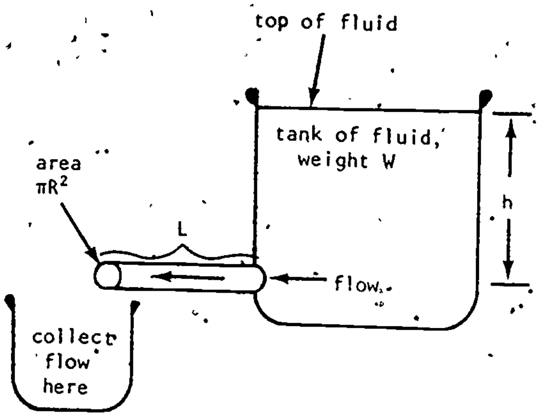


Figure 5.

If the fluid has weight density (weight per unit volume) ρ and the fluid depth is h as shown, the inlet pressure P_1 is $\rho g h$ where g is the gravitational constant.

11. EXERCISES

1. Show that $\int_0^R \frac{P}{4kL}(R^2 - r^2)\, 2\pi r\, dr = \frac{\pi R^4 P}{8kL}$.

   Notice that P, k, L and R are simply constants.

2. We have assumed that the fluid's velocity at the pipe wall is zero. There's no need to do that: The advanced derivation (see Section 12) that we have omitted in this paper in fact shows that the velocity is

   $$v(r) = \frac{-P}{4kL} r^2 + b$$

   where b is a constant we may choose.

   a) Show that v(R) = 0 leads to the formula (1) we have used.

   b) Suppose the velocity at the wall is one-half of the velocity at the center (r = 0). Find the function v(r) for this case.

   c) Use v(r) from (b) to find the total flow through the pipe of radius R.

3. The velocity v(r) varies from place to place in the pipe's cross-section, but has some average value $\bar{v}$.

   a) Explain how to find $\bar{v}$ from the total flow F and the principle in (2).

   b) The definition of the average value of the function v(r) is

   $$\bar{v} = \frac{\int_0^R v(r)\, 2\pi r\, dr}{\int_0^R 1 \cdot 2\pi r\, dr}.$$

   Calculate this and check against your work in (a). The two answers should agree.

   c) The largest velocity is V and occurs at r = 0. Check that $V = 2\bar{v}$.

4. a) Use a computer program to calculate the sum (5) for reasonable values of n, R, L, etc. Check the computer results against the algebraic result (6). Repeat with larger values of n.

   b) How large must n be to have the discrete sum within 1% of the integral result?

12. REFERENCE

If you know multivariable calculus and a little mathematical physics, you can read a clear derivation of Poiseuille's Law from basic ideas in elasticity and fluid flow:

Slater, J.C. and Frank N. H. Introduction to Theoretical Physics, McGraw-Hill, 1933. Or more recent books with similar titles.

13. SOLUTIONS OR HINTS TO EXERCISES

1. First convert to $\frac{2\pi PR^2}{4kL}\int_0^R r\,dr - \frac{2\pi P}{4kL}\int_0^R r^3\,dr$.

2. b) $v(r) = \frac{2PR^2 - Pr^2}{4kL}$.

   c) $\int_0^R \frac{2PR^2 - Pr^2}{4kL}\, 2\pi r\, dr = \frac{3\pi PR^4}{8kL}$.

3. a) If the fluid were moving at the same speed at all points in the cross-sectional circle of radius R, that constant speed would of course be the average of the Poiseuille's Law speeds: From (2), using $\Delta A = \pi R^2$, the full circular area,

$$\text{Total flow} = \bar{v}\cdot(\pi R^2) = \frac{\pi R^4 P}{8kL}$$

$$\Rightarrow \bar{v} = \frac{R^2 P}{8kL}.$$

   c) At $r = 0$, $v(0) = V = \frac{P}{4kL} R^2 = 2\bar{v}$.

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.

STUDENT FORM 2

Unit Questionnaire

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

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.

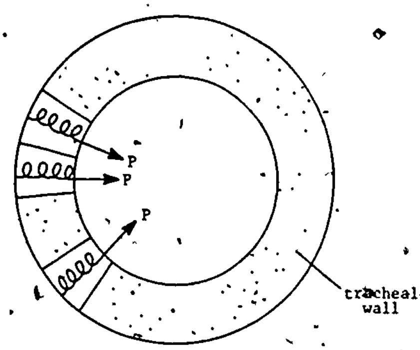
Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)

UMAP MODULE 211

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS

The Human Cough

by Philip Tuchinsky



Applications of Calculus to Physics, Biological and Medical Sciences

Birkhäuser Boston Inc.
380 Green Street
Cambridge, MA 02139

Intermodular Description Sheet: *UMAP Unit 211*

Title: THE HUMAN COUGH

Author: Philip Tuchinsky
7623 Charlesworth
Dearborn Heights, MI 48127

Dr. Tuchinsky is a computer scientist and mathematician at Ford Motor Company's Research and Engineering Center. He formerly taught in the Mathematical Sciences Department at Ohio Wesleyan University (where earlier editions of this paper were written).

Review Stage/Date: IV 7/30/80

Classification: APPL CALC/PHYSICS, BIO & MED SCI

Approximate Class Time: Less than one 50-minute class.

Intended Audience: Calculus students learning to use the derivative to compute extreme values of functions. The paper is suitable for independent reading or seminar presentation by more advanced students as well.

Prerequisite Skills:
1. Differentiation of polynomials.
2. Interpretation of dy/dx = 0 and the Second Derivative Test for identifying maxima and minima.
3. Operations on inequalities.
4. Basic curve sketching as taught in calculus.

Educational Objectives:
1. To see how a physical assumption may lead to a choice of domain for a function.
2. To see an application of maximization of a function on a closed interval domain.
3. To interrelate biology, physics, and calculus.

Related Units:
Viscous Fluid Flow and the Integral Calculus *(Unit 210)*

UMAP Editor for this Module: Solomon Garfunkel

The Project would like to thank L.M. Larsen of Kearney State College for his review, and all others who assisted in the production of this unit.

This material was field-tested and/or student reviewed in class by Simon Cohen of New Jersey Institute of Technology, and T.R. Hamlett of Arkansas Tech University, and has been revised on the basis of data received from these sites.

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 copyright holder.

© 1980 EDC/Project UMAP
All rights reserved.

ERIC Full Text Provided by ERIC

THE HUMAN COUGH

by

Philip Tuchinsky
7623 Charlesworth
Dearborn Heights, MI 48127

TABLE OF CONTENTS

ERIC Full Text Provided by ERIC

THE HUMAN COUGH

1. WHEN YOU COUGH . . .

When a foreign object in your trachea (windpipe) leads you to cough, your diaphragm thrusts sharply upward. As a result, the air in your lungs is suddenly compressed to a higher pressure than the air outside your body. A high-speed stream of air shoots upward through the trachea equalizing these pressures and, it is to be hoped, clearing the passage.

By Newton's law, the force exerted on the object to be cleared is due to the sudden acceleration of the air flowing through the trachea. The greater the velocity of the airstream during the cough, the greater the force on the foreigner and the more effective the cough. To increase the speed of the airflow, your body also contracts the windpipe during a cough, making a narrower channel for the air to flow through. For a given amount of air to escape in a fixed amount of time, it must move faster through a narrower channel than a wider one, just as a river flows rapidly where it is narrow but placidly where it is wide. In fact, x-rays show that the radius of the tracheal tube reduces to about two-thirds its usual radius during a cough.

2. NOTATION FOR A CALCULUS MODEL OF COUGHING

We can relate the speed of the airflow during a cough to the body's contraction of the trachea amazingly well by studying a simple mathematical model of the situation. We think of the trachea as a pipe with a circular cross section, and apply the differential calculus, using the following notation:

R_0 = the "rest radius" of the trachea (its usual radius when you are relaxed and not coughing) in centimeters;

R = the contracted radius of the trachea during a cough (thus $R < R_0$);

V = the average velocity of the air in the trachea when it is contracted to R cm. This depends on R and we wish to calculate R such that V(R) is maximal;

P = the extra pressure in the lungs during a cough, i.e., the difference $P_1 - P_2$ between the pressure P_1 in your lungs and the atmospheric pressure P_2 outside your mouth, measured in dyne/cm^2.

F = the total volume of air flowing through the trachea per second, in cm^3/sec.

We will make two physical assumptions, one about the airflow, the other about the flexibility of the trachea's wall.

3. LAMINAR FLOW

First, we assume that the airflow is *laminar*. This means that layers of air move at different speeds in the trachea. The thin layer of air right next to the pipe wall hardly moves at all because of friction with the wall. The layer, or lamina, just inside that one moves a little faster, and so on until the fastest airflow is found along the central axis of the trachea. It is as if the airstream were made of thin concentric tubes of air sliding over one another. See Figure 1.

Laminar flow is an appropriate model for the motion of any fluid through a confining pipe. In 1840, French physiologist Jean Poiseulle* established that the speed of the

*1797-1869. He was studying the flow of blood through veins and arteries.

ERIC Full Text Provided by ERIC

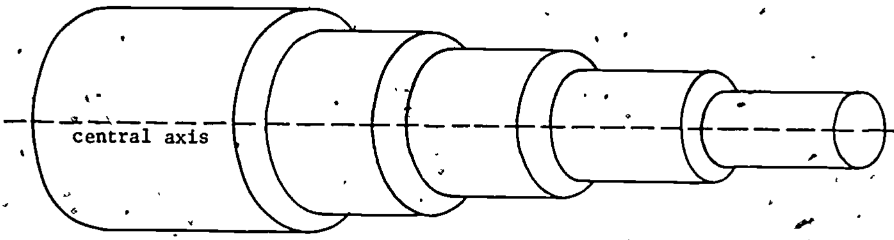


Figure 1. The air in the trachea is assumed to flow in thin concentric cylindrical layers called *laminae*. Inner layers move faster than outer ones, which are slowed by friction with the tracheal wall.

fluid (of air in the trachea in our case) at a point x cm out from the center axis of the pipe of radius R cm is

(1) $$v(x) = kP(R^2-x^2) \text{ cm/sec for } 0 \leq x \leq R.$$

Here k is a constant depending on the length of the pipe and the particular fluid involved. We defined P and R earlier. The average speed V is the average of these v(x) values over all points in the pipe.

Formula (1) is usually called *Poiseuille's Law* of viscous fluid flow. By using integral calculus, it is easy to deduce from (1) that the total flow per second through the trachea (when it is contracted to a radius of R cm) is

(2) $$F = cPR^4 \text{ cm}^3/\text{sec}.$$

The constant c again depends on the length of the pipe and the fluid involved. Formula (2) is derived from (1) in several ways in a companion paper to this one, *Viscous Fluid Flow and the Integral Calculus*, UMAP Unit 210. Laminar flow is discussed in more detail there, too.

4. AVERAGE VELOCITY AND TOTAL FLOW

We mentioned above that we could compute the average airspeed V in the trachea by using integral calculus to average the speeds v(x). However, we can relate V to the total flow per second F in a much simpler way.

Imagine air flowing through the trachea at a steady velocity of V cm/sec. In t seconds, each particle of air would travel Vt cm. Now, the cross-sectional area of the contracted tracheal tube is πR^2 cm^2. Therefore, a cylinder of air Vt cm long by πR^2 cm^2 would leave the tube during those t seconds. The flow of air through the tube, measured in *volume per second*, would be

$$(3) \qquad F = \frac{(Vt)(\pi R^2)}{t} = \pi R^2 V \quad cm^3/sec.$$

We can now write V in terms of P and the contracted radius R by using (2) and (3):

$$(4) \qquad V = \frac{F}{\pi R^2} = \frac{cPR^4}{\pi R^2} = c_1 PR^2,$$

where $c_1 = c/\pi$.

5. PERFECT ELASTICITY

The second assumption, about the flexibility or elasticity of the trachea's wall-tissue, is needed next. We assume that these tissues are "perfectly elastic." This means that the tissues contract so as to reduce the radius of the windpipe in proportion to the pressure-change P between the two ends of the pipe. That is,

$$(5) \qquad R_0 - R = aP,$$

for some constant a > 0. This is valid for fairly small pressure changes P, in fact for

ERIC Full Text Provided by ERIC

(6) $$0 \leq P \leq \frac{R_0}{2a}.$$

If larger values of P occur, the tracheal wall stiffens and the contracted radius R would be larger than the value predicted by (5). (This is fortunate—if the trachea were to contract too much, we would suffocate.)

Exercise 1. Use (5) to prove that the inequality

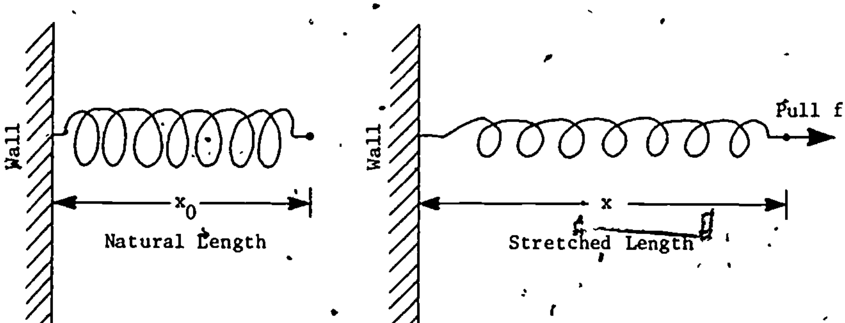
$$0 \leq P \leq \frac{R_0}{2a}$$

is equivalent to the inequality

$$\frac{R_0}{2} \leq R \leq R_0.$$

Thus, by assuming perfect elasticity, we are also assuming that the contracted radius R is at least 50 percent of the rest radius R_0.

You may be familiar with Hooke's Law, which says that the change $x-x_0$ in a spring's length when a pull, or force, of magnitude f is applied is proportional to f.



a. Unstretched b. Stretched

Figure 2. A spring stretched beyond its natural (unstressed) length by a force of magnitude f.

5

That is,

$$f = k(x-x_0),$$

for some constant k. This is really the principle behind perfect elasticity. The pressure change sucks in the tracheal wall with pressure P and the wall behaves as though it were made up of small springs, which stretch (Figure 3).

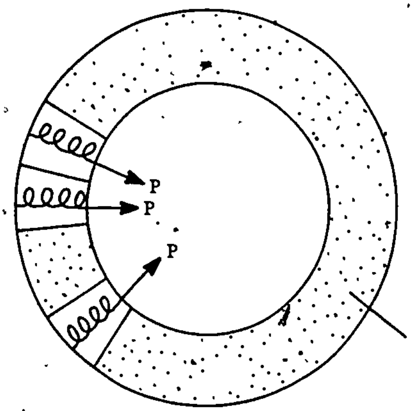


Figure 3. The tracheal wall is assumed to behave elastically as though it were made up of small springs which stretch as the trachea contracts.

As (5) says, the amount of stretch, R_0-R, is proportional to the magnitude of the force. Although this is a rather simplified explanation, it leads to a good working model, as you will see in the next section.

6. WHAT RADIUS R MAKES V THE LARGEST?

We can use (5), the formula for perfect elasticity, to express P in terms of R:

$$P = \frac{R_0 - R}{a}.$$

Inserting this in (4) gives us V in terms of R alone:

$$V = c_1\left(\frac{R_0 - R}{a}\right)R^2 = c_2(R_0 - R)R^2 \text{ cm/sec.} \qquad (7)$$

Here $c_2 = c_1/a$ and R_0 are constants. Equation (7) tells us that airspeed V is produced when the trachea contracts from R_0 to R cm.

Our original goal was to discover what value of R gives the largest value of V. Since V is a *differentiable* function of R, for R in the domain $[\frac{1}{2}R_0, R_0]$, V must assume its maximum at one of the endpoints $\frac{1}{2}R_0$ or R_0, or at an interior point where $dV/dR = 0$.

Exercise 2.

a. Show that $V = c_2(R_0-R)R^2$ satisfies $dV/dR = 0$ (has horizontal tangents) for $R = 2R_0/3$ and $R = 0$ but no other values.

b. Show that $R = 2R_0/3$ leads to $d^2V/dR^2 < 0$. Interpret this result: what sort of horizontal tangent is $R = 2R_0/3$?

c. Carefully explain how you know that V has its *absolute maximum* at $R = 2R_0/3$ when R is restricted to the domain $[\frac{1}{2}R_0, R_0]$.

As Exercise 2c shows, our model leads us to predict that our body can maximize the cough's effectiveness by contracting about 33 percent, from R_0 to $2/3R_0$. This agrees with experimental evidence as to how the body actually behaves! It is as though "Mother Nature" used calculus in designing the complex muscle-actions of coughing to maximize the airflow speed produced!

Exercise 3. Sketch the graph of $f(R) = (R_0-R)R^2$

a. for $0 \leq R \leq R_0$

b. for all real R.

Results from Exercise 2 will help, because V is just a constant multiple of the funtion f here.

7. ACKNOWLEDGEMENT

I first learned of this application from Alfred B. Willcox, Executive Director of the Mathematical Association of America, whom I wish to acknowledge and thank. Dr. Willcox presented most of this material under the title "Coughing with Calculus" as part of a talk at the Spring 1975 meeting of the Ohio Section of the MAA at Bowling Green State University.

8. SOLUTIONS TO EXERCISES

1. $0 \leq P \leq \frac{R_0}{2a} \iff 0 \leq aP \leq \frac{R_0}{2}$ (multiplication by a)

$$\iff 0 \leq R_0 - R \leq \frac{R_0}{2} \text{ (substitution from (5)).}$$

The left half, $0 \leq R_0 - R$, is equivalent to $R \leq R_0$ and the right half, $R_0 - R \leq \frac{R_0}{2}$, is equivalent to $\frac{R_0}{2} \leq R$. Together they give

$$\frac{R_0}{2} \leq R \leq R_0.$$

2. a. By the product rule (there are other ways)

$$\frac{dV}{dR} = c_2[(-1)R^2 + (R_0-R)2R] = c_2R(2R_0-3R).$$

b. $\frac{dV}{dR} = 0$ and $\frac{d^2V}{dR^2} < 0$ at a particular R indicates a *local* maximum.

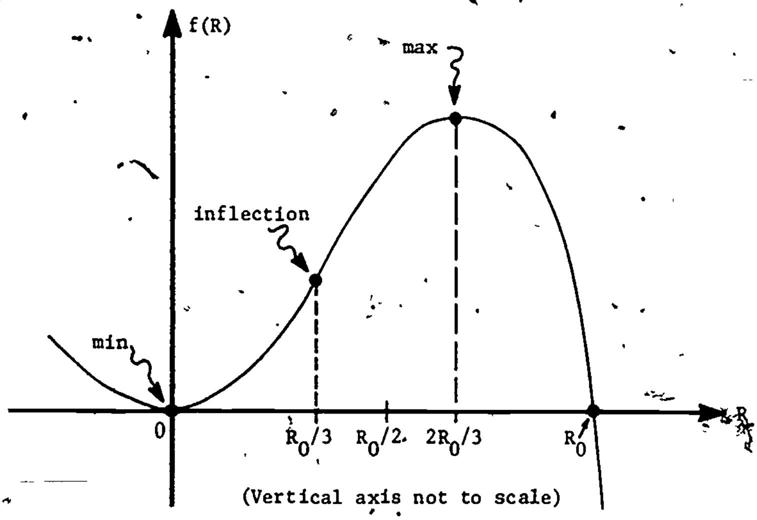
ERIC
Full Text Provided by ERIC

c. The absolute maximum needed here must occur at an endpoint of the domain or at an interior point where $dV/dR = 0$. Thus the candidates are

R	Corresponding Value of V	
endpoint $\frac{R_0}{2}$	$\frac{1}{8}c_2 R_0^3$	
endpoint R_0	0	
local maximum $2/3R_0$	$\frac{4}{27} c_2 R_0^3$	← the largest V

We ignore the horizontal tangent at $R = 0$ because it is outside the domain of our function.

3. The polynomial $f(R) = (R_0-R)R^2$ has a double root at $R = 0$, and a single root at $R = R_0$.



(Vertical axis not to scale)

SE 036 478

UMAP

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS

Birkhäuser Boston Inc.
380 Green Street
Cambridge, MA 02139

MODULE 215

Zipf's Law and His Efforts to Use Infinite Series in Linguistics

by Philip Tuchinsky

The terms of the series $\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1$ split up 1 in just about the way that the numbers of words appearing once, twice, thrice, etc., in James Joyce's <u>Ulysses</u> split up the total number of words in that novel!

Applications of Calculus to Social Science

ERIC Full Text Provided by ERIC

Intermodular Description Sheet: *UMAP Unit 215*

Title: ZIPF'S LAW AND HIS EFFORTS TO USE INFINITE SERIES IN LINGUISTICS

Author: Philip Tuchinsky
7623 Charlesworth
Dearborn Heights, MI 48127

Dr. Tuchinsky is a member of Engineering Computer Systems at Ford Motor Company's Research and Engineering Center. He formerly taught in the Mathematical Sciences Department at Ohio Wesleyan University (where earlier editions of this paper were written).

Review Stage/Date: IV 2/4/80

Classification: APPL CALC/SOCIAL SCIENCE

Suggested Support Material: Add one or more selections of English on which to do word-count experiments.

Approximate Class Time Needed: One 50 minute lecture plus out-of-class time for word-count experimentation and exercises.

Intended Audience: Calculus students studying series. By ignoring Exercises 4-7, the paper could be used at an intuitive level in pre-calculus or finite math or liberal arts mathematics courses. The unit is appropriate for independent study or seminar presentation by more advanced students.

Prerequisite Skills:
1. Definition of infinite series, and its sum.
2. Partial sums.
3. Geometric series summation.
4. Algebra on inequalities (for Exercises 1,2,3,5).
5. For Exercise 4 only: comparison, ratio and integral tests of convergence of series.
6. Algebra related to the logarithm function.
7. Log-log graph paper and its uses.
   (You can use this paper as a context in which to teach your students that $y=Ax^B$ will appear as a straight line on log-log paper, with A and B predictable geometrically or mathematically from the graph, and $y=A \cdot B^x$ will graph as a straight line on log-ordinary (semilog) paper. In my experience, many students are using these facts in science lab work without understanding why they work. They are delighted to have this enlightenment; their mistaken feeling that "none of this calculus is really useful for much" will be substantially reduced.)

Output Skills:
1. Use partial fractions to explain the summation of $\Sigma 1/k(k+1)$.
2. Calculate relative errors to measure quality of match-up between two sets of data.
3. Carry out a word-count study on any lengthy text in any language.
4. Convert item-count study data into rank-frequency data.
5. Use log-log paper to graphically test whether rank-frequency data obeys Zipf's Law.
6. Give an example of pure, apparently impractical research that has practical implications for a sophisticated system like human language.

UMAP Editor for this module: Solomon Garfunkel

© 1980 EDC/Project UMAP
All rights reserved.

ZIPF'S LAW AND HIS EFFORTS TO USE INFINITE SERIES IN LINGUISTICS

by

Philip Tuchinsky

TABLE OF CONTENTS

*This section is included in the instructional unit but omitted in the UMAP Journal version for the sake of brevity.

ERIC Full Text Provided by ERIC

ZIPF'S LAW AND HIS EFFORTS TO USE INFINITE SERIES IN LINGUISTICS

1. PARTIAL SUMS CAN HELP US ADD UP A SERIES

The partial sums of the series

$$\sum_{j=1}^{\infty} a_j$$

are, of course,

$$s_n = \sum_{j=1}^{n} a_j$$

for n = 1, 2, 3, ... The sum of the series is defined to be the limit of these partial sums as $n \to \infty$. Although that's a sound definition, it's almost useless when we want to calculate the sum of a series, because it is impossible to simplify the partial sums of most series into a form where the limit can be obtained. A classic exception to this rule is geometric series. The n-term partial sum $a + ar + ar^2 + \ldots + ar^{n-1}$ simplifies to

$$a\,\frac{1 - r^n}{1 - r}$$

(as you should be able to prove). In this simplified form, we can see what happens as $n \to \infty$: for r such that $|r| < 1$, we have $r^n \to 0$ and the series converges to

$$a\,\frac{1 - 0}{1 - r} = \frac{a}{1 - r},$$

while $|r| > 1 \Rightarrow r^n \to \pm\infty$ and the series diverges. (What happens when $r = \pm 1$?)

This paper is about another exception, another series whose partial sums can be directly analyzed.

ERIC Full Text Provided by ERIC

This series is not as important as geometric series (which has dozens of significant applications). However, our series played an interesting role in the linguistics research of George Kingsley Zipf in the 1920's and 1930's. We will examine that application and the later research about artificial languages that has made Zipf's work obsolete. A surprising interplay between the study of human languages and engineering research into communications networks and computer languages will be discussed.

We will see that the series we study is not completely successful as a mathematical model in Zipf's work. Several efforts to vary and improve the model will all lead to difficulties--no single accepted model will emerge. That sort of partial success is common when applied mathematicians work on actual complex problems; this deserves contrast against the experience of most students, who see one successful theorem proven after another as they study the established branches of mathematics.

2. SUMMING THE SERIES ZIPF USED

The series we consider here is

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots.$$

The key is to use *partial fractions*. Please check that

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

Now the partial sum through n terms is

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{(n-1)n} + \frac{1}{n(n+1)}$$

$$= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) + \left(\frac{1}{n} - \frac{1}{n+1}\right)$$

cancels cancels cancels cancels

$$= 1 - \frac{1}{n+1}.$$

ERIC Full Text Provided by ERIC

This partial sum is now so nicely simplified that we can see what happens as $n \to \infty$. Of course

$$\frac{1}{n+1} \to 0$$

and thus

$$\sum_{1}^{\infty} \frac{1}{k(k+1)} = \lim_{n\to\infty} \left(1 - \frac{1}{n+1}\right) = 1.$$

The original series adds up to 1.

3. WORD COUNTS IN JOYCE'S *ULYSSES*

This series gives a mathematical model of the occurrence of rare words in James Joyce's novel Ulysses. Among the 260,430 words in Ulysses there are N = 29,899 different words. Many are "rare" words appearing only once or twice. A few are common words that appear a thousand times or more. We'll study the rarely appearing words here. There are 16,432 words that appear exactly once each in Ulysses (about half of N), 4,776 words that appear exactly twice

$$(\text{about } \tfrac{1}{6} N = \tfrac{1}{2\cdot 3} N),$$

2,194 words that appear exactly 3 times each

$$(\text{about } \tfrac{1}{12} N = \tfrac{1}{3\cdot 4} N),$$

and so on.

In fact, if n_j is the number of words that appear exactly j times in Ulysses (j = 1, 2, 3, ...), these n_j words make up a fraction n_j/N (of the total N words) that is rather closely given by

$$\frac{1}{j(j+1)},$$

the j^{th} term of our series.

Thus we use the series to model n_j as

(1st model) $$n_j = \frac{N}{j(j+1)}.$$

This says that the terms of the series (which, you recall, add up to 1) split up 1 in just about the way that the words appearing once, twice, thrice, etc. in Ulysses split up the total of different words appearing in that novel.

4. HOW GOOD IS THIS SERIES MODEL?

The actual number of words appearing once, twice, ..., ten times in Ulysses is listed in Table 1 along with the number predicted by the series-model.

TABLE 1

j	n_j = actual # of words appearing exactly j times	Model's prediction of number of words that appear exactly j times	Relative error
1	16,432	14,950	9.0%
2	4,776	4,983	4.4%
3	2,194	2,492	13.6%
4	1,285	1,495	16.3%
5	906	997	10.0%
6	637	712	11.7%
7	483	534	10.6%
8	371	415	11.9%
9	298	332	11.5%
10	222	272	22.5%

Source: Zipf, Human Behavior and the Principle of Least Effort.

The last column provides a simple measurement of the extent to which predicted and actual values agree. The *relative error* is defined to be

$$RE = \left|\frac{\text{predicted value - actual value}}{\text{actual value}}\right| .$$

As an example, for j = 7, the RE is

$$\frac{|534 - 483|}{483} = \frac{51}{483} = .10559 = 10.6\%.$$

The predicted values are obtained from our series model as in this example: for j = 3, the model predicts that

$$n_3 = \frac{N}{3 \cdot 4} = \frac{29899}{12} = 2491.58,$$

which we round to 2,492.

The predicted values given by the terms of our series do follow the trend of the actual data quite well, but you may feel that the specific numbers (483 vs. 534 for j = 7, for example) are not as close as you might prefer. Shouldn't the model match reality better than that? The RE's in the last column average 12.5%. For most research in the natural sciences such relative errors would be considered large--repeated experiments done with laboratory equipment, for example, usually yield much more consistent results. Errors above even 5% make us wonder about the experimenter's measuring abilities or the design of the experiment. But we should not expect such hard-science accuracy in a "law" or model that concerns so complicated a social-science process as the choice of words by one human in creating one novel. Instead, we ask: Is this pattern obeyed by a wide range of language samples?

5. THE EXTENSIVE RESEARCH INTO WORD-COUNTS AND RELATED LANGUAGE PATTERNS

During the 1920's and 1930's, many word-count experiments were performed by psychologists and linguists, led by Professor George Kingsley Zipf of Harvard and his students. They found striking patterns in the frequency of occurrence of: rarely appearing words, the number of pages between appearances of a word, the number of and spacing between uses of individual letters, syllables, prefixes, suffixes, meanings, etc. Some of the language texts studied (not all for rare-word frequencies) were:

--- Ulysses by Joyce
--- Stretches of English language newspaper text
--- the plays of Plautus in Latin
--- the Iliad in Homeric Greek
--- works in Old English, and other medieval languages
--- part of a Bible in Gothic German
--- traditional oral legends in Dakota and Plains Cree (American Indian languages) and Nootka (an Eskimo language)

--- works in modern languages from German to Hebrew to Chinese
--- the speech of children at various ages
--- some schizophrenic speech.

This exceptionally broad selection of language samples all yielded very regular patterns that astonished the researchers. A few studies failed to support the patterns* but the evidence suggested that important cross-cultural properties of language were being found.

Linguists pursued this research in the search for fundamental structural properties of language. Psychologists hoped to explain just what process goes on in a human mind as it calls on its whole history of language experiences when crafting new sentences, paragraphs, or books. One of Zipf's books (see Section 10) contains a readable survey of these experiments. It also contains the extensive consequences for human behavior that Zipf put forward as implications of the research. A too-brief review of his logic: Zipf claimed that different amounts of mental effort are exerted by a speaker or writer in choosing words. Common words, very frequently encountered in the writer's past experiences, "come to mind" with little effort while words met less often in the past require more effort for their use. A human selects words to express an idea using the "principle of least effort." Zipf hoped to derive the specific quantitative patterns he had found from such a basic principle (in the same way that Newton, starting from a few basic assumptions such as the law of gravity, could derive the motion of the planets and many other results). Zipf offered situations analogous to writing or language usage where behavior obeying a law of least effort did lead to the patterns found, but he did not succeed in deriving the surprising patterns from language structure itself.

* One of the exceptions is another novel by James Joyce, Finnegan's Wake.

ERIC Full Text Provided by ERIC

6. ZIPF'S LAW (THE RANK-FREQUENCY LAW)

A central result of this research is "Zipf's Law" also called the "rank-frequency law." We have looked at the *number* n_j of rarely appearing words that appear with frequency (number of occurrences) j for j = 1, 2, 3, ... In a rank-frequency study, one looks instead at the *rank* of a word (1^{st}, 2^{nd}, 3^{rd}, etc.) when the words of a book are listed in order of decreasing frequency. Thus the most-repeated word has rank 1 and frequency f_1, the second-most-repeated word has rank 2 and appears f_2 times, and so on. Zipf's Law, also found empirically, is that

$$r \cdot f = \text{constant}$$

i.e., that the rank and corresponding frequency are inversely related. As an example, Table 2 gives various ranks, frequencies, and r·f products for Ulysses.

TABLE 2

Actual Rank-Frequency Data from Ulysses

Rank (r)	Frequency (f)	r·f products
10	2,653	26,530
20	1,311	26,220
30	926	27,780
40	717	28,680
50	556	27,800
100	265	26,500
200	133	26,600
300	84	25,200
400	62	24,800
500	50	25,000
1,000	26	26,000
2,000	12	24,000
3,000	8	24,000
4,000	6	24,000
5,000	5	25,000
10,000	2	20,000
20,000	1	20,000
29,899	1	29,899

Source: Zipf, Human Behavior and the Principle of Least Effort.

The approximate constancy of this third column is striking and intuitively unexpected. And the constant value obtained is roughly N = 29,899, the number of distinct words being ranked, or perhaps it is a bit less than N. This is discussed in comments following Exercise 1 in Section 8.*

7. A LOG-LOG GRAPH REVEALS OBEDIENCE TO ZIPF'S LAW

There is an easy way to graph the (r, f) pairs from Ulysses for r = 1, 2, 3, ..., 29,899 so that the closeness of fit to $r \cdot f = k$ becomes visible. On ordinary graph paper, $r \cdot f = k$ appears as a hyperbola; it is hard to look at the graph and determine that we have $f = k/r$ as opposed to some other similar curve, like $f = k/r^2$ or $f = k/r^{1.2}$. But these curves are easy to tell apart when graphed on log-log graph paper. Notice that $r \cdot f = k$ implies $\log r + \log f = \log k$. Thus the points (r, f) fall on the curve $r \cdot f = k$ if and only if the points $(x, y) = (\log r, \log f)$ fall on the *straight line with slope -1* $x + y = \log k$. On log-log graph paper (see Figure 1), the axes are labeled with values of r and f but, because of the special spacing of points along these axes, we are really plotting $y = \log f$ vs. $x = \log r$. We will have a good fit to $r \cdot f = k$ if the data fall along a straight line with slope -1, cutting both axes at 45°.

In Figure 1, the tendency of both curves A and B to follow the straight line C is very striking. (The "steps" at the bottom-right of both curves occur because, for high

*Zipf's Law $r \cdot f = k$ appears to fit many kinds of ranked data beyond our word counts. For example, when U.S. cities are ranked by population (so that r = 1 for New York, etc.) then $r \cdot f = k$ holds pretty well, where $f = f_r$ is the population of the city with rank r. The rule fails for cities world-wide, or for cities in much less urbanized, industrialized societies, and the extent of fit to this law has been proposed as a measure of a nation's urbanization. Consult the social science literature for more details and other examples.

ranks there are many ties, many occurrences of the rare frequencies 1, 2, 3, ...)

Researchers up to this point had not explained Zipf's Law, or the series model that we began with in this paper or other patterns.

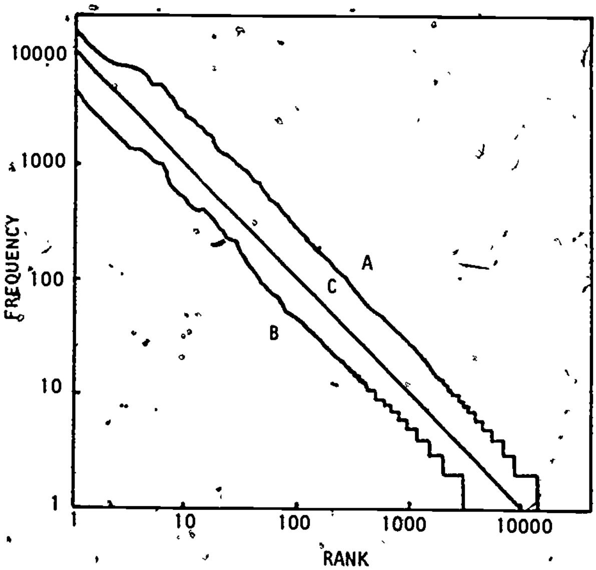


Figure 1.- Data that precisely obeys Zipf's Law would graph like C, having slope -1, to which curves A and B should be compared. Curve A consists of all the (r,f) data pairs for Ulysses, not just the few given in Table 2, connected together into a curve. Curve B is a similar rank-frequency graph for a sample of 43,989 running words of American newspaper text, studied by R.C. Eldridge. (The Ulysses data was created by Hanley and Joos, but first graphed by Zipf. Source: Zipf, Human Behavior and the Principle of Least Effort.)

8. EXERCISES: DERIVING A NUMBER-OF-WORDS LAW

We have studied two parts of Zipf's research, which we summarize as follows:

(A) The rank-frequency law $f_r = k/r$ gives the approximate frequency (number of appearances) f_r of the r^{th}-most-commonly-appearing word in the language sample, for $r = 1, 2, 3, \ldots, N$.

(B) The number-of-words law $n_j = N/j(j+1)$ tells how many words (among the N different words of the language sample) appear exactly j times, for $j = 1, 2, 3, \ldots$.

Both are empirical laws--they work quite well for a wide variety of language samples. So far we have no derivation of these laws from obvious or widely accepted facts, no clear explanation as to why they should be true.

These two laws are related to each other and that is worth our study--if one follows from the other, they are more believable together than either is by itself.

Therefore, let's assume that (A) is true and try to deduce a number-of-words law from it. Specifically, let's try to calculate n_1, the number of words that appear exactly once (i.e., that have $f = 1$).

The rank-frequency law predicts frequencies f_r between 1 and 2 for all words with ranks $k/2 + 1$ up to k:

$$1 \le f < 2 \Leftrightarrow 1 \le \frac{k}{r} < 2$$

$$\Leftrightarrow \frac{k}{2} < r \le k.$$

Thus, a total of $k/2$ words have theoretical frequencies f in the interval $[1, 2)$.

However, frequencies must be integers; fractional frequencies do not make sense. Let's decide that we will always round f downward to the next lower integer. Then

ERIC Full Text Provided by ERIC

$f \in [1, 2)$ becomes $f = 1$, and n_1, the number of words with $f = 1$, is

$$n_1 = \frac{k}{2} = \frac{k}{1 \cdot 2} .$$

This looks promising--if we are going to derive $n_j = N/j(j+1)$ from (A), we need that denominator $1 \cdot 2$ in n_1. But that k in the numerator? Maybe the correct constant k in the rank-frequency law is N? We'll have to test that idea later. First, extend our result for n_1 by doing Exercise 1.

Exercise 1

Assume that $f = k/r$ for $r = 1, 2, 3, \ldots$.

a) Show that $f \in [j, j+1)$ occurs exactly for ranks

$$r \in \left(\frac{k}{j+1}, \frac{k}{j}\right].$$

b) If we round $f \in [j, j+1)$ downward to the integer value $f = j$, show that

$$n_j = \frac{k}{j(j+1)}$$

for any j.

Thus we can deduce (B) from (A) if we agree to round f downwards and if $k = N$.

We should test whether $k = N$ empirically by trying it on many language samples. We can start here with Ulysses, which contains $N = 29,899$ different words. The r and f data in Table 2 can be used to get a comparable value of k. Let's exclude the data for $r = 10,000$, $20,000$, and $29,899$ because these (r, f) pairs are located in the "steps" of the (r, f) graph where r changes while f does not and those $r \cdot f$ products are not very constant. When we average the $r \cdot f$ products in Table 2 for $10 \leq r \leq 5,000$, we get $k = 25,874$. Thus $k \neq N$. We have k about 13.5% smaller than N in this one example.

Wait. This is no time to quit on the problem--the values $N/j(j+1)$ are also 10-15% too large for the actual n_j of *Ulysses* in Table 1 (except for j = 1, 2). We could correct that by decreasing $N/j(j+1)$ to $k/j(j+1)$. Thus we propose

$$(2^{nd}\ \text{model}) \quad n_j = \frac{k}{j(j+1)}$$

for all but the smallest j. We cannot apply this model for all j because

$$N = \sum_j n_j \ ;$$

however,

$$\sum_j \frac{k}{j(j+1)} = k \sum_j \frac{1}{j(j+1)} = k.$$

But the 2nd model may work well for all but the smallest few values of j, which are special cases requiring their own formula. *Ulysses* data comparable to that in Table 1 appears in Table 3. We must be cautious in concluding that the 2nd model will do this well for j > 10 or for language samples other than *Ulysses*. The second model does not seem to appear in the psychological literature, probably because Zipf deduced yet another number-of-words formula from the rank-frequency law.

TABLE 3

Additional Number-of-Words Predictions vs. *Ulysses* Data.

		2nd model*		3rd model*	
j	true n_j	predicted n_j	RE	predicted n_j	RE
1	16,432	12,937	21.3%	34,499	109.9%
2	4,776	4,312	9.7%	6,900	44.4%
3	2,194	2,156	1.7%	2,957	34.8%
4	1,285	1,294	1.0%	1,643	27.9%
5	906	862	4.9%	1,045	15.3%
6	637	616	3.3%	724	13.7%
7	483	462	4.3%	531	9.9%
8	371	359	3.2%	406	9.4%
9	298	287	3.7%	320	7.4%
10	222	235	5.9%	259	16.7%

*All calculations are based on k = 25,874.

ERIC Full Text Provided by ERIC

Surely you wanted to object to the "rounding" of $f \in [j, j+1)$ to $f = j$! After all, would you round 3.01, 3.3, 3.49, 3.51, 3.99 all to 3? It would also mean that $f \in [0, 1)$, which is predicted by $f = k/r$ for ranks $k < r \leq N$, is rounded to $f = 0$, although each of the words with these ranks appears in the language sample at least once. Zipf proposed instead to round $f \in [1/2, 3/2)$ to $f = 1$, $f \in [3/2, 5/2)$ to $f = 2$, etc.

Exercise 2.

Assume that $f = k/r$ for $r = 1, 2, 3, \ldots$

a) Show that $f \in [j - 1/2, j + 1/2)$ occurs exactly for words with ranks

$$\frac{k}{j + 1/2} < r \leq \frac{k}{j - 1/2}.$$

b) If we round $f \in [j - 1/2, j + 1/2)$ to $f = j$, show that

(3rd model) $$n_j = \frac{k}{(j - 1/2)(j + 1/2)} \text{ for any } j.$$

This third model is the one given by Zipf. It leads us to ask:

Exercise 3.

$$\sum_{j=1}^{\infty} n_j$$

should equal N, the total of different words in the book under study. Sum the series suggested in Exercise 2, formula,

$$\sum_{j=1}^{\infty} \frac{1}{(j - 1/2)(j + 1/2)}.$$

by simplifying the partial sums in much the way $\Sigma\, 1/j(j+1)$ was summed early in this paper.

Since Exercise 3 tells us that $\Sigma n_j = 2k \neq N$, we know we cannot use the 3rd model for all j, based on k and N from Ulysses. As with the 2nd model, for low j the predicted values are far too large. Table 3 shows a very poor fit

ERIC Full Text Provided by ERIC

between this model and the *Ulysses* data; for much larger values of j the fit may be much better.

So it goes! In three tries, we have not achieved a trouble-free model.

Exercise 4.

Without finding the sum, give more than one proof that

$$\sum_{j=1}^{\infty} \frac{1}{(j - 1/2)(j + 1/2)}$$

is a convenient series. Mention the convergence tests you use.

Exercise 5.

Suppose we decide to round upward: Assume $r \cdot f = k$ and decide to replace $f \in (j-1, j)$ by $f = j$. What rule for n_j follows? Is it a better model than the ones we have discussed? Prepare the equivalent of Table 1 for this 4^{th} model. How did you decide whether or not it is better than the first 3?

The series result

$$\sum_{1}^{\infty} \frac{1}{k(k+1)} = 1$$

can be used to find the sums of other series. Two examples appear as Exercises 6 and 7.*

Exercise 6.

First show that

$$\sum_{1}^{\infty} \frac{1}{k(k+1)} = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \frac{1}{4\cdot 5} + \ldots$$

$$= \frac{1}{2}\left(\frac{1}{1} + \frac{1}{3}\right) + \frac{1}{4}\left(\frac{1}{3} + \frac{1}{5}\right) + \ldots$$

$$= \sum_{n=1}^{\infty} \frac{2}{(2n-1)(2n+1)}.$$

Use this result to show

$$\frac{1}{1\cdot 3} + \frac{1}{3\cdot 5} + \frac{1}{5\cdot 7} + \ldots = \sum_{n=1}^{\infty} \frac{1}{(2n-1)(2n+1)} = \frac{1}{2}.$$

---

*Thanks go to William Glessner of Central Washington University for suggesting Exercises 6 and 7.

Exercise 7.

If we start with the result in Exercise 6:

$$\frac{1}{2} = \frac{1}{1\cdot 3} + \frac{1}{3\cdot 5} + \frac{1}{5\cdot 7} + \ldots$$

and use the same "sum up two terms at a time" method (as displayed in Exercise 6) on *it*, show that we get

$$\frac{1}{4} = \sum_{0}^{\infty} \frac{1}{(4n + 1)(4n + 5)} = \frac{1}{1\cdot 5} + \frac{1}{5\cdot 9} + \frac{1}{9\cdot 13} + \ldots.$$

9. MANDELBROT'S EXPLANATION OF THE LANGUAGE PATTERNS

Zipf's Law and other striking patterns found through word-count sorts of experiments on natural (i.e., human) languages were finally explained by scientists working on very different problems, problems related to artificial languages. Zipf and his colleagues had examined the structure of language and the process of writing or speaking; now Norbert Weiner and Claude Shannon led the study of communications channels. Human speech and writings, electronic signals sent over telephone lines, messages sent in Morse code, radar signals sent out and received after bouncing back, coded data moving from IBM cards into a computer's electronic memory, all are examples of information being coded and sent by a transmitter (speaker, writer, telegraph key user, etc.) then received, decoded, and interpreted by a receiver (listener, reader, etc.). The researchers asked: How could information be most efficiently coded and sent so that it would be received at lowest cost and with high accuracy? How much repetition ("redundancy") should be included as a check on the accuracy of the message received? Their main goals were the efficient design of high speed, high volume, high accuracy man-made data channels for use in computers, international telephone and microwave systems and military applications, but the linguists and psychologists noticed at once that this research was relevant to the study of human language communications, too.

ERIC Full Text Provided by ERIC

This research led to an anticlimatic completion of the project begun by Zipf and his team. In 1953-54, Benoît Mandelbrot showed that the number-frequency, rank-frequency and other patterns found by Zipf will always arise in any language satisfying these two assumptions:

1. The language is made up of words--small units of information separated by spaces.
2. The transmitter encodes and the receiver decodes word by word--that is, the speaker (or writer) formulates and speaks one word at a time and the listener (reader) listens and interprets one word at a time.

The main point is the presence of a space between units of information. By random processes this spacing, and the word by word handling of messages, accounts for the patterns. There is no need, in explaining the patterns, to claim that James Joyce, while writing Ulysses, was choosing words using unknown "universal laws" of language structure at some deep almost-unconscious level of thought. Instead, we simply claim that Joyce was choosing his words one at a time to convey his meanings. The space-between-words structure of English then suffices to produce the patterns. Mandelbrot showed this by using a lot of advanced mathematical statistics.

Zipf's ideas persisted for a while. The applicability of Shannon's work to human languages was challenged and some of Mandelbrot's assumptions were questioned, by H.S. Simon and others. Simon, in 1955, published alternative explanations of Zipf's Law and other patterns, using the idea that the more prior usage a word has had, the more likely it is to recur.

Mandlebrot has won the day, however. My most recent reference, in Mathematics and Psychology, edited by George A. Miller, John Wiley and Sons, New York, 1964, includes this quote from Bärbel Inhelder and Jean Piaget on page 249:

> ... during the 1930's G.K. Zipf stirred up considerable interest in various statistical regularities that he uncovered in his analysis of word frequencies. Twenty years later the mathematician Benoît Mandelbrot was able to demonstrate that Zipf's laws were attributable to random processes and implied no deep linguistic or psychological consequences.

10. SOURCES

I first met this application in the essay "The Sizes of Things" by H.A. Simon in Statistics: A Guide to the Unknown ed. Judith Tanur, Holden-Day, 1972, pp 195-202. This paperback contains many short essays that show the applicability and practical uses of statistics, especially the difficulties of statistical experiment design. Most are only modestly mathematical.

The work of Zipf and his colleagues is well summarized in G.K. Zipf, Human Behavior and the Principle of Least Effort, Addison-Wesley, Cambridge, Mass., 1949, Chapters 2, 3, and 4.

The original Ulysses data, complete, appears in M.L. Hanley et al, Word Index to James Joyce's *Ulysses*, Madison, Wisconsin, 1937.

The mathematics used by Zipf to relate his rank-frequency law to the number-frequency law for rare words, presented in Exercise 2, was presented in G.K. Zipf, "Homogeneity and heterogeneity in language", Psychological Record, 2 (1938), pp. 347-367. A more general argument by Martin Joos appears in a book review of Zipf's The Psycho-Biology of Language, Houghton Mifflin, Boston, 1935 in Language, 12 (1936) pp 196-210. Joos, while contributing to Zipf's rigor, is not uncritical.

A good summary of Mandelbrot's results and their meaning may be found on pp 60-69 of R.D. Luce, ed.,

ERIC Full Text Provided by ERIC

Developments in Mathematical Psychology: Information, Learning and Tracking, Free Press of Glencoe, Illinois, 1960. Part I (by Luce) is "The Theory of Selective Information and Some of Its Biological Implications" and covers Shannon's work and some brief mention of Zipf. I did not obtain the papers of Mandelbrot, Miller and Simon referenced there but relied on Luce's rendition of their work, which I hope I have not misrepresented. The bibliography on pp 110-119 of Luce (above) will direct you to the original literature.

The Project would like to thank William Glessner of Central Washington University, Ellensburg, Washington, and Mitchell Lazarus of Education Development Center, Inc., Newton, Massachusetts for their reviews, and all others who assisted in the production of this unit.

This material was class-tested and/or student reviewed in preliminary form by: Stephen Corder, Southern Baptist College, Walnut Ridge, Arkansas; Paul Nugent, Franklin College, Franklin, Indiana, and George C.T. Kung, University of Wisconsin, Stevens Point, Wisconsin, and has been revised on the basis of data received from these sites.

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 or the copyright holder.

11. ANSWERS TO EXERCISES

1. $r \cdot f = k$ and $j \le f < j+1 \Rightarrow j \le \frac{k}{r} < j+1 \Rightarrow \frac{k}{j+1} < r \le \frac{k}{j}$.

   Thus a total of

   $$n_j = \frac{k}{j} - \frac{k}{j+1} = \frac{k}{j(j+1)}$$

   ranks r have associated $f \epsilon [j, j+1)$.

2. Similarly $j - \frac{1}{2} \le f < j + \frac{1}{2} \Rightarrow \frac{k}{j + \frac{1}{2}} < r \le \frac{k}{j - \frac{1}{2}}$

   and

   $$n_j = \frac{k}{j - \frac{1}{2}} - \frac{k}{j + \frac{1}{2}} = \frac{k}{(j - \frac{1}{2})(j + \frac{1}{2})}.$$

3. Using partial fractions

   $$\frac{1}{(j - \frac{1}{2})(j + \frac{1}{2})} = \frac{1}{j - \frac{1}{2}} - \frac{1}{j + \frac{1}{2}},$$

   the partial sum is

   $$s_n = \left(\frac{1}{1/2} - \frac{1}{3/2}\right) + \left(\frac{1}{3/2} - \frac{1}{5/2}\right) + \ldots$$

   $$+ \left(\frac{1}{(2n-1)/2} - \frac{1}{(2n+1)/2}\right) = \frac{1}{1/2} - \frac{1}{(2n+1)/2}.$$

   Thus the series sums to 2. But $\Sigma n_j = 2k >> N$ makes Zipf's model also only partially useful.

4. Comparison and integral tests are easy enough.

5. The rule is $f = j+1$ for theoretical $f \epsilon (j, j+1]$, i.e., for ranks

   $$\frac{k}{j+1} \le r < \frac{k}{j}$$

   (using the solution method of Exercise 1). Then

$$n_{j+1} = \frac{k}{j} - \frac{k}{j+1} = \frac{k}{j(j+1)}$$

for $j = 1, 2, 3, \ldots$. Thus n_1 is excluded, which makes no sense, and the split-up of k totals

$$\sum_{2}^{\infty} \frac{1}{j(j+1)} = \frac{1}{2},$$

also not nicely interpretable. Shifting the terms $n_j \to n_{j+1}$ does not help us fit the Ulysses data better, as an eyeballing of Table 1 will show.

6. All that's missing is

$$1 = \sum_{1}^{\infty} \frac{1}{k(k+1)}$$

$$= \frac{1}{2}\left(\frac{1}{1} + \frac{1}{3}\right) + \frac{1}{4}\left(\frac{1}{3} + \frac{1}{5}\right) + \frac{1}{6}\left(\frac{1}{5} + \frac{1}{7}\right) + \ldots$$

$$= \sum_{n=1}^{\infty} \frac{1}{2n}\left(\frac{1}{2n-1} + \frac{1}{2n+1}\right)$$

$$= \sum_{n=1}^{\infty} \frac{1}{2n}\left(\frac{4n}{(2n-1)(2n+1)}\right) = \sum_{n=1}^{\infty} \frac{2}{(2n-1)(2n+1)}.$$

7. First we show

$$\frac{1}{2} = \frac{1}{1\cdot 3} + \frac{1}{3\cdot 5} + \frac{1}{5\cdot 7} + \ldots$$

$$= \frac{1}{3}\left(\frac{1}{1} + \frac{1}{5}\right) + \frac{1}{7}\left(\frac{1}{5} + \frac{1}{9}\right) + \frac{1}{11}\left(\frac{1}{9} + \frac{1}{13}\right) + \ldots$$

$$= \sum_{n=0}^{\infty} \frac{1}{4n+3}\left\{\frac{1}{(4n+1)} + \frac{1}{(4n+5)}\right\}.$$

$$= \sum_{n=0}^{\infty} \frac{1}{4n+3}\,\frac{8n+6}{(4n+1)(4n+5)} = \sum_{n=0}^{\infty} \frac{2}{(4n+1)(4n+5)}.$$

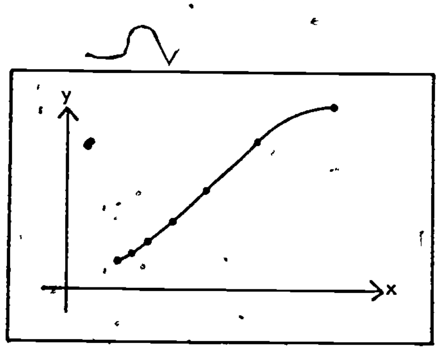
The result then follows at once.

UMAP MODULE 216

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS

Curves and their Parametrization

by Nelson L. Max



Introductory Topology

Birkhäuser Boston Inc.
380 Green Street
Cambridge, MA 02139

ED 036 478
ERIC

Intermodular Description Sheet: *UMAP Unit 216*

Title: CURVES AND THEIR PARAMETRIZATION

Author: Nelson L. Max
Department of Mathematics and Statistics
Case Western Reserve University
Cleveland, Ohio 44106

Review Stage/Date: IV 6/30/80

Classification: INTRO TOPOLOGY

Prerequisite Skills:
1. Understand the representation of points in the plane by Cartesian coordinates.
2. Understand the trigonometric functions of sine and cosine, and the measurement of angles in radians.
3. Understand the natural logarithm function.
4. Sum the geometric and harmonic infinite series.

Output Skills:
1. Define a parametrized curve, image of a curve, orientation of a curve.
2. Given the image of a curve, write the equation of one or more parametrized curves which trace the image with a given orientation.
3. Given the image of a curve, describe all its possible orientations.

Other Related Units:
The Alexander Horned Sphere *(Unit 231)*

The Project would like to thank Joseph Malkevitch of York College of CUNY and Anthony Phillips of SUNY at Stony Brook for their reviews, and all others who assisted in the production of this unit.

This unit was field-tested and/or student reviewed in preliminary form by W. Hugh Haynsworth of The College of Charleston, Charleston, South Carolina; C. Wagner of Pennsylvania State University, Middletown, Pennsylvania; Tom Haigh of Saint John's University, Collegeville, Minnesota; and Phillip Lestmann of Bryan College, Dayton, Tennessee, and has been revised on the basis of data received from these sites.

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 copyright holder.

© 1980 EDC/Project UMAP
All rights reserved.

ERIC

CURVES AND THEIR PARAMETRIZATION

by

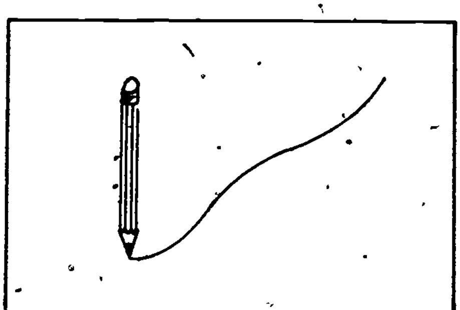
Nelson L. Max
Department of Mathematics and Statistics
Case Western Reserve University
Cleveland, Ohio 44106

TABLE OF CONTENTS

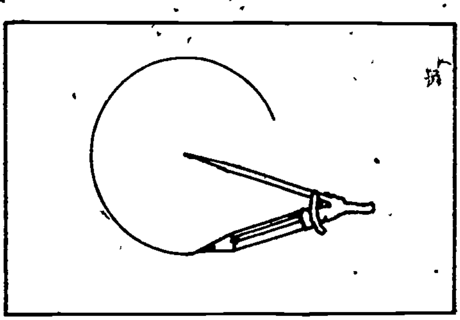
ERIC

1. THE DEFINITION OF "CURVE"

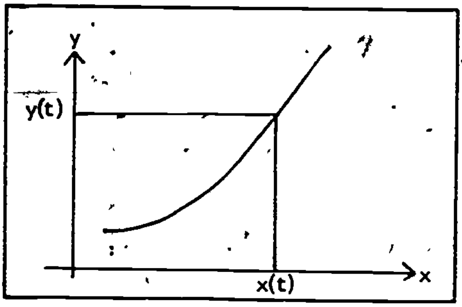
Webster's Dictionary defines a curve as "the path of a moving point." If the moving point were the point of a pencil, it could trace out the curve on paper.



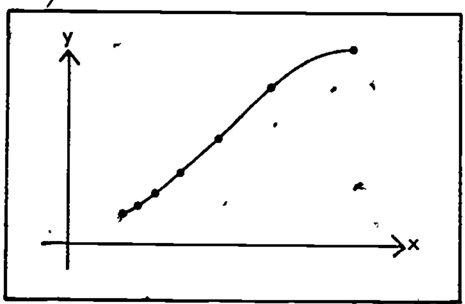
For example, the point of the pencil on a compass might trace out a circle.



Webster gives another more technical definition of a curve: "A line that may be precisely defined by an equation in such a way that its points are functions of a single independent variable or parameter." We can think of the variable or parameter as time and call it t. Then the coordinates of the moving point, x(t) and y(t), are the functions of time.



If we imagine the pencil as making a dot on the curve every second, these dots will show how the curve has been traced. In particular, their spacing will indicate the speed of the moving point. Here the point is speeding up as it moves to the right.

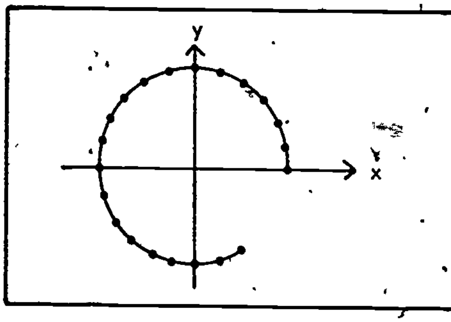


2. PARAMETRIZATIONS OF THE UNIT CIRCLE

Below is a circle which is traced counter-clockwise at a uniform speed of 15°, or $\frac{\pi}{12}$ radians, every second. When it is finished in 24 seconds, it will have 24 evenly spaced dots. The coordinates of the moving point are given by the equations

$$x(t) = \cos\left(\frac{\pi}{12}\, t\right);$$

$$y(t) = \sin\left(\frac{\pi}{12}\, t\right).$$



There are many other ways to trace the same circle. In the figure below we see only twelve evenly spaced dots, so the equations might be

$$x(t) = \cos\left(\frac{\pi}{6}t\right);$$

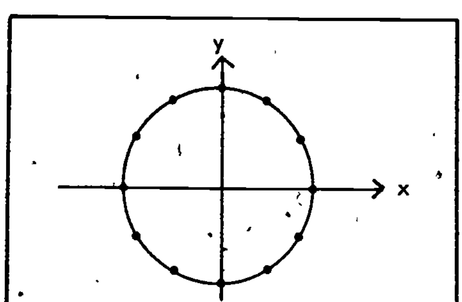
$$y(t) = \sin\left(\frac{\pi}{6}t\right).$$

However, they might also be

$$x(t) = \cos\left(\frac{\pi}{6}t\right);$$

$$y(t) = -\sin\left(\frac{\pi}{6}t\right)$$

which would trace the circle with the same constant speed in the opposite direction.



ERIC
Full Text Provided by ERIC

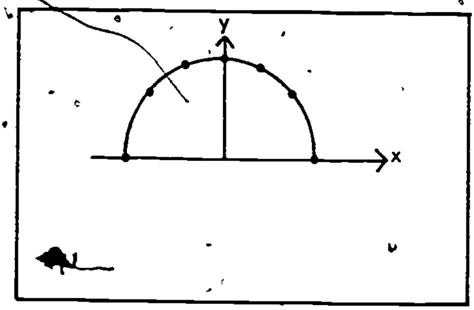
We might also trace the circle by having the x coordinate move at a uniform speed from 1 to - 1, for example,

$$x(t) = 1 - \frac{1}{3}t \qquad 0 \le t \le 6$$

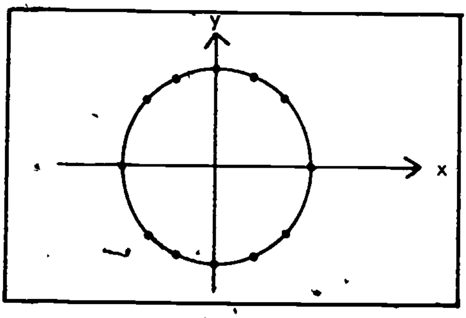
$$y(t) = \sqrt{1 - x^2} = \frac{1}{3}\sqrt{6t - t^2}$$

These equations work only to trace out a semicircle.

Here the dots are not evenly spaced. They are closest together at the top and bottom, indicating that the curve is traced most slowly there. The tracing point actually moves infinitely fast at the left and right sides.



QUESTION A: Can you find similar equations to trace out the bottom semicircle, for $6 \le t \le 12$?

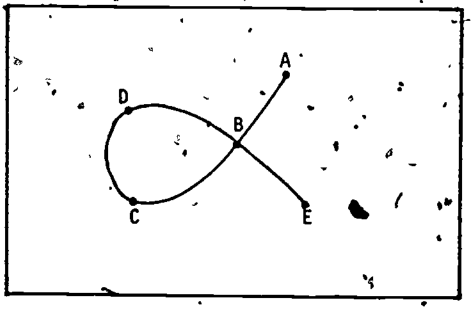


All these different functions define different *parametrizations* of the circle. We say that they define different *parametrized curves*. The set of points which a curve passes through is called its *image*. All the different parametrizations of the circle have the same image.

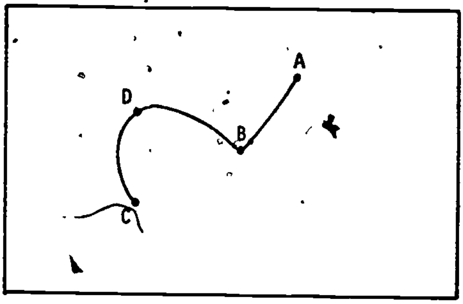
In addition to defining the speed, a parametrization also defines the order in which the points in the image are traced. Thus, a point tracing a clockwise circle moves in the opposite direction from a point tracing a counter clockwise circle, so it passes through the points in the image in the opposite order. Thus there are two *orientations* to the circle, clockwise and counter clockwise.

3. OTHER PARAMETRIZED CURVES

The situation becomes more complicated if the curve is not one-to-one, i.e., if it passes through some points more than once. Here is a curve which crosses itself, passing through the point B twice. One orientation would be to pass through the points on the image in the order ABCDBE.

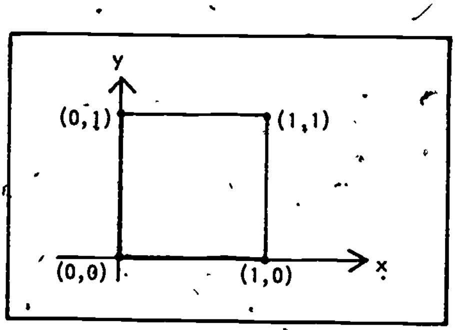


Another method of tracing the same image, shown partly completed here, would pass through the points in the order ABDCBE, making two corners at B. Two more orientations would start at E and end at A. QUESTION B: What are they?



Although a curve can pass through certain points on its image more than once, it should not cover whole sections more than once. Thus, ABCDBDCBE would not give a valid orientation for the curve.

There is nothing wrong with a corner in a curve. A mathematical curve is not necessarily a smoothly curving line, but may have corners, and can even consist entirely of straight lines. For example, a square is a curve. Can you find a set of equations which describe this curve?



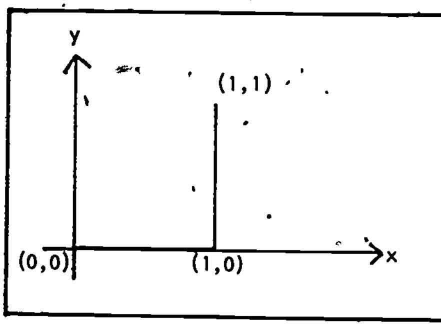
The trick is to find separate formulas for the different sides of the square, just as separate formulas could be used for the two semicircles making up a circle.

ERIC
Full Text Provided by ERIC

The functions below define two sides of the square.

$$x(t) = \begin{cases} t, & 0 \le t \le 1 \\ 1, & 1 \le t \le 2 \end{cases}$$

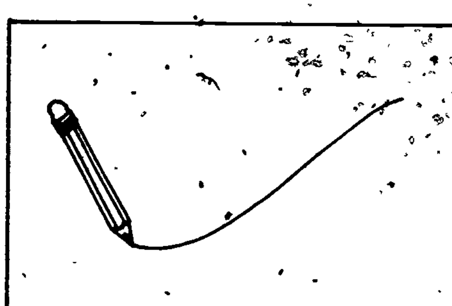
$$y(t) = \begin{cases} 0, & 0 \le t \le 1 \\ t-1, & 1 \le t \le 2 \end{cases}$$



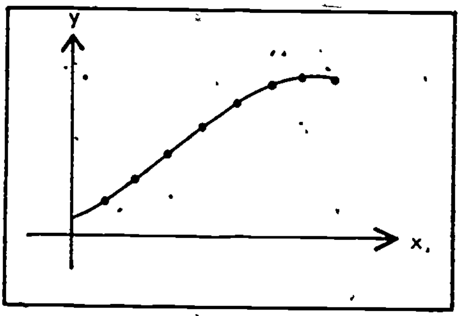
QUESTION C: Can you continue these functions for $2 \le t \le 4$ to define the other two sides?

4. CONTINUOUS CURVES

A natural subcollection of the class of parametrized curves are ones for which the tracing point moves continuously, without jumping. This condition is equivalent to requiring that the curve can be drawn without lifting the pencil from the paper.



It is also equivalent to requiring that the two coordinates x(t) and y(t) be continuous functions of the time parameter t.

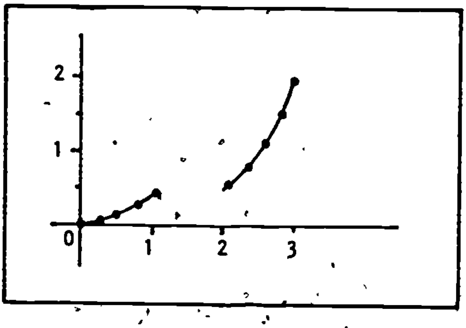


If one of the coordinate functions is discontinuous, for example,

$$x(t) = \begin{cases} t, & 0 \le t \le 1 \\ t+1, & 1 \le t \le 2 \end{cases}$$

$$y(t) = \frac{1}{2} t^2$$

the resulting image, shown below, may have a gap in it. If both x(t) and y(t) are continuous, the result will be a continuous parametrized curve, called simply a curve for short.



5. CURVES WITH UNUSUAL PROPERTIES

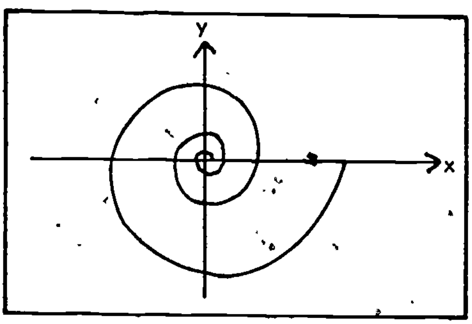
There are many strange examples which satisfy this definition of curve.

Example 1

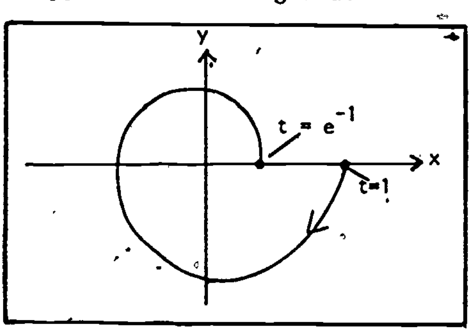
For example, if $\ln x$ denotes the natural logarithm of x, then the equations

$$x = t\cos(2\pi \ln t)$$
$$y = t\sin(2\pi \ln t)$$

which make sense on the interval $0 < t \le 1$, can be extended to a continuous function on $0 \le t \le 1$ by defining $x(0) = y(0) = 0$. This gives a curve, called the logarithmic spiral, which has infinitely many (similar) turns near $t = 0$. Nevertheless we will prove that it has finite length.

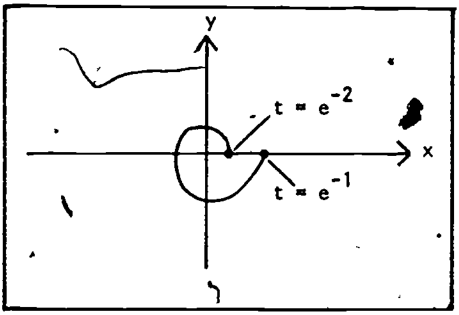


Consider the first turn of the spiral, from $t = 1$ to $t = e^{-1}$. Suppose it has length L.



9

The next turn of the spiral, from $t = e^{-1}$ to $t = e^{-2}$ looks exactly similar, but e^{-1} is large, so its length is Le^{-1}. Similarly the length of the next turn is Le^{-2}. Thus the length of the whole spiral is $L + Le^{-1} + Le^{-2} + Le^{-3} \ldots$, a geometric series which converges to $L/(1 - e^{-1})$, a finite length.



Problem

Verify that the turn of the spiral from $t = e^{-1}$ to $t = e^{-2}$ is similar to the turn from $t = 1$ to $t = e^{-1}$ with constant of proportionality $\frac{1}{e}$.

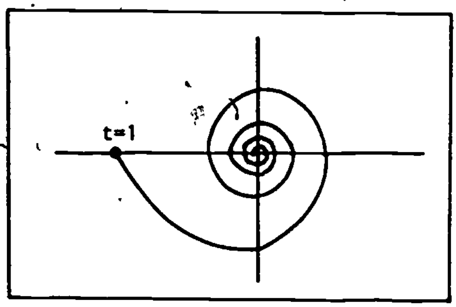
Example 2

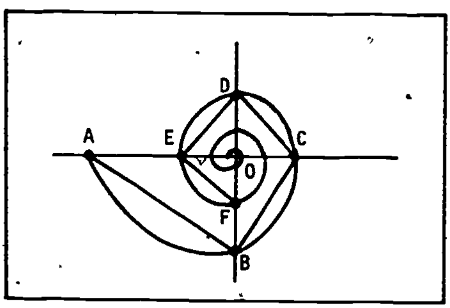
There is also a spiral which winds toward the origin in such a way that it has infinite length. It is the hyperbolic spiral

$$x = t \cos\left(\frac{\pi}{t}\right)$$

$$y = t \sin\left(\frac{\pi}{t}\right)$$

which can again be defined for $0 \leq t \leq 1$ by letting $x(0) = y(0) = 0$. The length of the spiral must be at least as long as the length of the inscribed polygon ABCDE... which we will show is infinite. If 0 is the origin, then AB and BC, are both longer than OB, while CD and DE are both longer than OD...and so forth. So the length of the spiral is greater than twice the sum of the lengths of the line segments from the origin to the "y crossings," the points where the curve crosses the y-axis, the first three of which are B, D, F.





How long is OB? Consider the intersections of the spiral with the y-axis (x=0). Since $t \neq 0$, we have

$$\cos\left(\frac{\pi}{t}\right) = 0$$

so that,

$$\frac{\pi}{t} = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}, \ldots$$

i.e.,

$$t = 2, \frac{2}{3}, \frac{2}{5}, \frac{2}{7}, \ldots$$

Since $t \leq 1$, the acceptable values for $\frac{\pi}{t}$ are $\frac{3\pi}{2}, \frac{5\pi}{2}, \ldots$. The length of OB is the y value when

$$\frac{\pi}{t} = \frac{3\pi}{2} \left(t = \frac{2}{3}\right): y\left(\frac{2}{3}\right) = -\frac{2}{3}, \text{ so } OB = \frac{2}{3}.$$

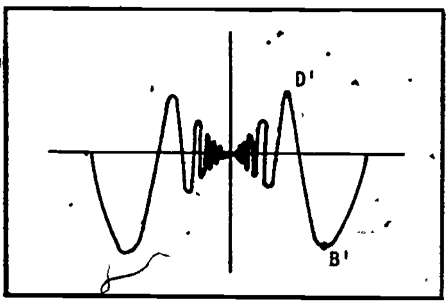
Below is the graph of $y = t \sin\left(\frac{\pi}{t}\right)$, with the point $B' = \left(\frac{2}{3}, -\frac{2}{3}\right)$ giving the y coordinate of the point B, i.e., the length of OB. Similarly $D' = \left(\frac{2}{5}, \frac{2}{5}\right)$ gives the y coordinate for the point D. The length of the spiral, which is greater than the length of the polygon, is thus greater than the series

$$2\left(\frac{2}{3}\right) + 2\left(\frac{2}{5}\right) + 2\left(\frac{2}{7}\right) + \ldots = 4\left(\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots\right)$$

which diverges. So the length is infinite.

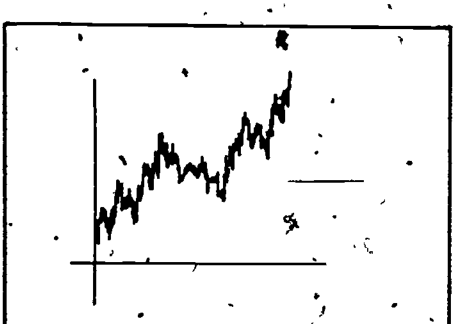
Example 3

The infinitely wiggly graph of $y = t \sin\left(\frac{\pi}{t}\right)$ also has infinite length by a similar argument.



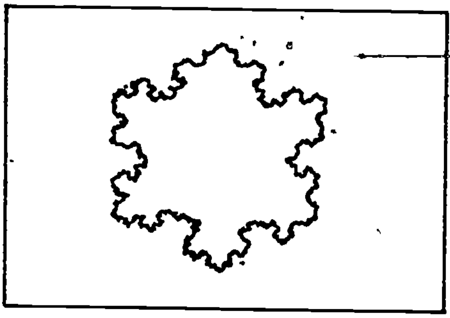
Other Examples

There are functions whose graphs have infinitely many wiggles, and infinite length, between any two points.

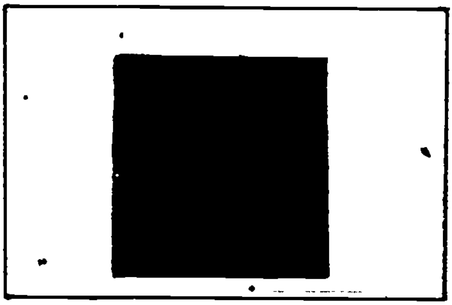


ERIC
Full Text Provided by ERIC

The snowflake curve also has infinite length between any two of its points.

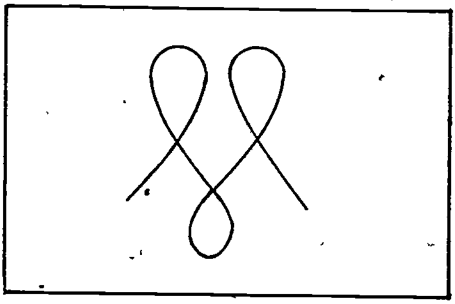


Among the strangest examples of curves, are the "space filling curves," which pass through every point in an area such as a square.



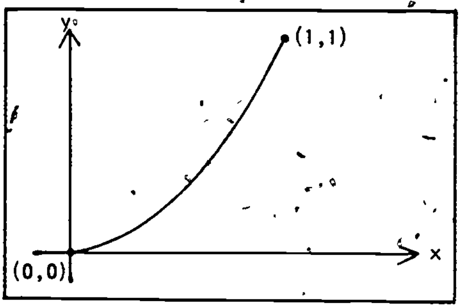
6. EXERCISES

1. How many different parameterized curves have the image shown below?

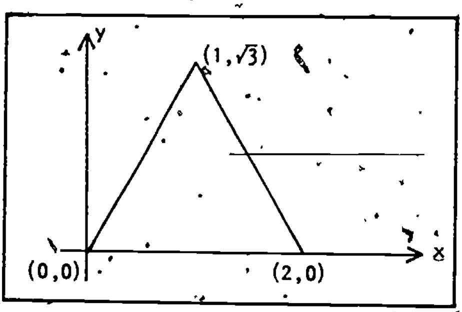


2. How many different oriented curves have this image?

3. Find two different parametrized curves, defined for $0 \leq t \leq 1$, which have this piece of a parabola as an image?



3. Find the equations for the parametrized curve which traces this equilateral triangle at uniform speed in a counter-clockwise direction, starting at the vertex (0,0), in the time interval $0 \leq t \leq 3$.



7. MODEL EXAM

1. State whether each of the curves described below is an oriented curve, a parametrized curve, or the image of a curve.

   a) The contrail left by a jet plane.

   b) The script letter m, drawn from left to right.

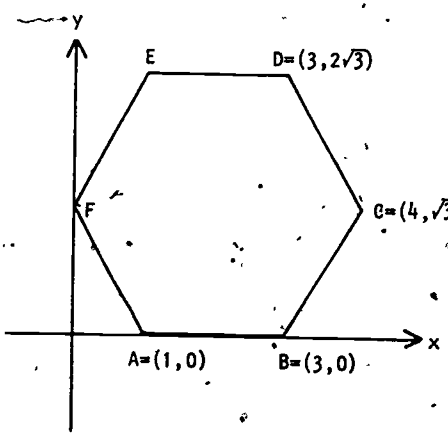
   c) A marathon course.

   d) The straight path of an automobile, accelerating uniformly from 0 to 60 miles per hour in ten seconds.

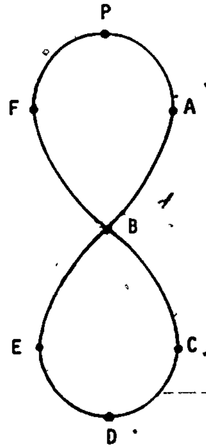
   e) A figure eight.

   f) The path of the tip of a second hand on a wall clock.

2. If ABCD is the curve defined by tracing the first three sides of the hexagon below, with constant speed in the time interval $0 \leq t \leq 3$, find the formulas for $x(t)$ and $y(t)$.



3. Describe all possible orientations of the figure 8, starting at the top point P.



4. What is meant by the image of a parametrized curve?

5. Give a parametrized curve whose image is $\{(x,y) \mid 0 \le x \le 1,\ 0 \le y \le 1,\ x^2 = y^3\}$.

ERIC Full Text Provided by ERIC

8. ANSWERS TO QUESTIONS IN TEXT

A. $x(t) = \frac{1}{3}t - 3$.

$$y(t) = -\sqrt{1 - x^2} = -\sqrt{1 - \frac{1}{3}(t-3)^2} = -\frac{1}{3}\sqrt{18t - t^2 - 72}.$$

B. The order EBDCBA and the order EBCDBA.

C. A set of the equations for all four sides of the square are:

$$x(t) = \begin{cases} t, & 0 \le t \le 1 \\ 1, & 1 \le t \le 2 \\ 3-t, & 2 \le t \le 3 \\ 0 & 3 \le t \le 4 \end{cases} \qquad y(t) = \begin{cases} 0, & 0 \le t \le 1 \\ t-1, & 1 \le t \le 2 \\ 1, & 2 \le t \le 3 \\ 4-t, & 3 \le t \le 4 \end{cases}$$

9. ANSWERS TO EXERCISES

1. Infinitely many.

2. Sixteen.

3. (Possible answers)

$x(t) = 1 - t, \quad y(t) = (1-t)^2;$

$x(t) = t, \quad y(t) = t^2;$

$x(t) = t^2, \quad y(t) = t^4;$

$x(t) = \sqrt{t}, \quad y(t) = t.$

4. $$x(t) = \begin{cases} 2t, & 0 \le t \le 1 \\ 3-t, & 1 \le t \le 3 \end{cases} \qquad y(t) = \begin{cases} 0, & 0 \le t \le 1 \\ \sqrt{3}(t-1), & 1 \le t \le 2 \\ \sqrt{3}(3-t), & 2 \le t \le 3 \end{cases}$$

10. ANSWERS TO MODEL EXAM

1. a) image b) oriented curve c) oriented curve
   d) parametrized curve e) image f) parametrized curve

2.

$$x(t) = \begin{cases} 1 + 2t, & 0 \le t \le 1 \\ 2 + t, & 1 \le t \le 2 \\ 6 - t, & 2 \le t \le 3 \end{cases} \qquad y(t) = \begin{cases} 0 & 0 \le t \le 1 \\ \sqrt{3}\,(t-1) & 1 \le t \le 3 \end{cases}$$

3. PABCDEBFP, PABEDCBFP, PFBEDCBAP, and PFBCDEBAP.

5. Possible answers

$$x(t) = t,\ y(t) = t^{2/3} \qquad 0 \le t \le 1$$

$$x(t) = t^3,\ y(t) = t^2 \qquad 0 \le t \le 1$$

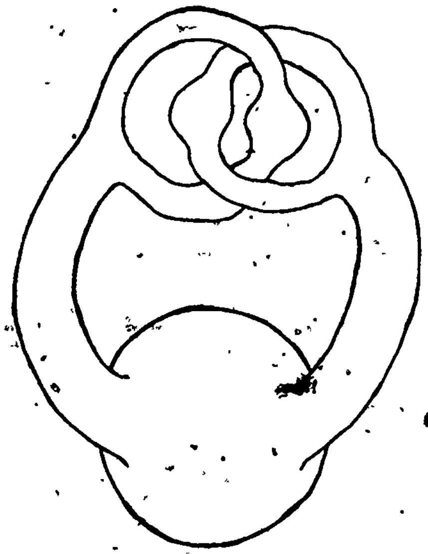
umap

UNIT 231

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS PROJECT

THE ALEXANDER HORNED SPHERE

by Nelson L. Max



INTRODUCTORY TOPOLOGY

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

THE ALEXANDER HORNED SPHERE

by

Nelson L. Max
Department of Mathematics and Statistics
Case Western Reserve University
Cleveland, Ohio 44106

TABLE OF CONTENTS

ED036478
ERIC

Intermodular Description Sheet: *UMAP Unit 231*

Title: THE ALEXANDER HORNED SPHERE

Author: Nelson L. Max
Department of Mathematics and Statistics
Case Western Reserve University
Cleveland, Ohio 44106

Review Stage/Date: III 6/20/77

Classification: INTRO TOPOLOGY

Suggested Support Material:
Films:
The Alexander Horned Sphere, 2½ minutes, silent, color.
Zooms, 7 minutes, color, musical track.
Available from:
International Film Bureau
332 South Michigan Avenue
Chicago, Illinois 60604

Prerequisite Skills:
1. Parametrization of simple closed curves.
2. Topological definitions of connectedness, open and closed sets, continuous functions, homeomorphisms.

Output Skills:
1. Understand the construction of the Alexander horned sphere.
2. Discover properties of the horned sphere as a counterexample to Shoenfliess Theorem for the standard sphere S^1.

Other Related Units:
Curves and Their Parametrizations *(Unit 216)*
Turning a Sphere Inside Out *(Unit 289)*

© 1978 EDC/Project UMAP
All Rights Reserved.

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 education research in the U.S. and abroad.

PROJECT STAFF

Ross L. Finney	Director
Solomon Garfunkel	Associate Director/Consortium Coordinator
Felicia Weitzel	Associate Director for Administration
Barbara Kelczewski	Coordinator for Materials Production
Dianne Lally	Project Secretary
Paula M. Santillo	Financial/Administrative Secretary
Jack Alexander	Editorial Consultant
Edwina Michener	Editorial Consultant
Louise Raphael	Editorial Consultant

NATIONAL STEERING COMMITTEE

W.T. Martin	MIT (Chairman)
Steven J. Brams	New York University
Llayron Clarkson	Texas Southern University
James D. Forman	Rochester Institute of Technology
Ernest J. Henley	University of Houston
Donald A. Larson	SUNY at Buffalo
William F. Lucas	Cornell University
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 Joseph Malkevitch and Solomon Garfunkel for their reviews and all others who assisted in the production of this unit.

This material was prepared with the support of Nation 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.

1. INTRODUCTION

In this unit we will describe the image of a homeomorphism from the standard sphere into three dimensional space, whose exterior is not homeomorphic to the exterior of a standard sphere. It is called the Alexander horned sphere because it was discovered by J.W. Alexander in 1924, and looks as if it has grown horns. We will start by discussing the situation for simple closed curves in the plane. Then we will describe the horned sphere, and suggest the idea behind the proof that it has a non-standard exterior.

2. THE JORDAN CURVE AND SHOENFLIESS THOEREMS

A simple closed curve is a closed curve which does not cross itself. If it is parametrized by a continuous function f from the interval [0,1] to the plane R^2, then $f(a) = f(b)$, for $a < b$, if and only if $a = 0$ and $b = 1$. (See Figure 1.)

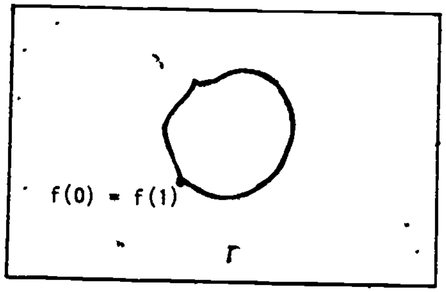


Figure 1

If S^1 stands for the unit circle, $\{(x,y)\varepsilon R^2 | x^2+y^2=1\}$, we may also think of our curve as a homeomorphism g of S^1 into the plane. This means that g is a homeomorphism of S^1 onto its image $g(S^1)$, although $g(S^1)$ is not necessarily the whole plane.

Suppose we have such a simple closed curve g. The Jordan curve theorem states that $g(S^1)$ separates the plane into the union of two non-empty connected open sets A and B. That is, $R^2 - g(S^1) = A \cup B$, A and B are both non-empty and open, and in particular, $g(S^1)$ is the complete frontier of both A and B. The previous set, A, is called the interior of the curve, and the unbounded one, B, is called the exterior. (See Figure 2.)

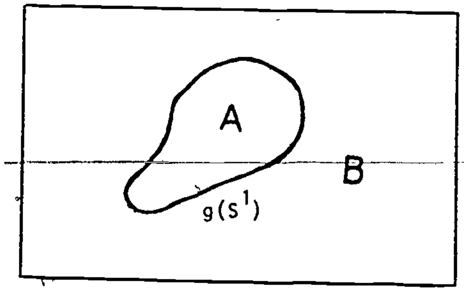


Figure 2

The Shoenfliess Theorem states in addition that the homeomorphism g, which is defined only on the unit circle S^1, can be extended to the whole plane, so that it takes the interior of S^1 to A and the exterior to B. Thus A and B are homeomorphic to the standard "round" regions.

We will not prove either of these theorems here.

3. THE HORNED SPHERE

Let R^3 denote the three dimensional space of triples of real numbers (x,y,z), let $S^2 = \{(x,y,z,)\varepsilon R^3 | x^2 + y^2 + z^2 = 1\}$ be the surface of standard round sphere in R^3, and let g be a homeomorphism of S^2 into R^3. Then the generalization of the Jordan Curve Theorem, sometimes called the Jordan Separation Theorem, states that $R^3 - g(S^2) = A \cup B$, the union of two non-empty connected open sets, and $g(S^2)$ is the complete frontier of each. Again, A represents the interior of the distorted sphere $g(S^2)$ and B represents the exterior. (See Figure 3.)

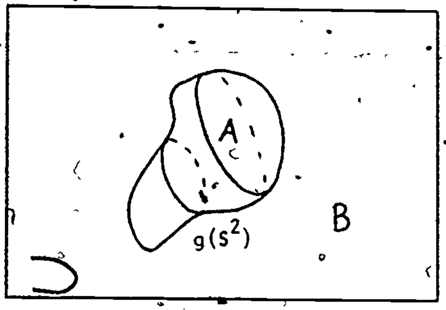


Figure 3

The analogous generalization is actually true in any number of dimensions.

However, the generalization of the Shoenfliess Theorem is not true in the case of $g(S^2)$ and the Alexander Horned Sphere is a counterexample. In this section, we will construct a homeomorphism g of S^2 into R^3, such that the exterior B of $g(S^2)$ is not homeomorphic to the exterior of the standard sphere S^2. In particular, then, it will not be possible to extend g to a homeomorphism of the exterior of S^2 onto B.

To construct the horned sphere, we start with a round sphere as the first approximation and push out a pair of horns to make the second approximation. We can do this by taking two pairs of concentric discs on the sphere, $D_0 \subset C_0$, and $D_1 \subset C_1$. Then we keep $S^2 - (C_0 \cup C_1)$ fixed, push $C_0 - D_0$ and $C_1 - D_1$ to the tubular sides of the horns, leaving circular caps made from D_0 and D_1, as shown in Figure 4.

From the flat ends of these horns, we push out two new branches in the same way to get the third approximation. It looks like a pair of crab's claws partially interlocked but not closed or touching. To do this, we

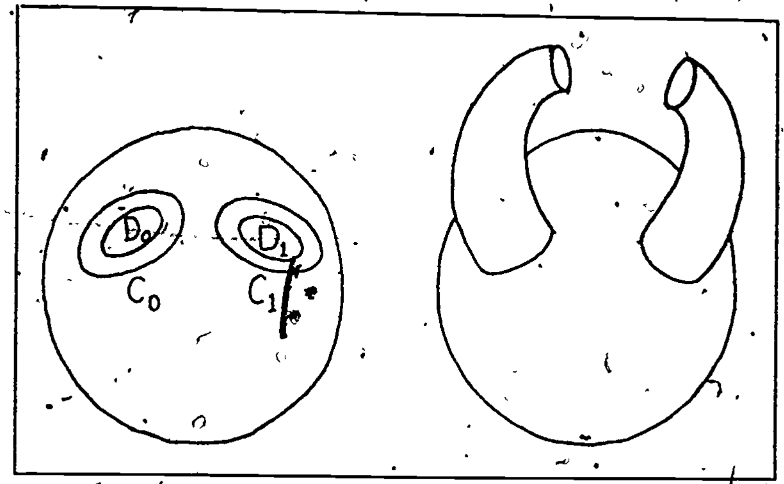


Figure 4

need only move points which lie inside the four discs C_{00}, C_{01}, C_{10}, and C_{11}. (Figure 5.)

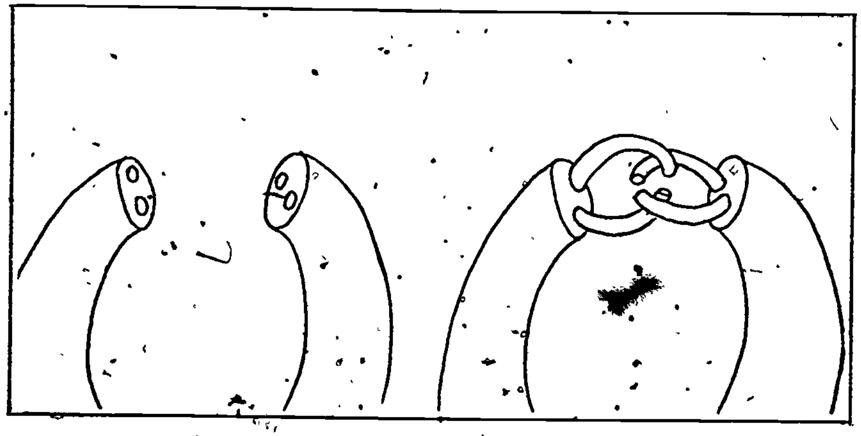


Figure 5

We repeat again and again, growing new branches on the tops of each of the old branches. Since each new pair of claws is a reduced version of the previous pair, the total amount any point moves is dominated by a geometric progression. Therefore, the approximations converge uniformly to a continuous limit function g from S^2 to R^3. By the way the construction is arranged, g is also one-to-one, so it can be proved that g is a homeomorphism. (Figure 6.)

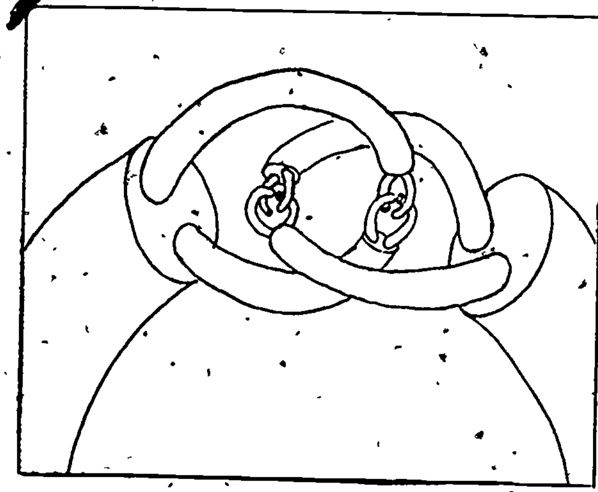


Figure 6

We can round the corners of our surface to make a new function g which is smooth (Figure 7), except at the points which belong to an infinite number of the discs C_i. These exceptional points are called *wild* points. If we take any infinite binary expansion, say .01100110..., we can get a corresponding contracting sequence of discs $C_0 \supset C_{01} \supset C_{011} \supset C_{0110} \supset \ldots$ which contains a wild point P in common. Thus there is at least one wild point for every real number between 0 and 1, so that the collection of wild points is uncountable.

Let $C = \{(x,y,z) \in R^3 | x^2+y^2+z^2 | < 1\}$ be the interior of the unit sphere S^2. Then we could pull C along as we push out S^2, so the function g can be extended to C, giving a homeomorphism of the closed ball $S^2 \cup C$ into R^3. Therefore the interior A of $g(S^2)$ is homeomorphic to the round ball C.

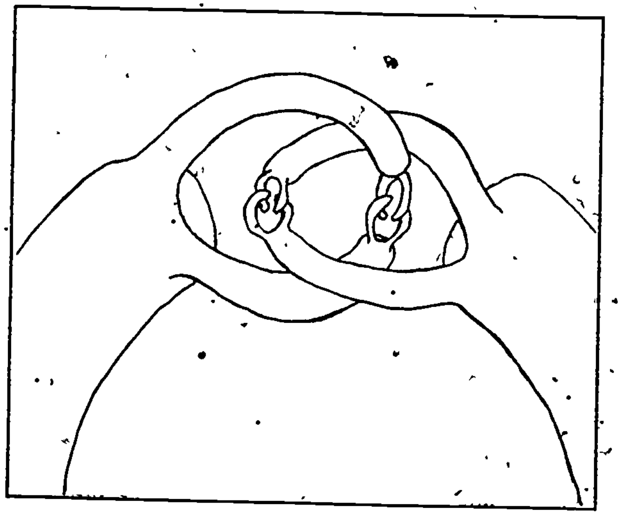


Figure 7

But what about the exterior B of $g(S^2)$? We will show in the next section why B is not homeomorphic to the exterior D of the unit sphere S^2.

4. SIMPLY CONNECTED SETS

The demonstration that B is not homeomorphic to D uses the following topological property. A set X is called *simply connected* if every closed curve in X (called a loop in X for short) can be shrunk continuously in X until its image is a single point.

For example, the exterior D of S^2 in R^3 is simply connected, because every loop L can be pulled off the sphere and collapsed to a single point P. A number of intermediate positions are shown in Figure 8.

Suppose the loop L is parametrized by a continious function f(s) from [0,1] to X, and that the shrinking motion takes place for t in the time interval [0,1]. Then for each fixed t, we get an intermediate curve

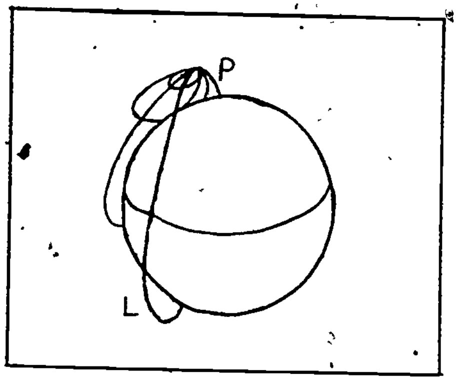


Figure 8

curve $f_t(s)$, which is also closed and continuous, and these intermediate curves depend continuously on t. The intermediate curve must agree with $f(s)$, when $t = 0$, and stay fixed at P when $t = 1$. Thus, a parametrization of the shrinking motion is a continuous function $F(s,t) = f_t(s)$ of two variables, $s \varepsilon [0,1]$, which marks distance along each curve, and $t \varepsilon [0,1]$, which marks the different intermediate curves in the motion.

It must satisfy:

a. $F(s,0) = f(s)$ for all s,
b. $f(0,t) = F(1,t)$ for all t, and
c. $F(s,1) = P$ for all S.

Such a function F is called a *homotopy*. It is said to shrink the loop f in X to the point P.

If T denotes the solid donut, or *torus*, shown in Figure 9, then its exterior $Y = R^3 - T$ is not simply connected. The loop L, which wraps around the hole, cannot be shrunk to a point without crossing T.

Suppose there were a homeomorphism h from the exterior D of S^2 to the exterior Y of T. Then, knowing D is simply connected, we could prove Y to be simply connected as follows. Let f parametrize a closed loop

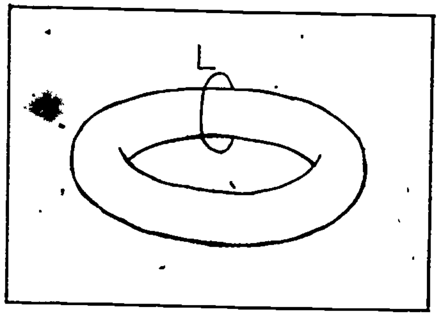


Figure 9

in Y. Then $h^{-1} \circ f$ is a closed loop in D. Since D is simply connected, there is a homotopy F which shrinks the loop $h^{-1} \circ f$ in D to a point P. Then $h \circ F$ will shrink the loop f in Y to a point $h(P)$. Since this works for any loop f in Y, Y is simply connected.

We say simply connectedness is a *topological property*, because it is preserved by homeomorphisms.

5. THE EXTERIOR OF THE HORNED SPHERE

We can prove similarly that the exterior of B of the Alexander Horned Sphere is not homeomorphic to D, if we can show that it is not simply connected.

At first, this might seem difficult, because the claws never touched, so the exterior of each approximation is simply connected. However, a property which is true of each of a sequence of approximations is not necessarily true of the limit. In fact, we can define the horned sphere differently, so that the exterior of each approximation is not simply connected.

Imagine you are carving the solid horned sphere $g(S^2 \cup C)$ out of a piece of wood. The first approximation K_1 will be a torus with two bulges, one for the original sphere, and one to contain the claws, as shown in Figure 10.

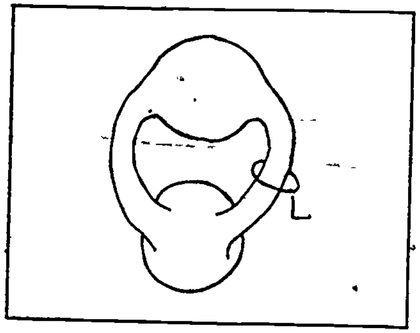


Figure 10

The exterior K_1 is not simply connected, since the loop L cannot be shrunk.

The next step will be to carve out two claws from the upper bulge (see Figure 11), leaving their tips connected by two smaller bulges. The result K_2 has a non-simply-connected exterior, since the loop L still cannot be shrunk.

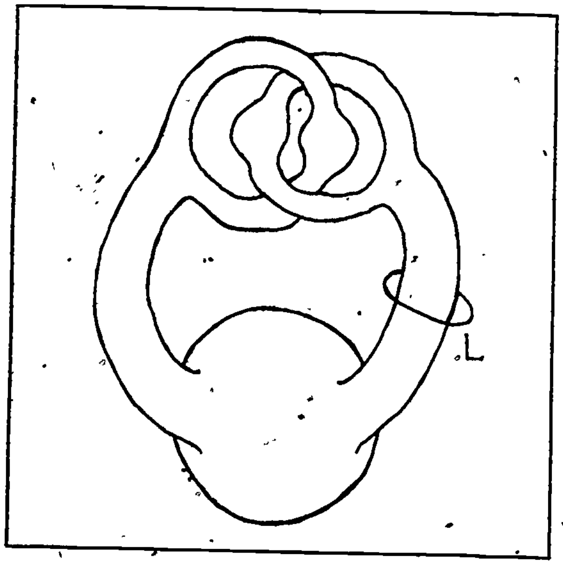


Figure 11

If we continue, we get a sequence of closed sets $K_1 \supset K_2 \supset K_3 \supset \ldots \supset g(S^2 \cup C)$, each of whose exteriors is non-simply connected. (See Figure 12.)

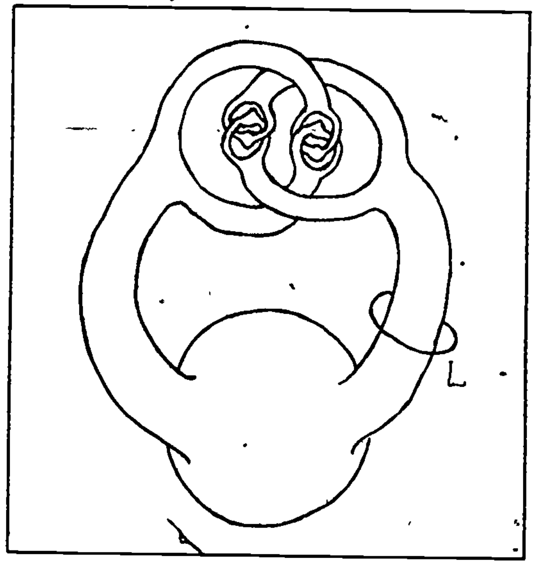


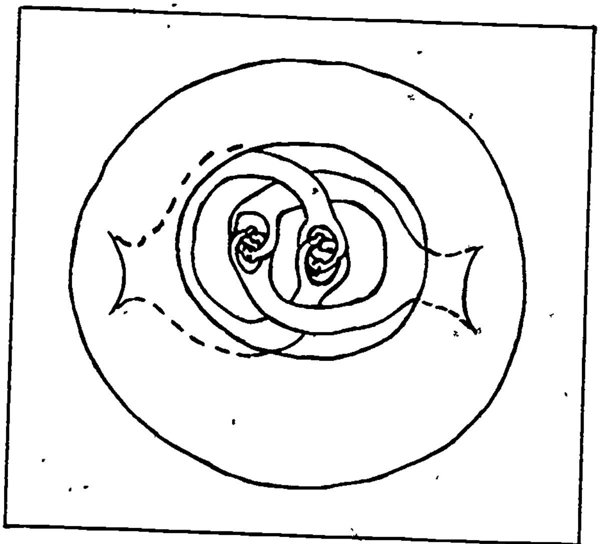
Figure 12

Now suppose the loop L could be shrunk to a point in the exterior B of $g(S^2)$, using a homotopy $F(s,t)$. Since the image of F does not meet $g(S^2 \cup C)$, it must remain a finite distance ϵ away. But now we find a solid approximation K_n within ϵ of $g(S^2 \cup C)$, and the image of the homotopy will also miss K_n. This contradicts the fact that L cannot be shrunk to a point on the exterior of K_n.

6. PROBLEM

Draw a sphere $g(S^2)$ such that its interior A is not homeomorphic to the interior C of a round sphere.

Solution: Push the horns into the inside of the sphere. A hole has been cut away from the surface to make them visible.



380

ERIC

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.

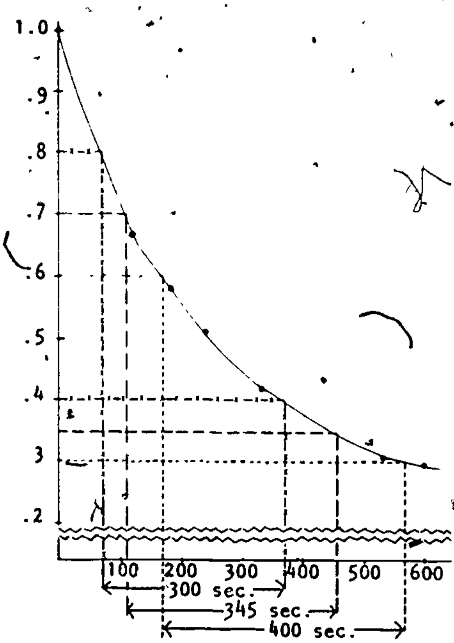
Please describe anything that you found particularly helpful. (Please use the back of this sheet if you need more space.)

UMAP MODULE 232

MODULES AND MONOGRAPHS IN UNDERGRADUATE MATHEMATICS AND ITS APPLICATIONS

Kinetics of Single Reactant Reactions

by Brindell Horelick and Sinan Koont



Applications of Calculus to Chemistry

Birkhäuser Boston Inc.
380 Green Street
Cambridge, MA 02139

SE 036 478

ERIC

KINETICS OF SINGLE REACTANT REACTIONS

by

Brindell Horelick
Department of Mathematics
University of Maryland Baltimore County
Baltimore, MD 21228

and

Sinan Koont
Department of Economics
University of Massachusetts
Amherst, MA 01003

TABLE OF CONTENTS

ERIC Full Text Provided by ERIC

Intermodular Description Sheet: *UMAP Unit 232*

Title: KINETICS OF SINGLE REACTANT REACTIONS

Author: Brindell Horelick and Sinan Koont
Department of Mathematics, University of Maryland Baltimore County, Baltimore, Maryland 21228
Department of Economics, University of Massachusetts, Amherst, Massachusetts 01003

Review Stage/Date: IV 7/2/79

Classification: APPL CALC/CHEM-KINETICS

References:

Bares, J., Cerny, C., Fried, V., and J. Pick (1962). Collection of Problems in Physical Chemistry. Addison Wesley, Reading, MA.
Capellos, C., and B. Bielski (1972). Kinetic Systems. Wiley-Interscience, NY.
Frost, A.A., and R.G. Pearson (1961). Kinetics and Mechanism. (2nd ed.) John Wiley and Sons, Inc. NY.
Laidler, K.J. (1965). Chemical Kinetics. McGraw-Hill, Inc., NY.
Stevens, B. (1970). Chemical Kinetics. Chapman and Hall, London.
Weston, R.E. Jr., and H.A. Schwarz (1972). Chemical Kinetics. Prentice-Hall, Englewood Cliffs, NY.

Prerequisite Skills:

1. Be familiar with the Cartesian coordinate system.
2. Understand that $a'(t)$ describes the rate of change of $a(t)$.
3. Be able to integrate
$$\int_0^t \frac{a'(t)}{a^n(t)}dt$$
for $n = 0, \frac{1}{2}, 1, 2, 3$.
4. Be able to solve an exponential equation.

This unit is intended for Calculus students with an active interest in and some background knowledge of chemistry. This background may be represented by concurrent registration in a college level chemistry course.

Output Skills:

1. Be able to describe single reactant irreversible reactions, including definitions of rate constant, reaction order, and half-life.
2. Be able to find explicit formulas for $a(t)$ and for the half-life for a reaction of order n.
3. Be able to determine the reaction order and rate constant of a reaction, given data on $a(t)$, provided the reaction is of order 0, 1, or 2.

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 copyright holder.

© 1979 EDC/Project UMAP
All rights reserved.

1. SINGLE REACTANT IRREVERSIBLE REACTIONS

1.1 Definition and Some Examples

Suppose we have a chemical reaction of a particularly simple sort, one which involves only one substance (let us call it A) as a reactant, and which is irreversible, therefore going to completion. It may be represented by writing:

$$A \rightarrow B_1 + B_2 + \ldots + B_n ,$$

where B_1, B_2, ... B_n are the products. Suppose at time $t = 0$ we have a certain concentration a_0 of A (measured, for example, in moles per liter). It is possible to observe and record the concentration $a(t)$ of A at various later times t.

TABLE I

Experimental Data from Three Single-Reactant Irreversible Reactions.

(a)

t (seconds)	0	51	206	454	751	1132	1575	2215
a(t) (mm Hg)	15.03	14.58	13.32	11.49	9.73	7.79	6.08	4.17

(b)

t (minutes)	0	1	4	10	30	40
a(t) (mm Hg)	55	50	38	21	3	1.5

(c)

t (seconds)	0	120	180	240	330	530	600
$a(t)/a_0$	1	.6705	.5825	.512	.4195	.310	.2965

ERIC Full Text Provided by ERIC

Table I gives three sets of such observations. Part (a) is for an experiment conducted at 280°C involving the decomposition of trichloromethyl chloroformate into phosgene:

$$Cl-\overset{O}{\overset{\|}{C}}-O-CCl_3 \rightarrow 2Cl-\overset{O}{\overset{\|}{C}}-Cl.$$

Part (b) is for the decomposition at 500°C of ethylamine into ethylene and ammonia:

$$C_2H_5NH_2 \rightarrow C_2H_4 + NH_3.$$

Part (c) is for alkaline hydrolysis of ethyl nitrobenzoate at an initial concentration of 0.05 moles per liter.

The reactants in parts (a) and (b) are gaseous. At constant temperature and volume, a(t) is proportional to its partial pressure, and it is this figure, in millimeters of mercury (mm Hg) that appears in Table I. In part (c), a(t) is given as a fraction of a_0.

In the conversion of trichloromethyl chloroformate to phosgene, both the reactant and the product are gaseous, and the total pressure actually increases as the reaction proceeds because each trichloromethyl chloroformate molecule gives rise to two phosgene molecules. The partial pressure of the trichloromethyl chloroformate is deduced from the total pressure by taking the reaction equation and the original pressure into account. In many reactions, however, the amount of the reactant is determined by techniques based on its absorption of light.

1.2 Graphs of the Results

We have plotted these results in Figures 1, 2, and 3. In that all of the curves decrease as t increases, these curves look very similar. But there is at least one significant difference (aside from the differences of scales). In each figure we have selected various concentrations of A and determined graphically approximately how long it takes for a(t) to decrease from the selected concentration to half of it. For example, Figure 1 shows us

ERIC Full Text Provided by ERIC

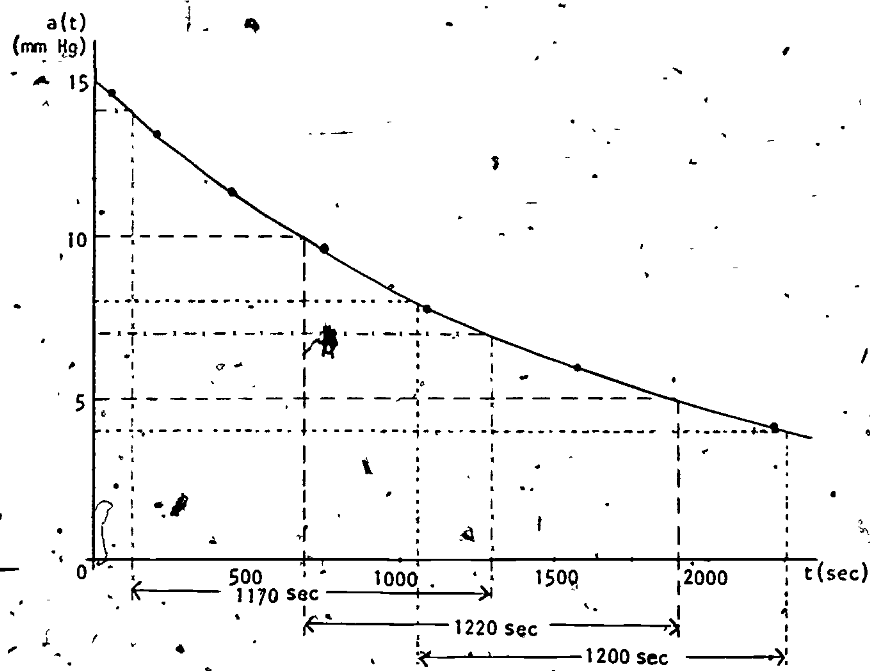


Figure 1. Decomposition of Trichloromethyl Chloroformate (from Table I(a)).

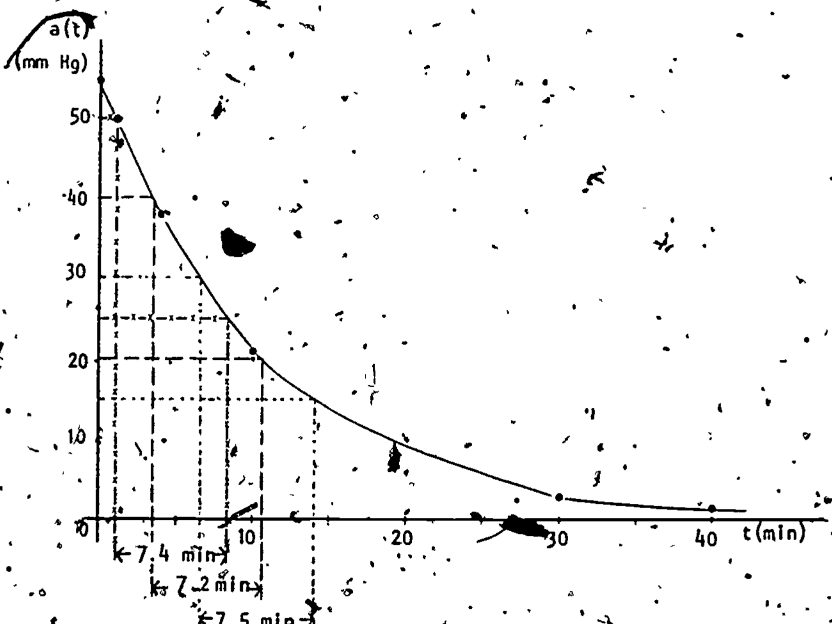


Figure 2. Decomposition of Ethylamine (from Table I(b)).

3

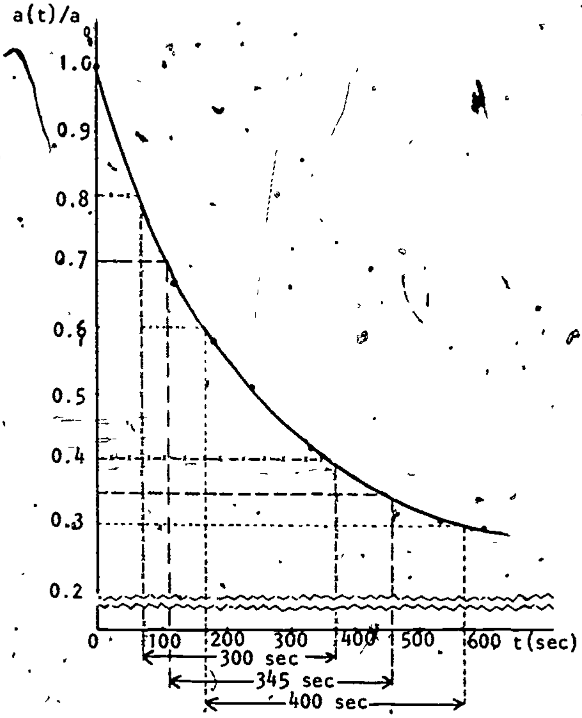


Figure 3. Alkaline Hydrolysis of Ethyl Nitrobenzoate (from Table I(c)).

that it takes approximately 1170 seconds for a(t) to decrease from 14 to 7 mm Hg, or 1220 seconds for it to decrease from 10 to 5 mm Hg. In each of the first two figures the measured time intervals are approximately equal, but in Figure 3 they are not.

1.3 Questions

Can we explain this difference in terms of the reactions? Or, turning the question around, can we draw any conclusions based on these observations, about the nature of the reactions?

1.4 Chemical Kinetics

Questions such as these are part of a branch of chemistry known as *chemical kinetics*. Chemical kinetics is concerned with the rates and mechanisms of chemical reactions. The name reflects the fact that "kinetics" is concerned with the changing aspects of systems, as distinguished from "statics" which concerns systems at *equilibrium*. We should also point out here that the rate at which a chemical process takes place and the mechanism of the process (i.e., what exactly happens during the transformation of A into $B_1 + B_2 + \ldots B_n$) are two different things. The study of reaction mechanisms lies at a higher theoretical level than the study of reaction rates. In general, experimentally determined reaction rates can be used to rule out a proposed mechanism if they are inconsistent with it. But experimental data that are consistent with a proposed mechanism can only serve as supporting evidence for it; they cannot be used directly to *prove* its correctness.

2. REACTION ORDER

2.1 Definitions

To make the question in Section 1.3 more specific, we shall summarize some background information about the reaction rates in reactions of this type. If substance A (in gas or liquid form) is uniformly distributed, and if the temperature and volume are kept constant, then it usually turns out that the *rate* $a'(t)$ at which A decomposes is proportional to a non-negative integer power (0,1,2,...) of the concentration $a(t)$. In other words

(1) $$a'(t) = -k[a(t)]^n$$

where k is a positive constant and n is a non-negative integer. We call k the *rate constant* and n the *order* of the reaction. Equation (1) with n established is called the *rate law* for the reaction.

We shall consider reaction orders 0, 1 and 2 in detail. Higher reaction orders for reactions of the type we are discussing are considerably more rare.

2.2 Zero-order Reactions

Setting n = 0 in Equation (1) gives

(2) $$a'(t) = -k_0$$

where we have introduced the subscript to denote the reaction order. The rate is independent of the concentration of A. It is determined by other factors such as temperature, the intensity of light in light-induced reactions, the surface area available in surface-catalyzed reactions, or the amount of catalyst in homogeneous catalysis. (A catalyst is a chemical substance that controls the rate of a reaction without undergoing any net change in itself over the course of the reaction.)

2.3 First-order Reactions

In this case, we have

(3) $$a'(t) = -k_I a(t).$$

Most simple decomposition reactions involving a single reactant are of first-order. This is not surprising if we imagine the reaction process to consist of molecules of A decomposing randomly. If, for example, each molecule has 1 chance in 10 of decomposing in the next second, then about $\frac{1}{10}$th of those present will in fact decompose in that second. In other words, the *change* in a(t) *in that second* is about $-\frac{1}{10}$ a(t). We describe this by writing

$$a'(t) = -\frac{1}{10} a(t).$$

2.4 Second-order Reactions

The rate law for second-order reactions is:

(4) $$a'(t) = -k_{II} a^2(t).$$

In general, elementary reactions which require the collision of two molecules are good candidates for this category.

2.5 Statement of the Problem

Equation (1) has been confirmed for many reactions by numerous experiments, and also explained theoretically.

We shall not get into the theoretical explanation except to say (as has already been indicated in Sections 2.2, 2.3 and 2.4) that different reaction orders are the result of different underlying reaction mechanisms. So if we have a reaction and want to know more about its mechanism a very useful first step is to determine its reaction order experimentally.

Can we use data such as that given in Table I to determine whether a reaction has one of the orders we have discussed, and, if so, which one?

3. DETERMINING THE REACTION ORDER

3.1 Solving for a(t)

To begin with, we can use Equations (2), (3), and (4) to obtain explicit formulas for a(t) in the three cases.

(a) *Zero-order reactions:* If $a'(t) = -k_0$ then $a(t) = -k_0 t + C$ where C is a constant of integration. Using the fact that $a(0) = a_0$ we see that $C = a_0$ and

(5) $$a(t) = a_0 - k_0 t.$$

(b) *First-order reactions:* Starting with Equation (3), divide both sides by a(t) (which is never zero):

$$\frac{a'(t)}{a(t)} = -k_I$$

$$\int_0^t \frac{a'(t)}{a(t)}\,dt = -\int_0^t k_I\,dt$$

$$\ln a(t)\Big|_0^t = -k_I t\Big|_0^t$$

$$\ln a(t) = -k_I t + \ln a_0$$

(6) $$a(t) = a_0 e^{-k_I t}$$

(c) *Second-order reactions:* In Equation (4) we divide each side by $a^2(t)$, and conclude that

$$\frac{a'(t)}{a^2(t)} = -k_{II}$$

$$\int_0^t \frac{a'(t)}{a^2(t)}\,dt = -\int_0^t k_{II}\,dt$$

$$-\frac{1}{a(t)}\Big|_0^t = -k_{II}t\Big|_0^t$$

$$\frac{1}{a(t)} = k_{II}t + \frac{1}{a_0} = \frac{a_0 k_{II} t + 1}{a_0}$$

(7) $$a(t) = \frac{a_0}{a_0 k_{II} t + 1}.$$

Exercise 1

Find $a(t)$ explicitly for a third-order reaction.

Exercise 2

Assume two reactions are of first and second order respectively:

$$a'(t) = -k_I a(t)$$

$$b'(t) = -k_{II} b^2(t)$$

Assume they begin with the same amount of reactant ($a_0 = b_0$), and their initial rates are the same [$a'(0) = b'(0)$]. Prove that $a(t) < b(t)$ for all $t > 0$.

(Hint: Note that $\frac{a(t)}{b(t)} = 1$ when $t = 0$ and show that it is strictly decreasing for $t > 0$.)

3.2 The Difficulty

The rate constant (k_0, k_I, k_{II}) is of course not known, so we cannot get away with anything so naive as plugging our data into Equations (5), (6), and (7) to see which one checks out. It is true that the graphs of these equations have three distinctive "shapes", whatever the constants are (for example, Equation (5) is a straight line). So we could consider graphing our experimental data and trying to determine which "shape" curve fits it best. In this unit however, we present a method of determining the reaction order that does not depend on graphing, and which also gives us the rate constant at no extra cost.

3.3 Solving the Difficulty

The method starts with solving Equations (5), (6), and (7) for k_0, k_I, and k_{II}:

(8) $$k_0 = \frac{a_0 - a(t)}{t}, \qquad t > 0$$

(9) $$k_I = \frac{1}{t} \ln \frac{a_0}{a(t)}, \qquad t > 0$$

(10) $$k_{II} = \frac{1}{t} \left(\frac{1}{a(t)} - \frac{1}{a_0}\right), \qquad t > 0.$$

Now if, for example, the reaction order is zero, then all the data points should satisfy Equation (5) for some constant k_0. Thus whenever we substitute any data point (t, a(t)) to the right side of Equation (8) we should get more or less the same value (namely k_0). Naturally there will be small variations due to experimental error. Similar comments apply to Equation (9) if the reaction order is one, and Equation (10) if the reaction order is two.

So all we need to do is compute three rows of figures -- the right sides of Equations (8), (9), and (10) -- for our data points, and see if any row remains more or less constant. If so, that row gives us the reaction order, and its constant value is the rate constant (k_0, k_I, or k_{II}).

3.4 An Example

As an example, let's go back to part (a) of Table I. In Table II, we have repeated the data and also tabulated the right sides of Equations (8), (9), and (10). The figures in the row corresponding to Equation (9) are nearly constant ($\approx 5.8 \times 10^{-4} sec^{-1}$) while those in the other rows are not. So this reaction is apparently a first order reaction with $k_I \approx 5.8 \times 10^{-4} sec^{-1}$.

TABLE II

Calculation of Rate Constant and Reaction Order for Data of Table I(a).

t	sec	0	51	206	454	751	1132	1575	2215
a(t)	mm Hg	15.03	14.58	13.32	11.49	9.73	7.79	6.08	4.17
$\frac{a_0 - a(t)}{t} \times 10^3$	$\frac{mm\ Hg}{sec}$		8.82	8.30	7.80	7.06	6.40	5.68	4.90
$\frac{1}{t} \ln \frac{a_0}{a(t)} \times 10^4$	$\frac{1}{sec}$		5.96	5.86	5.92	5.79	5.81	5.75	5.79
$\frac{1}{t}\left(\frac{1}{a(t)} - \frac{1}{a_0}\right) \times 10^5$	$\frac{1}{mm\ Hg\ sec}$		4.03	4.15	4.52	4.83	5.46	6.22	7.82

Exercise 3

Determine the reaction order and rate constant from the data given in part (b) of Table I.

Exercise 4

Determine the reaction order and rate constant from the data given in part (c) of Table I.

ERIC Full Text Provided by ERIC

4. HALF-LIFE

4.1 Definition

The half-life $t_{1/2}$ of a certain amount of a reactant is the length of time required for exactly half of it to be used up. In other words, if the amount of reactant is a_0 at time $t = 0$, and if $a(t)$ is the amount at a later time t, then $t_{1/2}$ is the solution of the equation

$$a(t) = \frac{1}{2} a_0 .$$

In Section 1.2 we determined graphically the half-lives of various amounts of three reactants, and discovered that for two of the reactants $t_{1/2}$ did not seem to depend upon the initial amount, but for the third reactant it did. Let us see if this phenomenon can shed a little more light on the concept of reaction order.

4.2 Formulas for Half-Life

To start with, let us compute $t_{1/2}$ for each of the three reaction orders we are considering. All we need to do is set $a(t) = \frac{1}{2}a_0$ in each of Equations (5), (6), and (7) and solve for t:

(11) $$t_{1/2} = \frac{1}{2k_0} a_0 \quad \text{(Zero-order)}$$

(12) $$t_{1/2} = \frac{\ln 2}{k_I} \quad \text{(First-order)}$$

(13) $$t_{1/2} = \frac{1}{k_{II}\, a_0} \quad \text{(Second-order)}$$

Exercise 5

Find $t_{1/2}$ as a function of a_0 for a third-order reaction.

Exercise 6

We define $t_{3/4}$ as the time required for $\frac{3}{4}$ of a reactant to be used up. That is, $a(t_{3/4}) = \frac{1}{4} a_0$. Find $t_{3/4}$ as a function of a_0 for reactions of zero, first, and second-order.

ERIC Full Text Provided by ERIC

Exercise 7

Find the ratio $\frac{t_{3/4}}{t_{\frac{1}{2}}}$ for reactions of zero, first, and second-order.

Exercise 8

Table III gives $t_{\frac{1}{2}}$ and $t_{3/4}$ for three initial amounts of the reactant in the reaction

$$CH_3CHO \longrightarrow CH_4 + CO$$

acetaldehyde methane carbon monoxide

Determine if possible whether the reaction has one of the three orders discussed in this unit and, if so, which one.

TABLE III

Half-life and $3/4$-Life Data for the Reaction $CH_3CHO \rightarrow CH_4 + CO$ (Exercise 8)

a_0 (mm Hg)	420	225	184
$t_{\frac{1}{2}}$ (seconds)	385	572	665
$t_{3/4}$ (seconds)	1135	1710	1920

Exercise 9

Suppose, for *every* x between 0 and 1, we write t_x for the time required for fraction x of a reactant to be used up. (In Exercise 6 $t_{3/4}$ is an example of t_x with $x = 3/4$.) Show that in a first-order reaction t_x is independent of the initial amount no matter what x is.

4.3 Zero-order Reactions

For a zero-order reaction, half-life is *proportional to* initial amount. The greater the amount, the longer

ERIC Full Text Provided by ERIC

the half-life. To help yourself understand and remember this, think of a very large number of marbles, from which we remove, say, 10 each second ($k_0 = 10$). The more there are originally, the longer it will take to remove half of them.

4.4 First-order Reactions

For a first-order reaction, half-life is *independent of* initial amount!! To help understand and remember this, think again of a very large number of marbles. This time *remove one half of the pile in the first second,* then one half of the remaining pile in the next second, etc. ($k_I = \frac{1}{2}$). No matter how many we start with, it will take one second to remove half of them. Also, at any later stage it will take one second to remove half of what remains.

4.5 Second-order Reactions

For a second-order reaction, half-life is *proportional to* the reciprocal of the initial amount. Another way of saying this is that $a_0 t_{1/2}$ is a constant. The *more* of A there is, the *less* time it takes for one half of it to decompose! Although this may seem paradoxical we invite you to consider the fact that second-order reactions depend upon collisions of pairs of molecules. Equation (13) says that the more molecules there are, the more likely they will collide, and the faster the reaction will proceed.

Exercise 10

The following data were obtained by F. Daniels and E.H. Johnston (J. Am. Chem. Soc., 43, 53 (1921)) for the decomposition of nitrogen pentoxide (N_2O_5) in solution in carbon tetrachloride (CCl_4) at 45° C:

$$2N_2O_5 \rightarrow 2N_2O_4 + O_2.$$

ERIC Full Text Provided by ERIC

t (seconds)	0	184	319	526	867	1198	1877	2315	3144
concentration of N_2O_5 (mole/ℓ)	2.33	2.08	1.91	1.67	1.36	1.11	.72	.55	.34

Determine the reaction order and the rate constant, as well as the half-life $t_{\frac{1}{2}}$. How long would it take for 87.5% of the reactant to be used up?

The Project would like to thank Scott C. Mohr of Boston University, Andrew Jorgensen of Indiana State University at Evansville, Bernice Kastner of Montgomery College, Rockville, Maryland, Barbara Juister of Elgin Community College, Elgin, Illinois for their reviews, and all others who assisted in the production of this unit.

This unit was field-tested and/or student reviewed in preliminary form at Indiana State University at Evansville, Southern Oregon State College, Ashland, Colorado School of Mines, Golden, and Northern Illinois University, DeKalb, and has been revised on the basis of data received from these sites.

5. MODEL EXAM

1. For some reactions the reaction order is found to be fractional. Find a(t) explicitly (in terms of a_0) for a reaction with reaction order $n = ½$.

2. Define t_x as that time for which $a(t_x) = (1 - x)a_0$. Find t_x for a second order reaction. Is this t_x independent of a_0?

3. Determine the reaction order and rate constant from the following data for a hypothetical reaction.

t (seconds)	0	2	4	6	8	10
a(t) (moles/l)	10.0	3.98	2.51	1.82	1.44	1.19

15

ERIC Full Text Provided by ERIC

ANSWERS TO EXERCISES

1. $a(t) = a_0\left(\frac{1}{2a_0^2 k_{III} t + 1}\right)^{\frac{1}{2}}$

2. The information given to us is:
   1. $a'(t) = -k_I a(t)$
   2. $b'(t) = -k_{II} b^2(t)$
   3. $a(0) = b(0)$
   4. $a'(0) = b'(0)$

   To see that $\frac{a(t)}{b(t)}$ is a decreasing function of t, we show that the derivative of the quotient is negative.

$$\frac{d}{dt}\left(\frac{a(t)}{b(t)}\right) = \frac{b(t)\,a'(t) - a(t)\,b'(t)}{b^2(t)}$$

$$= \frac{b(t)(-k_I a(t)) - a(t)(-k_{II} b^2(t))}{b^2(t)}$$

(14) $$= a(t)\left(k_{II} - \frac{k_I}{b(t)}\right).$$

Now, $a'(0) = b'(0)$ means that

$$k_I a(0) = k_{II} b^2(0)$$

and $a(0) = b(0)$ means further that

$$k_I b(0) = k_{II} b^2(0)$$
$$k_I = k_{II} b(0).$$

When we substitute this value of k_I in Equation (14) we obtain

$$\frac{d}{dt}\left(\frac{a(t)}{b(t)}\right) = a(t)\left(k_{II} - \frac{k_{II} b(0)}{b(t)}\right)$$

$$= k_{II} a(t)\left(1 - \frac{b(0)}{b(t)}\right).$$

Since $b(t) < b(0)$ for $t > 0$,

$$\frac{b(0)}{b(t)} > 1$$

and

$$\left(1 - \frac{b(0)}{b(t)}\right) < 0.$$

3. Reaction order = 1

$k \approx 9.4 \times 10^{-2} \text{min}^{-1}$.

4. Reaction order = 2

$a_0 k \approx 4.1 \times 10^{-3} \text{sec}^{-1}$ or, since $a_0 = 0.05$

$k = 8.2 \times 10^{-2}\ \ell\ \text{mole}^{-1} \text{sec}^{-1}$.

5. $t_{\frac{1}{2}} = \dfrac{3}{2a_0{}^2 k_3}$.

6. Zero order: $t_{3/4} = \dfrac{3a_0}{4k_0}$

First-order: $t_{3/4} = \dfrac{1}{k_I} \ln 4$.

Second-order: $t_{3/4} = \dfrac{3}{a_0 k_{II}}$

7. Zero-order: $\frac{3}{2}$

First-order: 2

Second-order: 3

8. Second-order.

10. First order, $k \approx 6.2 \times 10^{-4}$ sec^{-1}, $t_{\frac{1}{2}} \approx 1120$ sec, $3t_{\frac{1}{2}} \approx 3360$ sec

ANSWERS TO MODEL EXAM

1. $a(t) = a_0 - \sqrt{a_0}\, kt + \dfrac{k^2t^2}{4}$

2. $t_x = \dfrac{1}{a_0 k_{II}} \left(\dfrac{x}{1 - x}\right)$; No.

3. $k = 7.5 \times 10^{-1}\ \ell\ \text{mole}^{-2} \text{sec}^{-1}$

Order = 2.