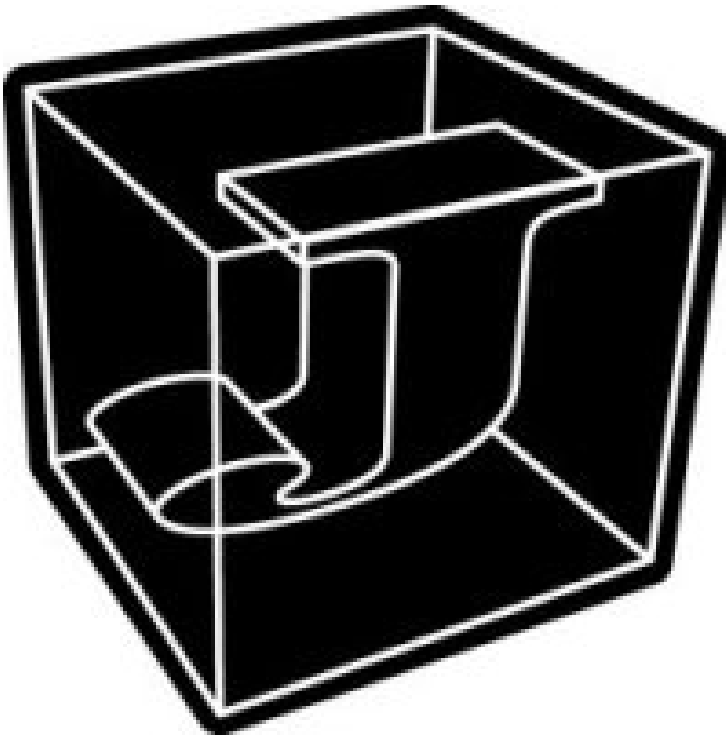


# Calculus



Kenneth E. Iverson

Copyright © 2002 Jsoftware Inc. All rights reserved.

## Preface

Calculus is at once the most important and most difficult subject encountered early by students of mathematics; introductory courses often succeed only in turning students away from mathematics, and from the many subjects in which the calculus plays a major role.

The present text introduces calculus in the informal manner adopted in my *Arithmetic* [1], a manner endorsed by Lakatos [2], and by the following words of Lanczos from his preface to [3]:

Furthermore, the author has the notion that mathematical formulas have their “secret life” behind their Golem-like appearance. To bring out the “secret life” of mathematical relations by an occasional narrative digression does not appear to him a profanation of the sacred rituals of formal analysis but merely an attempt to a more integrated way of understanding. The reader who has to struggle through a maze of “lemmas”, “corollaries”, and “theorems”, can easily get lost in formalistic details, to the detriment of the essential elements of the results obtained. By keeping his mind on the principal points he gains in depth, although he may lose in details. The loss is not serious, however, since any reader equipped with the elementary tools of algebra and calculus can easily interpolate the missing details. It is a well-known experience that the only truly enjoyable and profitable way of studying mathematics is the method of “filling in the details” by one’s own efforts.

The scope is broader than is usual in an introduction, embracing not only the *differential* and *integral* calculus, but also the *difference* calculus so useful in approximations, and the *partial derivatives* and the *fractional* calculus usually met only in advanced courses. Such breadth is achievable in small compass not only because of the adoption of informality, but also because of the executable notation employed. In particular, the array character of the notation makes possible an elementary treatment of partial derivatives in the manner used in tensor analysis.

The text is paced for a reader familiar with polynomials, matrix products, linear functions, and other notions of elementary algebra; nevertheless, full definitions of such matters are also provided.

# Table Of Contents

<b>Introduction .....</b>	<b>6</b>
A. Calculus .....	6
B. Notation and Terminology .....	10
C. Role of the Computer and of Notation .....	14
D. Derivative, Integral, and Secant Slope .....	14
E. Sums and Multiples .....	15
F. Derivatives of Powers .....	16
G. Derivatives of Polynomials .....	17
H. Power Series .....	18
I. Conclusion .....	20
<b>Differential Calculus .....</b>	<b>23</b>
A. Introduction .....	23
B. The derivative operator .....	24
C. Functions Defined by Equations (Relations) .....	24
D. Differential Equations .....	26
E. Growth $F' d.1 = F$ .....	26
F. Decay $F' d.1 = -@F$ .....	27
G. Hyperbolic Functions $F' d.2 = F$ .....	28
H. Circular Functions $F' d.2 = -@F$ .....	29
I. Scaling .....	30
J. Argument Transformations .....	31
K. Table of Derivatives .....	31
L. Use of Theorems .....	33
M. Anti-Derivative .....	34
N. Integral .....	35
<b>Vector Calculus .....</b>	<b>37</b>
A. Introduction .....	37
B. Gradient .....	38
C. Jacobian .....	40
D. Divergence And Laplacian .....	42
E. Symmetry, Skew-Symmetry, and Orthogonality .....	42
F. Curl .....	45
<b>Difference Calculus .....</b>	<b>47</b>
A. Introduction .....	47
B. Secant Slope Conjunctions .....	47
C. Polynomials and Powers .....	48
D. Stope Functions .....	50

E. Slope of the Stope .....	51
F. Stope Polynomials.....	52
G. Coefficient Transformations.....	53
H. Slopes as Linear Functions .....	54
<b>Fractional Calculus .....</b>	<b>59</b>
A. Introduction .....	59
B. Table of Semi-Differintegrals .....	61
<b>Properties of Functions .....</b>	<b>65</b>
A. Introduction .....	65
B. Experimentation.....	67
C. Proofs.....	70
D. The Exponential Family .....	70
E. Logarithm and Power.....	71
F. Trigonometric Functions .....	73
G. Dot and Cross Products .....	77
H. Normals .....	79
<b>Interpretations and Applications .....</b>	<b>83</b>
A. Introduction .....	83
B. Applications and Word Problems .....	84
C. Extrema and Inflection Points.....	85
D. Newton's Method.....	87
E. Kerner's Method.....	89
F. Determinant and Permanent .....	90
G. Matrix Inverse.....	92
H. Linear Functions and Operators .....	92
I. Linear Differential Equations.....	94
J. Differential Geometry .....	95
K. Approximate Integrals .....	97
L. Areas and Volumes .....	101
M. Physical Experiments .....	103
<b>Analysis.....</b>	<b>107</b>
A. Introduction .....	107
B. Limits .....	108
C. Continuity .....	111
D. Convergence of Series .....	111
<b>Appendix .....</b>	<b>117</b>
A. Polynomials .....	117
B. Binomial Coefficients .....	119
C. Complex Numbers .....	119
D. Circular and Hyperbolic Functions.....	120

E. Matrix Product and Linear Functions .....	120
F. Inverse, Reciprocal, And Parity .....	121
<b>Index .....</b>	<b>126</b>

# Chapter 1

## Introduction

### A. Calculus

Calculus is based on the notion of studying any phenomenon (such as the position of a falling body) together with its *rate of change*, or *velocity*. This simple notion provides insight into a host of familiar things: the growth of trees or financial investments (whose rates of change are proportional to themselves); the vibration of a pendulum or piano string; the shape of the cables in a powerline or suspension bridge; and the logarithmic scale used in music.

In spite of the simplicity and ubiquity of its underlying notion, the calculus has long proven difficult to teach, largely because of the difficult notion of *limits*. We will defer this difficulty by first confining attention to the polynomials familiar from high-school algebra.

We begin with a concrete experiment of dropping a stone from a height of twenty feet, and noting that both the position and the velocity (rate of change of position) appear to depend upon (are *functions* of) the elapsed time. However, because of the rapidity of the process, we are unable to observe either with any precision.

More precise observation can be provided by recording the fall with a video camera, playing it back one frame at a time, and recording the successive positions in a vertical line on paper. A clearer picture of the motion can be obtained by moving the successive points to a succession of equally spaced vertical lines to obtain a *graph* or *plot* of the position against elapsed time.

The position of the falling stone can be described approximately by an algebraic expression as follows:

$$p(t) = 20 - 16 * t * t$$

We will use this definition in a computer system (discussed in Section B) to compute a table of times and corresponding heights, and then to plot the points detailed in the table. The computer expressions may be followed by comments (in Roman font) that are not executed:

```
i.11                               First eleven integers, beginning at zero  
0 1 2 3 4 5 6 7 8 9 10
```

```
t=:0.1*i.11                         Times from 0 to 1 at intervals of one-tenth
```

```
h=:20-16*t*t                       Corresponding heights
```

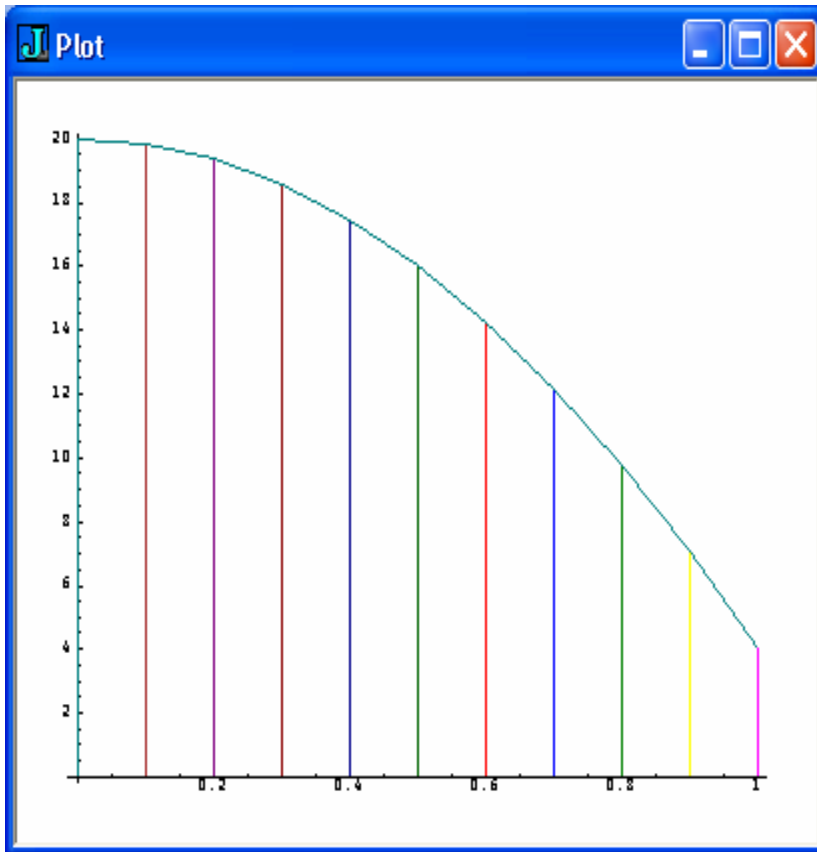
```

t, .h
0    20
0.1 19.84
0.2 19.36
0.3 18.56
0.4 17.44
0.5  16
0.6 14.24
0.7 12.16
0.8  9.76
0.9  7.04
1    4

load 'plot'

PLOT=: 'stick,line' &plot
PLOT t;h

```



The plot gives a graphic view of the velocity (rate of change of position) as the slopes of the lines between successive points, and emphasizes the fact that it is rapidly increasing in magnitude. Moreover, the table provides the information necessary to compute the average velocity between any pair of points.

For example, the last two rows appear as:

```

0.9  7.04
1    4

```

and subtraction of the first of them from the last gives both the change in time (the elapsed time) and the corresponding change in position:

```

1 4 - 0.9 7.04

```

0.1 3.04

Finally, the change in position divided by the change in time gives the average velocity:

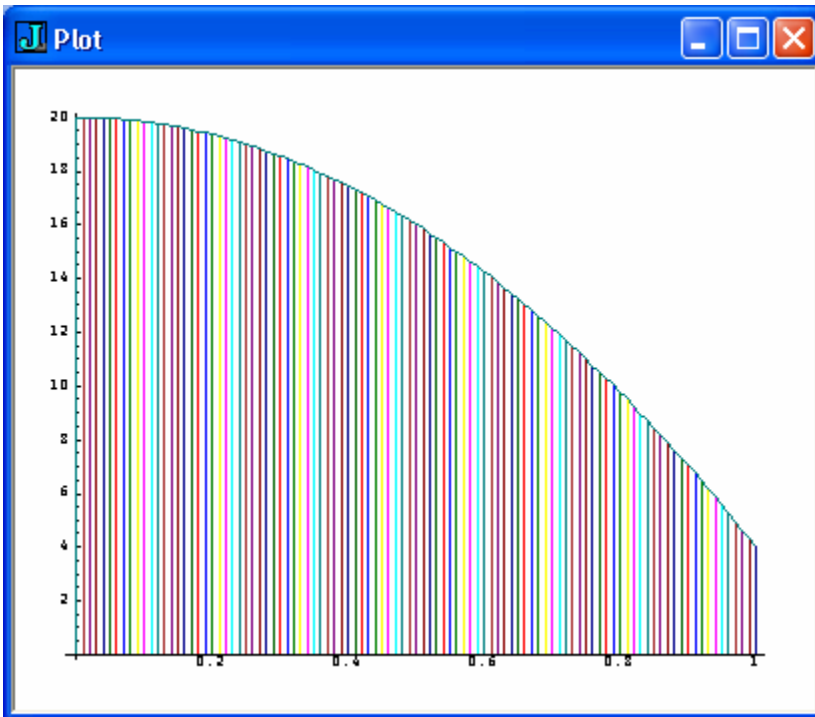
$$\frac{-3.04}{0.1} \quad \text{Division is denoted by } \% \\ \text{The } - \text{ denotes a negative number}$$

The negative value of this velocity indicates that the velocity is in a downward direction.

Both the table and the plot suggest abrupt changes in velocity, but smaller intervals between points will give a truer picture of the actual continuous motion:

$$\begin{aligned} \mathbf{t} &:= 0.01 * \mathbf{i} . 101 && \text{Intervals of one-hundredth over the same range} \\ \mathbf{h} &:= 20 - 16 * \mathbf{t} * \mathbf{t} \end{aligned}$$

PLOT t;h



This plot suggests that the actual (rather than the average) rate of change at any point is given by the slope of the tangent (touching line) to the curve of the graph. In terms of the table, it suggests the use of an interval of zero.

But this would lead to the meaningless division of a zero change in position by a zero change in time, and we are led to the idea of the "limit" of the ratio as the interval "approaches" zero.

For many functions this limit is difficult to determine, but we will avoid the problem by confining attention to polynomial functions, where it can be determined by simple algebra.

The velocity (rate of change of position) is also a function of  $t$  and, because it is *derived* from the function  $p$ , it is called the *derivative* of  $p$ . It also can be expressed algebraically as follows:  $v(t) = -32*t$ .

Moreover, since the velocity is also a function of  $t$ , it has a derivative (the *acceleration*) which is also called the *second* derivative of the original function  $p$ .



Various notations (with various advantages) have been used for the derivative:

newton	$\dot{p}$	$\ddot{p}$	
leibniz	$dy/dt$	$d^2y/d^2t$	$d^ny/d^n t$ ( $y = p(t)$ )
modern	$p'$	$p''$	$p^n$
heaviside ( <b>J</b> )	$p \mathbf{D}.1$	$p \mathbf{D}.2$	$p \mathbf{D}.n$

Heaviside also introduced the notion of  $\mathbf{D}$  as a *derivative operator*, an entity that applies to a function to produce another function. This is a new notion not known in elementary algebra.

In the foregoing we have seen that calculus requires three notions that will not have been met by most students of high school algebra:

1. The notion of the rate of change of a function.
2. The notion of an *operator* that applies to a function to produce a function.
3. The notion of a *limit* of an expression that depends upon a parameter whose limiting value leads to an indeterminate expression such as  $0/0$ .

Although the notion of an operator that produces a function is not difficult in itself, its first introduction as the *derivative* operator (that is, jointly with another new notion of rate of change) makes it more difficult to embrace. We will therefore begin with the use of simpler (and eminently useful) operators before even broaching the notion of rate of change.

A further obstacle to the teaching of calculus (common to other branches of mathematics as well) is the absence of *working models* of mathematical ideas, models that allow a student to gain familiarity through concrete and accurate experimentation. Such working models are provided automatically by the adoption of mathematical notation that is also executable on a computer.

In teaching mathematics, the necessary notation is normally introduced in context and in passing, with little or no discussion of notation as such. Notation learned in a simple context is often expanded without explicit comment. For example, although the significance of a fractional power may require discussion, the notations  $x^{1/2}$  and  $x^{m/n}$  and  $x^{p^2}$  used for it may be silently adapted from the more restricted integer cases  $x^2$  and  $x^n$ .

Although an executable notation must differ somewhat from conventional notation (if only to resolve conflicts and ambiguities), it is important that it be introducible in a similarly casual manner, so as not to distract from the mathematical ideas it is being used to convey. The subsequent section illustrates such use of the executable notation **J** (available free from website [jsoftware.com](http://jsoftware.com)) in introducing and using vectors and operators.

## B. Notation and Terminology

The terminology used in J is drawn more from English than from mathematics:

- a) Functions such as  $+$  and  $*$  and  $^$  are also referred to as *verbs* (because they *act* upon *nouns* such as **3** and **4**), and operators such as  $/$  and  $\&$  are accordingly called *adverbs* and *conjunctions*, respectively.
- b) The symbol  $=:$  used in assigning a name to a referent is called a *copula*, and the names **credits** and **sum** used in the sentences **credits=: 24.5 17 38** and **sum=:+ /** are referred to as *pronouns* and *proverbs* (pronounced with a long o), respectively.
- c) Vectors and matrices are also referred to by the more suggestive terms *lists* and *tables*.

Because the notation is executable, the computer can be used to explore and elucidate topics with a clarity that can only be appreciated from direct experience of its use. The reader is therefore urged to use the computer to do the exercises provided for each section, as well as other experiments that may suggest themselves.

To avoid distractions from the central topic of the calculus, we will assume a knowledge of some topics from elementary math (discussed in an appendix), and will introduce the necessary notation with a minimum of comment, assuming that the reader can grasp the meaning of new notation from context, from simple experiments on the computer, from the on-line Dictionary, or from the study of more elementary texts such as *Arithmetic* [1]. The remainder of this section is a computer dialog (annotated by comments in a different font) that introduces the main characteristics of the notation.

The reader is urged to try the following sentences (and variants of them) on the computer:

<b>3.45+6.78+0.01</b> 10.24	Plus
<b>2*3</b> 6	Times
<b>2^3</b> 8	Power (product of three twos)
<b>1 2 3 * 4 5 6</b> 4 10 18	Lists or vectors
<b>2 &lt; 3 2 1</b> false) 1 0 0 2 <. 3 2 1 2 2 1	Less than (1 denotes true, and 0 denotes false) Lesser of (Minimum) Related spellings denote related verbs
<b>(+: , -: , *: , %:) 16</b> 32 8 256 4	Double, halve, square, square root
<b>+ / 4 5 6</b> 15	The symbol / denotes the adverb <i>insert</i>

```
4+5+6
15
```

```
*/4 5 6
120
```

```
3-5
_2
  -5
_5
```

Verbs are ambivalent, with a meaning that depends on context; the symbol - denotes subtraction or negation according to context

```
2^1 2 3
2 4 8
^1 2 3
2.71828 7.38906 20.0855
```

The power function

The exponential function

```
*/4 5 6
120
```

A derived verb produced by an adverb is also ambivalent; the

```
1 2 3 */ 4 5 6
4 5 6
8 10 12
12 15 18
```

dyadic case of \*/ produces a multiplication table

```
a=: 1 2 3
b=: 4 5 6 7
powertable=: ^/
c=: a powertable b
c
1 1 1 1
16 32 64 128
81 243 729 2187
```

The copula (=:) can be used to assign names to nouns, verbs, adverbs, and conjunctions

```
+/ c
98 276 794 2316
```

Adds together items (rows) of the table c

```
+/"1 c
4 240 3240
```

The rank conjunction " applies its argument (here the function +/) to each rank-1 cell (*list*)

```
3"1 c
3 3 3
```

The constant function 3 applied to each list of c

```
3"1 b
3
```

The constant function 3 applied to the list b

```
3"0 b
3 3 3
```

The constant function 3 applied to each atom of b

```
x=: 4
1+x*(3+x*(3+x*(1)))
125
1+x*3+x*3+x*1
125
(3*4)+5
17
3*4+5
27
```

Parentheses provide punctuation as in high-school algebra. However, there is no precedence or hierarchy among verbs; each applies to the result of the entire phrase to its right

```

tithe=: %&10
tithe 35
3.5
log=: 10&^.
log 10 20 100
1 1.30103 2

sin=: 1&o.
sin 0 1 1r2p1
0 0.841471 1

x=:1 2 3 4
^&3 x
1 8 27 64

```

The conjunction `&` bonds a dyad to a noun; result is a corresponding function of one argument (a *monad*)

Sine (of radian arguments)  
Sine of 0, 1, and one-half pi

Cube of x

We will write informal proofs by writing a sequence of sentences to imply that each is equivalent to its predecessor, and that the last is therefore equivalent to the first. For example, to show that the sum of the first  $n$  odd numbers is the square of  $n$ , we begin with:

```

] odds=: 1+2*i.n=: 8
1 3 5 7 9 11 13 15
|.odds
15 13 11 9 7 5 3 1

odds + |.odds
16 16 16 16 16 16 16 16

n#n
8 8 8 8 8 8 8 8

```

The identity function `]` causes display of result

and then write the following sequence of equivalent sentences:

```

+/.odds
+/.|.odds
-:(+/.odds) + (+/.|.odds)
-:+/ (odds+|.odds)
+/-:(odds+|.odds)
+/.n#n
n*n
*:n

```

### Exercises

Solutions or hints appear in bold brackets. Make serious attempts before consulting them.

- B1 To gain familiarity with the keyboard and the use of the computer, enter some of the sentences of this section and verify that they produce the results shown in the text. Do not enter any of the comments that appear to the right of the sentences.
- B2 To test your understanding of the notions illustrated by the sentences of this section, enter variants of them, but try to predict the results before pressing the Enter key.
- B3 Enter `p=: 2 3 5 7 11` and predict the results of `+/p` and `*/p`; then review the discussion of parentheses and predict the results of `-/p` and `%/p`.

B4 Enter `i. 5` and `#p` and `i.#p` and `i.-#p`. Then state the meanings of the primitives `#` and `i.`.

B5 Enter `asp=: p * _1 ^ i. # p` to get a list of primes that alternate in sign (enter `asp` alone to display them). Compare the results of `-/p` and `+/asp` and state in English the significance of the phrase `-/`.

[ `-/` yields the alternating sum of a list argument]

B6 Explore the assertion that `%/a` is the alternating product of the list `a`.

[ Use `arp=: p^_1^i.#p` ]

B7 Execute (by entering on the computer) each of the sentences of the informal proof preceding these exercises to test the equivalences. Then annotate the sentences to state why each is equivalent to its predecessor (and thus provide a formal proof).

B8 Experiment with, and comment upon, the following and similar sentences:

```
s=: '4%5'
```

```
|.s
```

```
do=: "
```

```
do s
```

```
do |.s
```

```
|.i.5
```

```
|. 'I saw'
```

[ Enclosing quotes produce a list of characters that may be manipulated like other lists and may, if they represent proper sentences, be *executed* by applying the verb `".` ]

B9 Experiment with and comment upon:

```
]a=: <1 2 3
```

```
>a
```

```
2*a
```

```
2*>a
```

```
]b=: (<1 2 3) , (<'pqrs')
```

```
|.b
```

```
#b
```

```
1 2 3; 'pqrs'
```

[ `<` boxes its argument to produce a *scalar encoding*; `>` opens it.]

B10 Experiment with and comment upon:

```
power=: ^
```

```
with=: &
```

```
cube=: ^ with 3
```

```
cube 1 2 3 4
```

```
1 8 27 64
```

```
cube
```

```
^&3
```

[ Entering the name of a function alone shows its definition in *linear* form;

the *foreign* conjunction ! : provides other forms]

B11 Press the key F1 (in the top row) to display the J vocabulary, and click the mouse on any item (such as -) to display its definition.

### C. Role of the Computer and of Notation

Seeing the computer determine the derivatives of functions such as the square might well cause a student to forget the mathematics and concentrate instead on the wonder of how the computer does it. A student of astronomy might likewise be diverted by the wonders of optics and telescopes; they are respectable, but they are *not* astronomy.

In the case of the derivative operator, the computer simply consults a given table of derivatives and an associated table of rules (such as the *chain* rule). The details of the computer calculation of the square root of **3.14159** are much more challenging.

The important point for a student of mathematics is to treat the computer as a *tool*, being clear about *what* it does, not necessarily *how* it does it. In particular, the tool should be used for convenient and accurate experimentation with mathematical ideas.

The study of notation itself can be fascinating, but the student of calculus should concentrate on the mathematical ideas it is being used to convey, and not spend too much time on byways suggested by the notation. For example, a chance application of the simple factorial function to a fraction (! 0.5) or the square root to a negative number (% :-4) might lead one away into the marvels of the gamma function and imaginary numbers.

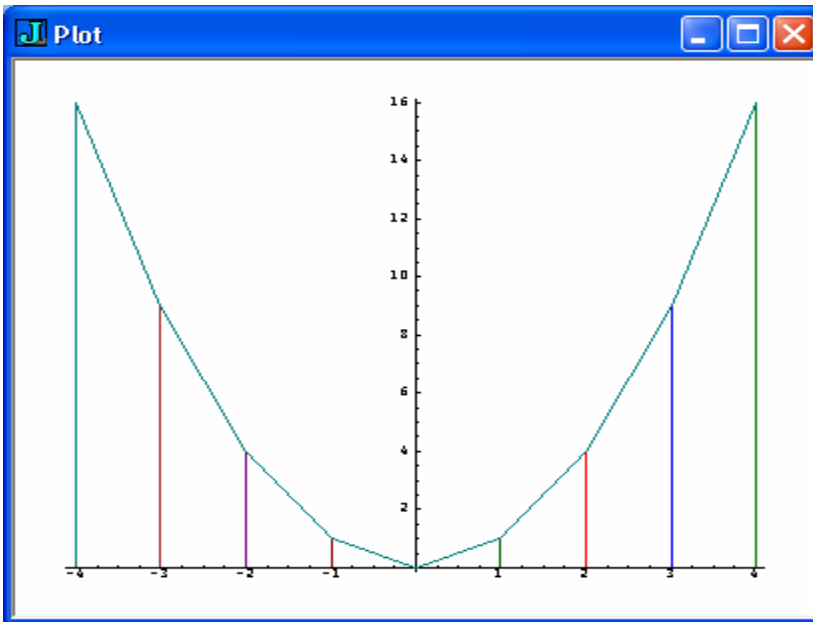
A student must, of course, learn *some* notation, such as the use of ^ for *power* (first used by de Morgan) and of + and \* for *plus* and *times*. However, it is best not to spend too much conscious effort on memorizing vocabulary, but rather to rely on the fact that most words will be used frequently enough in context to fix them in mind. Moreover, the definition of a function may be displayed by simply entering its name *without* the usual accompanying argument, as illustrated in Exercise B10.

### D. Derivative, Integral, and Secant Slope

The central notions of the calculus are the *derivative* and the *integral* or *anti-derivative*. Each is an *adverb* in the sense that it applies to a function (or *verb*) to produce a *derived* function. Both are illustrated (for the square function  $x^2$ ) by the following graph, in which the slope of the tangent at the point  $x, x^2$  as a function of  $x$  is the derivative of the square function, that is  $2x$ . The area under the graph is the integral of the square, that is, the function  $x^3/3$ , a function whose derivative is the square function.

Certain important properties of a function are easily seen in its graph. For example, the *square* has a *minimum* at the point **0 0**; increases to the right of zero at an accelerating rate; and the area under it can be estimated by summing the areas of the trapezoids:

```
PLOT x;*: x=:i:4
```



These properties concern the *local* behavior of a function in the sense that they concern how rapidly the function value is changing at any point. They are not easily discerned from the expression for the function itself, but *are* expressed directly by its derivative.

More surprisingly, a host of important functions can be defined simply in terms of their derivatives. For example, the important *exponential* (or *growth*) function is completely defined by the fact that it is equal to its derivative (therefore growing at a rate equal to itself), and has the value 1 for the argument 0.

The *difference* calculus (Chapter 4) is based upon *secant slopes*, such as illustrated by the lines in the foregoing plot of the square function. The slope of the *secant* (from *ligne secante*, or *cutting line*) through the points  $x, f(x)$  and  $(x+r), (f(x+r))$  is obtained by dividing the *rise*  $(f(x+r) - f(x))$  by the *run*  $r$ ; the result of  $((f(x+r) - f(x)) / r)$  is called the *r-slope* of  $f$  at the point  $x$ .

The difference calculus proves useful in a wide variety of applications, including approximations to arbitrary functions, and financial calculations in which events (such as payments) occur at fixed intervals.

The function used to plot the square must be prepared as follows:

```
load 'graph plot'

PLOT='stick,line' &plot
```

## E. Sums and Multiples

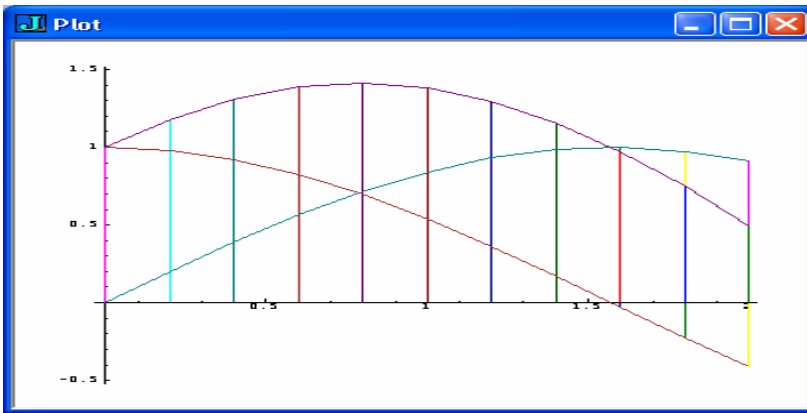
The derivative of the function  $p+q$  (the sum of the functions  $p$  and  $q$ ) is the sum of their derivatives. This may be seen by plotting the functions together with their sum. We will illustrate this by the sine and cosine functions:

```
p=:1&o.          The sine function

q=:2&o.          The cosine function

x=: (i.11)%5

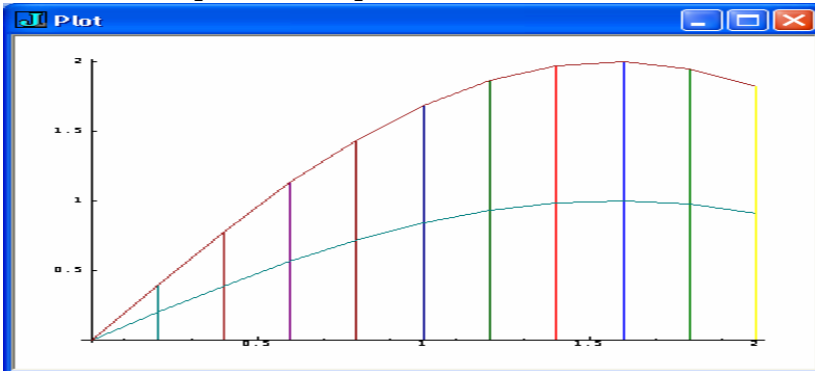
PLOT x;>(p x) ; (q x) ; ((p x)+(q x))
```



Since each value of the sum function is the sum of the component functions, the slopes of its secants are also the sum of the corresponding slopes. Since this is true for every secant, it is true for the derivative.

Similarly, the slopes of a multiple of a function  $p$  are all the same multiple of the slopes of  $p$ , and its derivative is therefore the same multiple of the derivative of  $p$ . For example:

```
PLOT x;>(p x);(2 * p x)
```



## F. Derivatives of Powers

The derivative of the square function  $f = x^2$  can be obtained by algebraically expanding the expression  $f(x+r)$  to the equivalent form  $(x^2) + (2*x*r) + (r^2)$ , as shown in the following *proof*, or list of identical expressions:

$$\begin{aligned} & ((f(x+r) - (f(x))) / r \\ & (((x+r)^2) - (x^2)) / r \\ & ((x^2) + (2*x*r) + (r^2)) - (x^2) / r \\ & ((2*x*r) + (r^2)) / r \\ & (2*x) + r \end{aligned}$$

Moreover, if  $r$  is set to zero in the final expression  $(2*x) + r$ , the result is  $2*x$ , the value of the derivative of  $x^2$ .

Similar analysis can be performed on other power functions. Thus if  $g = x^3$ :

$$\begin{aligned} & ((g(x+r) - (g(x))) / r \\ & ((3*(x^2)*r) + (3*x*r^2) + (r^3)) / r \\ & (3*x^2) + (3*x*r) + (r^2) \end{aligned}$$

Again the derivative is obtained by setting  $r$  to zero, leaving  $3*x^2$ .



Similar analysis shows that the derivative of  $x^4$  is  $4x^3$  and, in general, the derivative of  $x^n$  is  $n x^{n-1}$ . Since the first term of the expansion of  $(x+r)^n$  is cancelled by the subtraction of  $x^n$ , and since all terms after the second include powers of  $x$  greater than 1, the only term relevant to the derivative is the second, that is,  $n x^{n-1}$ .

## G. Derivatives of Polynomials

The expression  $(8x^0) + (-20x^1) + (-3x^2) + (2x^3)$  is an example of a polynomial. We may also express it as  $\text{poly}(x)$ , using the polynomial function denoted by `poly`. The elements of the list `[8 -20 -3 2]` are called the *coefficients* of the polynomial. For example:

```
x=:2
(8*x^0)+(-20*x^1)+(-3*x^2)+(2*x^3)
_28
8 -20 -3 2 p. x
_28
c=:8 -20 -3 2
x=:0 1 2 3 4 5
(8*x^0) + (-20*x^1) + (-3*x^2) + (2*x^3)
8 _13 _28 _25 8 83
c p. x
8 _13 _28 _25 8 83
```

The expression  $(8x^0) + (-20x^1) + (-3x^2) + (2x^3)$  is a sum whose derivative is therefore a sum of the derivatives of the individual terms. Each term is a multiple of a power, so each of these derivatives is a multiple of the derivative of the corresponding power. The derivative is therefore the sum:

$$(0 \cdot 8) + (-20 \cdot 1 \cdot x^0) + (-3 \cdot 2 \cdot x^1) + (2 \cdot 3 \cdot x^2)$$

This is a polynomial with coefficients given by `c*i.#c`, with the leading element removed to reduce each of the powers by 1:

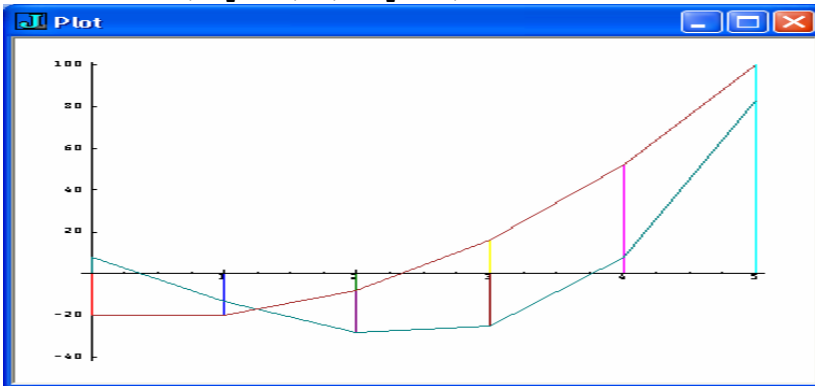
```
c
8 -20 -3 2
i.#c
0 1 2 3
c*i.#c
0 _20 _6 6
dc=:}.c*i.#c
dc
_20 _6 6
dc p. x
_20 _20 _8 16 52 100
x,.(c p. x),.(dc p. x)
0 8 _20
1 _13 _20
```

```

2  _28  _8
3  _25  16
4  _ 8  52
5  83 100

```

```
PLOT x;>(c p. x);(dc p. x)
```

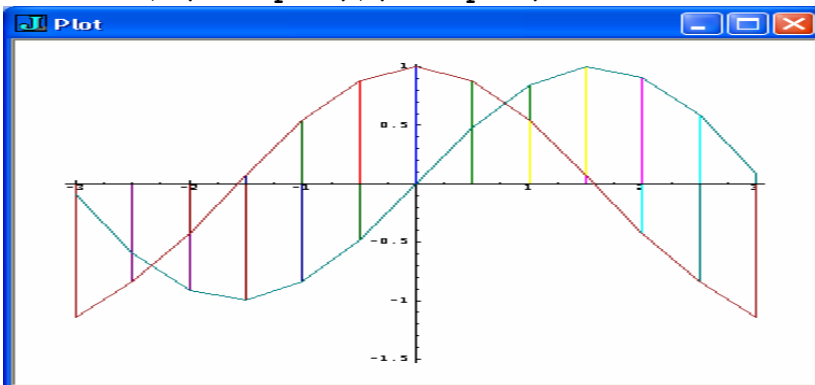


As remarked in Section A, " ... the functions of interest in elementary calculus are easily approximated by polynomials ... ". The following illustrates this for the sine function and its derivative (the cosine), using `_1r6` for the rational fraction negative one-sixths:

```

csin=:0 1 0 _1r6 0 1r120 0 _1r5040
ccos=:).csin*i.#csin
x=(i:6)%2
PLOT x;>(csin p. x);(ccos p. x)

```



## H. Power Series

We will call `s` a *series* function if `s n` produces a list of `n` elements. For example:

```

s1=: $&0 1    Press F1 for the vocabulary, and see the definition of $
s2=: _1&^@s1
s1 5
0 1 0 1 0
s2 8
1 _1 1 _1 1 _1 1 _1

```

A polynomial with coefficients produced by a series function is a sum of powers weighted by the series, and is called a *power series*. For example:

```

x=:0.5*i.6
(s1 5) p. x    Sum of odd powers
0 0.625 2 4.875 10 18.125

```

```

(s2 8) p. x      Alternating sum of powers
1 0.664063 0 _9.85156 _85 _435.68

```

We will define an adverb `PS` such that `n (s PS) x` gives the  $n$ -term power series determined by the series function `s`:

```

PS=:1 : (';' ' (u. x.) p. y. ') See definition of : (Explicit definition)

```

```

5 s1 PS x
0 0.625 2 4.875 10 18.125

```

```

8 s2 PS x
1 0.664063 0 _9.85156 _85 _435.68

```

```

S1=:s1 PS
5 S1 x
0 0.625 2 4.875 10 18.125

```

Power series can be used to approximate the functions needed in elementary calculus. For example:

```

s3=:%@!@i.      Reciprocal of factorial of integers
s4=:$&0 1 0 _1
s5=:s3*s4

s3 7
1 1 0.5 0.166667 0.0416667 0.00833333 0.00138889

s4 7
0 1 0 _1 0 1 0

s5 7
0 1 0 _0.166667 0 0.00833333 0

```

```

S3=:s3 PS
S4=:s4 PS
S5=:s5 PS

7 S3 x      Seven-term power series approximation to
1 1.64872 2.71806 4.47754 7.35556 12.0097

```

```

^x          the exponential function
1 1.64872 2.71828 4.48169 7.38906 12.1825

```

```

10 S5 x      Ten-term power series approximation to
0 0.479426 0.841471 0.997497 0.909347 0.599046

```

```

1&o. x      the sine function
0 0.479426 0.841471 0.997495 0.909297 0.598472

```

Since `c=:s5 10` provides the coefficients of an approximation to the sine function, the expression `. c * i.10` provides (according to the preceding section) the coefficients of an approximation to its derivative (the cosine). Thus:

```

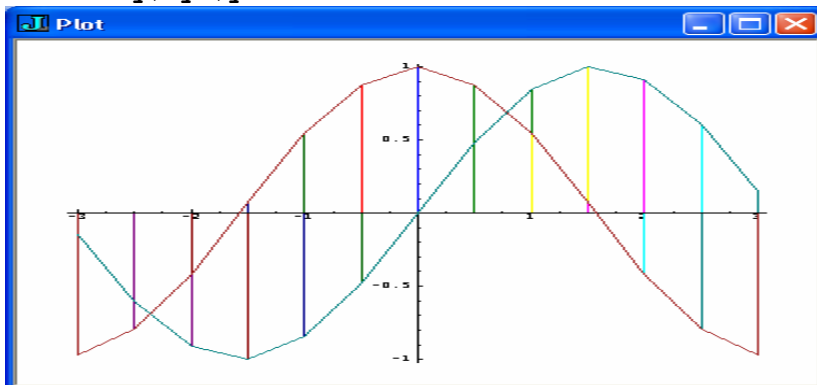
c=:s5 10

```

```
y1=:c p. y=:0.5*i:6
```

```
y2=:({).c*i.10) p. y
```

```
PLOT y;>y1;y2
```



## I. Conclusion

We conclude with a brief statement of the ways in which the present treatment of the calculus differs from most introductory treatments. For the differential calculus of Chapter 2, the important difference is the avoidance of problems of limits by restricting attention to polynomials, and the use of power series to extend results to other functions.

Moreover:

1. In Vector Calculus (Chapter 3), *Partial* derivatives are treated in a simpler and more general way made possible by the use of functions that deal with arguments and results of arbitrary rank; this in contrast to the restriction to *scalars* (single elements) common in elementary treatments of the calculus.
2. The *Calculus of Differences* (Chapter 4) is developed as a topic of interest in its own right rather than as a brief way-station to integrals and derivatives.
3. *Fractional* derivatives (Chapter 5) constitute a powerful tool that is seldom treated in calculus courses. They are an extension of derivatives of integral order, introduced here in a manner analogous to the extension of the power function to fractional exponents, and the extension of the *factorial* and *binomial coefficient* functions to fractional arguments.
4. Few formal proofs are presented, and proofs are instead treated (as they are in *Arithmetic* [1]) in the spirit of Lakatos in his *Proofs and Refutations* [2], of which the author says:

"Its modest aim is to elaborate the point that informal, quasi-empirical, mathematics does not grow through the monotonous increase of the number of indubitably established theorems but through the incessant improvement of guesses by speculation and criticism, by the logic of proofs and refutations."

5. The notation used is unambiguous and executable. Because it is executable, it is used for experimentation; new notions are first introduced by leading the student to see them in action, and to gain familiarity with their use before analysis is attempted.
6. As illustrated at the end of Section B, informal proofs will be presented by writing a sequence of expressions to imply that each is equivalent to its predecessor, and that the last is therefore equivalent to the first.

7. The exercises are an integral part of the development, and should be attempted as early as possible, perhaps even before reading the relevant sections. Try to provide (or at least sketch out) answers without using the computer, and then use it to confirm your results.
8. Two significant parts may be distinguished in treatments of the calculus:
  - a) A *body* comprising the central notions of *derivative* and *anti-derivative* (integral), together with their important consequences.
  - b) A *basis* comprising the analysis of the notion of *limit* (that arises in the transition from the secant slope to the tangent slope) needed as a foundation for an axiomatic deductive treatment.

The common approach is to treat the basis first, and the body second. For example, in Johnson and Kiokemeister *Calculus with analytic geometry* [6], the section on *The derivative of a function* occurs after eighty pages of preliminaries.

The present text defers discussion of the analytical basis to Chapter 8, first providing the reader with experience with the derivative and the importance of its fruits, so that she may better appreciate the point of the analysis.



# Chapter 2

## Differential Calculus

### A. Introduction

In Chapter 1 it was remarked that:

- The power of the calculus rests upon the study of functions together with their *derivatives*, or *rates-of-change*.
- The difficult notion of *limits* encountered in determining derivatives can be deferred by restricting attention to functions expressible as polynomials.
- The results for polynomials can be extended to other functions by the use of power series.
- The derivative of the polynomial  $c_i x^i$  is the polynomial  $d_i x^{i-1}$ , where  $d_i = i \cdot c_i$ .

We begin by defining a function `deco` for the derivative coefficients, and applying it repeatedly to a list of coefficients that represent the cube (third power):

```

deco=:}.@([ * i.@#)

c=:0 0 0 1
x=:0 1 2 3 4 5 6
c p. x
0 1 8 27 64 125 216
x^3
0 1 8 27 64 125 216

]cd=:deco c           Coefficients of first derivative of cube
0 0 3
cd p. x
0 3 12 27 48 75 108
3*x^2
0 3 12 27 48 75 108
#cd                   Number of elements
3

]cdd=:deco cd        Coefficients of second derivative of cube
0 6
cdd p. x
0 6 12 18 24 30 36
2*3*x^1
0 6 12 18 24 30 36

```

```

#cdd      Number of elements
2

]cddd=:deco cdd      Coefficients of third derivative of cube
6
  cddd p. x          A constant function
6 6 6 6 6 6
  1*2*3*x^0
6 6 6 6 6 6
#cddd      Number of elements
1

]cdddd=:deco cddd    Coefficients of fourth derivative of cube (empty list)

  cdddd p. x         Sum of an empty list (a zero constant function)
0 0 0 0 0 0
#cdddd      Number of elements
0

```

## B. The derivative operator

If  $f=:c&p.$  is a polynomial function, then  $g=:(\text{deco } c)\&p.$  is its derivative. For example:

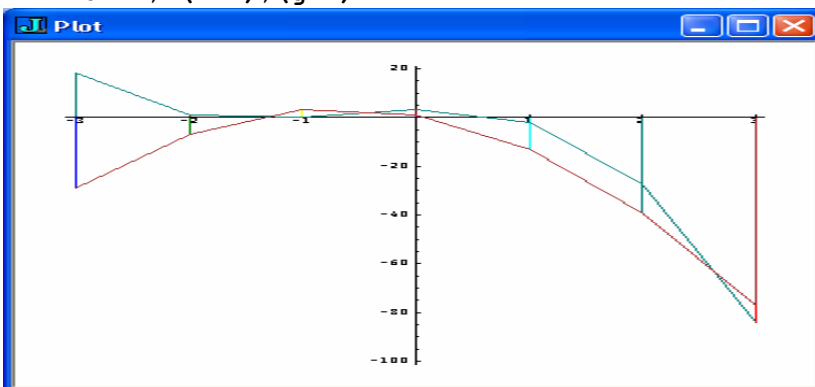
```

c=:3 1 _4 _2
f=:c&p.
g=:(\text{deco } c)\&p.
]x=:i:3
_3 _2 _1 0 1 2 3

f x
18 1 0 3 _2 _27 _84
g x
_29 _7 3 1 _13 _39 _77

PLOT x;>(f x);(g x)

```



Since `deco` provides the computations for obtaining the derivative of  $f$  in terms of its defining coefficients, it can also provide the basis for a derivative operator that applies directly to the function  $f$ . For example:

```

f d. 1 x
_29 _7 3 1 _13 _39 _77

```

In the expression  $f \text{ d. } 1$ , the right argument determines the order of the derivative, in this case giving the *first* derivative. Successive derivatives can be obtained as follows:



```

f d. 2 x
28 16 4 _8 _20 _32 _44

```

```

(deco deco c) p. x
28 16 4 _8 _20 _32 _44

```

```

f d. 3 x
_12 _12 _12 _12 _12 _12 _12
(deco deco deco c) p. x
_12 _12 _12 _12 _12 _12 _12

```

## C. Functions Defined by Equations (Relations)

A function may be defined directly, as in  $f = x^3$  or  $g = 0 \ 0 \ 0 \ 1 \ p$ . It may also be defined *indirectly* by an equation that specifies some relation that it must satisfy. For example:

1. **invcube** is the inverse of the cube.

A function that satisfies this equation may be expressed directly in various ways. For example:

```

cube =: ^&3
cube x =: 1 2 3 4 5
1 8 27 64 125

```

```

invcube =: ^&(%3)

```

```

invcube cube x
1 2 3 4 5

```

```

cube invcube x
1 2 3 4 5

```

```

altinvcube =: cube ^: _1          Inverse operator
altinvcube cube x
1 2 3 4 5

```

2. **reccube** is the reciprocal of the cube.

```

reccube =: %@cube
reccube x
1 0.125 0.037037 0.015625 0.008
(reccube * cube) x
1 1 1 1 1

```

3. The derivative of **s** is the cube.

```

s =: 0 0 0 0 0.25&p.
s x
0.25 4 20.25 64 156.25
s d.1 x
1 8 27 64 125

```

A stated relation may not specify a function completely. For example, the equation for Example 3 is also satisfied by the alternative function  $as = 8^0 + s$ . Thus:

```
as=:8"0 + s

as x
8.25 12 28.25 72 164.25
as d.1 x
1 8 27 64 125
```

Further conditions may therefore be stipulated to define the function completely. For example, if it is further required that  $s^2$  must be 7, then  $s$  is completely defined. Thus:

```
as=:3"0 + s
as 2
7
as d.1 x
1 8 27 64 125
```

C1 Experiment with the expressions of this section.

## D. Differential Equations

An equation that involves derivatives of the function being defined is called a *differential equation*. The remainder of this chapter will use simple differential equations to define an important collection of functions, including the *exponential*, *hyperbolic*, and *circular* (or *trigonometric*).

We will approach the solution of differential equations through the use of polynomials. Because a polynomial includes one more term than its derivative, it can never exactly equal the derivative, and we consider functions that approximate the desired solution. However, for the cases considered, successive coefficients decrease rapidly in magnitude, and approximation can be made as close as desired. Consideration of the convergence of such approximations is deferred to Chapter 8.

## E. Growth $f' = f$

If the derivative of a function is equal to (or proportional to) the function itself, it is said to grow *exponentially*. Examples of exponential growth include continuous compound interest, and the growth of a well-fed colony of bacteria.

If  $f$  is the polynomial  $c \cdot x^n$ , then the derivative of  $f$  is the polynomial with coefficients  $n \cdot c \cdot x^{n-1}$ . Thus:

```
]c=:1, (%1), (%1*2), (%1*2*3), (%!4), (%!5), (%!6)
1 1 0.5 0.166667 0.0416667 0.00833333 0.00138889
c*i.#c
0 1 1 0.5 0.166667 0.0416667 0.00833333
}. c*i.#c
1 1 0.5 0.166667 0.0416667 0.00833333
deco c
1 1 0.5 0.166667 0.0416667 0.00833333
```

In this case the coefficients of the derivative polynomial agree with the original coefficients except for the missing final element. The same is true for any coefficients produced by the following *exponential coefficients* function:

```

ec=: %@!

]c=: ec i. n=: 7
1 1 0.5 0.166667 0.0416667 0.00833333 0.00138889
  deco c
1 1 0.5 0.166667 0.0416667 0.00833333

```

Consequently, the function  $c&p.$  is approximately equal to its derivative. For example:

```

c&p. x=: 0 1 2 3
1 2.71806 7.35556 19.4125

(deco c)&p. x
1 2.71667 7.26667 18.4

```

The primitive *exponential* function, denoted by  $\wedge$ , is defined as the limiting case for large  $n$ . For example:

```

c=: ec i. n=: 12
c&p. x
1 2.71828 7.38905 20.0841
  ^x
1 2.71828 7.38906 20.0855

```

The related function  $\wedge@(r&*)$  grows at a rate *proportional* to the function, the ratio being  $r$ . For example:

```

r=:0.1
q=: ^@(r&*)

q d.1 x
0.1 0.110517 0.12214 0.134986
  r * q x
0.1 0.110517 0.12214 0.134986

```

## F. Decay $F' d.1 = -@F$

A function whose derivative is equal to or proportional to its negation is *decaying* at a rate proportional to itself. Interpretations include the charge of water remaining in a can punctured at the bottom, and the electrical charge remaining in a capacitor draining through a resistor; the rate of flow (and therefore of loss) is proportional to the pressure provided by the remaining charge at any time.

The coefficients of a polynomial defining such a function must be similar to that for growth, except that the elements must alternate in sign. Thus:

```

eca=: _1&^ * ec
eca i.7
1 _1 0.5 _0.166667 0.0416667 _0.00833333 0.00138889

  deco eca 7
_1 1 _0.5 0.166667 _0.0416667 0.00833333

```

```
(eca 20) &p. x
1 0.367879 0.135335 0.0497871 0.0183153
```

```
(deco eca 20) &p. x
_1 _0.367879 _0.135335 _0.0497871 _0.0183175
```

The relation between the growth and decay functions will be explored in exercises and in Chapter 6.

F1 Define a function `pp` such that `(a pp b) &p.` is equivalent to the product `(a&p.*b&p.)`; test it for `a=:1 2 1` [ `b=:1 3 3 1`.

```
[ pp=: +//.@(*) ]
```

F2 Predict the value of a few elements of `(ec pp eca) i.7` and enter the expression to validate your prediction.

F3 Enter `x=:0.1*i:30` and `y1=:^ x` and `y2=:^-x`. Then enter `PLOT x;>y1;y2`.

F4 Predict and confirm the result of the product `y1*y2`.

## G. Hyperbolic Functions $F d. 2 = F$

The second derivative of a function may be construed as its *acceleration*, and many phenomena are described by functions defined in terms of their acceleration.

We will again use polynomials to approximate functions, first a function that is equal to its second derivative. Since the *second* derivative of the exponential  $e^x$  is also equal to itself, the coefficients `ec i.n` would suffice. However, we seek new functions and therefore add the restriction that `f d.1` must *not* equal `f`.

Coefficients satisfying these requirements can be obtained by suppressing (that is, replacing by zeros) alternate elements of `ec i.n`. Thus:

```
2|i.n=: 9
0 1 0 1 0 1 0 1 0
```

```
hsc=: 2&| * ec
]c=: hsc i.n
0 1 0 0.166667 0 0.00833333 0 0.000198413 0
```

```
deco c
1 0 0.5 0 0.0416667 0 0.00138889 0
```

```
deco deco c
0 1 0 0.166667 0 0.00833333 0
```

The result of `deco c` was shown above to make clear that the *first* derivative differs from the function. However, it should also be apparent that *it* qualifies as a second function that equals its second derivative. We therefore define a corresponding function `hcc`:

```
hcc=: 0&@(2&|) * ec
hcc i.n
1 0 0.5 0 0.0416667 0 0.00138889 0 2.48016e_5
```

```

deco deco hcc i.n
1 0 0.5 0 0.0416667 0 0.00138889

```

The limiting values of the corresponding polynomials are called the *hyperbolic sine* and *hyperbolic cosine*, respectively. They are the functions defined by `hsin=: 5&o.` and `hcos=: 6&o.` Thus:

```

hsin=:5&o.
hcos=:6&o.

(hsc i.20)&p. x=: 0 1 2 3 4
0 1.1752 3.62686 10.0179 27.2899

```

```

hsin x
0 1.1752 3.62686 10.0179 27.2899

```

```

(hcc i.20)&p. x
1 1.54308 3.7622 10.0677 27.3082

```

```

hcos x
1 1.54308 3.7622 10.0677 27.3082

```

It should also be noted that each of the hyperbolic functions is the derivative of the other. Further properties of these functions will be explored in Chapter 6. In particular, it will be seen that a plot of one against the other yields a hyperbola. The more pronounceable abbreviations *cosh* and *sinh* (pronounced *cinch*) are also used for these functions.

G1 Enter `x=:0.1*i:30` and `y1=:hsin x` and `y2=:hcos x`. Then plot the two functions by entering `PLOT x;>y1;y2`.

G2 Enter `PLOT y1;y2` to plot *cosh* against *sinh*, and comment on the shape of the plot.

G3 Predict the result of  $(y_2^2 - y_1^2)$  and test it on the computer.

## H. Circular Functions $F d.2 = -@F$

It may be noted that the hyperbolics, like the exponential, continue to grow with increasing arguments. This is not surprising, since their acceleration increases with the increase of the function.

We now consider functions whose acceleration is *opposite* in sign to the functions themselves, a characteristic that leads to *periodic* functions, whose values repeat as arguments grow. These functions are useful in describing periodic phenomena such as the oscillations in a mechanical system (the motion of a weight suspended on a spring) or in an electrical system (a coil connected to a capacitor).

Appropriate polynomial coefficients are easily obtained by alternating the signs of the non-zero elements resulting from `hsc` and `hcc`. Thus:

```

sc=: _1&^@(3&=)@(4&|) * hsc
cc=: _1&^@(2&=)@(4&|) * hcc
sc i.n
0 1 0 _0.166667 0 0.00833333 0 _0.000198413 0
cc i.n

```

```
1 0 _0.5 0 0.0416667 0 _0.00138889 0 2.48016e_5
```

```
(sc i.20)&p. x
0 0.841471 0.909297 0.14112 _0.756803
sin=:1&o.
cos=:2&o.
```

```
sin x
0 0.841471 0.909297 0.14112 _0.756802
```

```
(cc i.20)&p. x
1 0.540302 _0.416147 _0.989992 _0.653644
```

```
cos x
1 0.540302 _0.416147 _0.989992 _0.653644
```

It may be surprising that these functions defined only in terms of their derivatives are precisely the *sine* and *cosine* functions of trigonometry (expressed in terms of arguments in *radians* rather than *degrees*); these relations are examined in Section 6F.

H1 Repeat Exercises G1-G3 with modifications appropriate to the circular functions.

H2 Use the "power series" operator **PS** and other ideas from Section 1G in experiments on the hyperbolic and circular functions.

## I. Scaling

The function  $\wedge@(\mathbf{x}\&*)$  used in Section B is an example of *scaling*; its argument is first multiplied by the *scale factor*  $\mathbf{x}$  before applying the main function  $\wedge$ . Such scaling is generally useful, and we define a more convenient conjunction for the purpose as follows:

```
AM=: 2 : 'x. @ (y.&*)'      Atop Multiplication
```

For example:

```
^&(0.1&*) x=: 0 1 2 3 4
1 1.10517 1.2214 1.34986 1.49182
```

```
^ AM 0.1 x
1 1.10517 1.2214 1.34986 1.49182
```

Thus,  $\mathbf{f} \text{ AM } \mathbf{x}$  may be read as " $\mathbf{f}$  atop multiplication (by)  $\mathbf{x}$ ". Also:

```
^ AM 0.1 d.1 x
0.1 0.110517 0.12214 0.134986 0.149182
0.1 * ^ AM 0.1 x
0.1 0.110517 0.12214 0.134986 0.149182
```

## J. Argument Transformations

Scaling is only one of many useful argument transformations; we define two further conjunctions, *atop addition* and *atop polynomial*:

```
AA=: 2 : 'x. @ (y.&+)'
AP=: 2 : 'x. @ (y.&p.)'
```

In Section H it was remarked that the circular functions `sin` and `cos` "repeat" their values after a certain period. Thus:

```
per=: 6.28
cos x
1 0.540302 _0.416147 _0.989992 _0.653644

cos AA per x
0.999995 0.54298 _0.413248 _0.989538 _0.656051
```

Experimentation with different values of `per` can be used to determine a better approximation to the true period of the cosine.

The conjunction `AP` provides a more general transformation. Thus:

```
f AA 3 AM 4 is f AP 3 4
f AM 3 AA 4 is f AP 12 3
```

A function `FfC` to yield Fahrenheit from Celsius can be used to further illustrate the use of argument transformation:

```
FfC=: 32"0 + 1.8"0 * ] Uses Constant functions (See Section 1B)

fahr=: _40 0 100
FfC fahr
_40 32 212

] AA 32 AM 1.8 fahr
_40 32 212

] AP 32 1.8 fahr
_40 32 212
```

The following derivatives are easily obtained by substitution and the use of the table of Section K:

Function	Derivative
<code>f AA r</code>	<code>f D AA r</code>
<code>f AM r</code>	<code>(f D AM r * r"0)</code>
<code>f AP c</code>	<code>(f D AP c * (d c) &amp;p.)</code>

## K. Table of Derivatives

The following table lists a number of important functions, together with their derivatives. Each function is accompanied by a phrase (such as *Identity*) and an index that will be used to refer to it, as in *Theorem 2* or  $\theta 2$  (where  $\theta$  is the Greek letter *theta*).

$\theta$	NAME	FUNCTION	DERIVATIVE
1	Constant function	$a^0$	$0^0$
2	Identity	$] $	$1^0$
3	Constant Times	$a^0 * ]$	$a^0$
4	Sum	$f+g$	$(f \text{ d.1})+(g \text{ d.1})$
5	Difference	$f-g$	$(f \text{ d.1})-(g \text{ d.1})$
6	Product	$f*g$	$(f*(g \text{ d.1}))+((f \text{ d.1})*g)$
7	Quotient	$f\%g$	$(f\%g)*((f \text{ d.1})\%f)-((g \text{ d.1})\%g)$
8	Composition	$f@g$	$(f \text{ d.1})@g * (g \text{ d.1})$
9	Inverse	$f \text{ INV}$	$\%@(f \text{ d.1 } @(f \text{ INV}))$
10	Reciprocal	$\%@f$	$-\@(f \text{ d.1 } \% (f*f))$
11	Power	$\wedge n$	$n\&p. * \wedge(n-1)$
12	Polynomial	$c\&p.$	$(\text{deco } c)\&p.$

**Legend:**

Functions  $f$  and  $g$  and constants  $a$  and  $n$ , and list constant  $c$

Polynomial derivative  $\text{deco}=: ] . @ ( ] * i . @\#)$

Inverse adverb  $\text{INV}=: \wedge :\_1$

Although more thorough analysis will be deferred to Chapter 8, we will here present arguments for the plausibility of the theorems:

- $\theta 1$  Since  $a^0 x$  is  $a$  for any  $x$ , the rise is the *zero* function  $0^0$ .
- $\theta 2$  Since  $(] a+x) - (] x)$  is  $(a+x) - x$ , the rise is  $a$ , and the slope is  $a\%a$
- $\theta 3$  Multiplying a function by  $a$  multiplies all of its rises, and hence its slopes, by  $a$  as well.
- $\theta 4,5$  The rise of  $f+g$  (or  $f-g$ ) is the sum (or difference) of the rises of  $f$  and  $g$ . Also see the discussion in Section 1D.
- $\theta 6$  If the result of  $f$  is fixed while the result of  $g$  changes, the result of  $f*g$  changes by  $f$  times the change in  $g$ ; conversely if  $f$  changes while  $g$  is fixed. The total change in  $f*g$  is the sum of these changes.
- $\theta 7$  If  $h=: f\%g$ , then  $g*h$  is  $f$ , and, using  $\theta 6$ :
- $$f \text{ d.1}$$
- $$(g*h) \text{ d.1}$$
- $$(g*(h \text{ d.1}))+((g \text{ d.1})*h)$$
- The equation  $(f \text{ d.1})=(g*(h \text{ d.1}))+((g \text{ d.1})*h)$  can be solved for  $h \text{ d.1}$ , giving the result of  $\theta 7$ .
- $\theta 8$  The derivative of  $f@g$  is the derivative of  $f$  "applied at the point  $g$ " (that is,  $(f \text{ d.1})@g$ ), multiplied by the rate of change of the function that is applied first (that is,  $g \text{ d.1}$ )



- $\theta 9$   $f \circ (f^{-1}) \cdot 1$  is the product  $(f \cdot 1) \circ (f^{-1}) \cdot 1$  (from  $\theta 6$ ). But since  $f \circ (f^{-1})$  is the identity function, its derivative is  $1 \cdot p$ . and the second factor  $(f^{-1}) \cdot 1$  is therefore the reciprocal of the first.
- $\theta 10$  This can be obtained from  $\theta 7$  using the case  $f = 1$ .
- $\theta 11$  Since  $x^5$  is equivalent to the product function  $1 \cdot x^4$ , its derivative may be obtained from  $\theta 6$  and the result for the derivative of  $x^4$ . Further cases may be obtained similarly; that is, by *induction*.
- $\theta 12$  This follows from  $\theta 3$  and  $\theta 11$ .
- K1** Enter  $f = x^2$  and  $f = x^3$  and  $x = 1 \ 2 \ 3 \ 4$ ; then test the equivalence of the functions in the discussion of Theorem 7 by entering each followed by  $x$ , being sure to parenthesize the entire sentence if need be.
- K2** If  $a$  is a noun (such as 2.7), then  $a''0$  is a constant function. Prove that  $((a''0 + f) \cdot 1 = f \cdot 1)$  is a tautology, that is, gives 1 (true) for every argument.

## L. Use of Theorems

The product of the identity function ( $1$ ) with itself is the square ( $x^2$  or  $x \cdot x$ ), and the expression for the derivative of a product can therefore be used as an alternative determination of the derivative of the square and of higher powers:

$$(1 \cdot 1) \cdot 1$$

$$(1 \cdot (1 \cdot 1)) + ((1 \cdot 1) \cdot 1) \quad \text{Theorem 6}$$

$$(1 \cdot 1''0) + (1''0 \cdot 1) \quad \text{Theorem 2}$$

$$1 + 1$$

$$2''0 \cdot 1 \quad \text{Twice the argument}$$

Further powers may be expressed as products with the identity function. Thus:

$$f4 = 1 \cdot f3 = 1 \cdot f2 = 1 \cdot f1 = 1 \cdot f0 = 1''0$$

$$x = 0 \ 1 \ 2 \ 3 \ 4$$

$$>(f0;f1;f2;f3;f4) \ x$$

$$1 \ 1 \ 1 \ 1 \ 1$$

$$0 \ 1 \ 2 \ 3 \ 4$$

$$0 \ 1 \ 4 \ 9 \ 16$$

$$0 \ 1 \ 8 \ 27 \ 64$$

$$0 \ 1 \ 16 \ 81 \ 256$$

Their derivatives can be analyzed in the manner used for the square:

$$f3 \cdot 1$$

$$(1 \cdot f2) \cdot 1$$

$$(((1 \cdot 1) \cdot f2) + (1 \cdot (f2 \cdot 1)))$$

$$((1''0 \cdot f2) + (1 \cdot 2''0 \cdot 1))$$

```
(f2+2"0 * f2)
```

```
(3"0 * f2)
```

## M. Anti-Derivative

The anti-derivative is an operator defined by a relation: applied to a function  $f$ , it produces a function whose derivative is  $f$ . Simple algebra can be applied to produce a function **adeco** that is inverse to **deco**.

Since **deco** multiplies by indices and then drops the leading element, the inverse must divide by one plus the indices, and then append an arbitrary leading element. We will try two different leading elements, and then define **adeco** as a *dyadic* function whose left argument specifies the arbitrary element (known as the *constant of integration*):

```
f1=: 5"1 , ] % >:@i.@#[] Constant of integration is 5
```

```
c=: 3 1 4 2
```

```
f1 c
5 3 0.5 1.33333 0.5
deco f1 c
3 1 4 2
```

```
f2=: 24"1 , ] % >:@i.@#[] Constant of integration is 24
```

```
f2 c
```

```
24 3 0.5 1.33333 0.5
deco f2 c
3 1 4 2
```

```
adeco=: [ , ] % >:@i.@#[] Constant specified by left argument
```

```
4 adeco c
```

```
4 3 0.5 1.33333 0.5
deco 4 adeco c
3 1 4 2
```

```
zadeco=: 0&adeco
```

Monadic for common case of zero

```
zadeco c
```

```
0 3 0.5 1.33333 0.5
deco zadeco c
3 1 4 2
```

Just as **deco** provides a basis for the derivative operator **d.**, so does **adeco** provide the basis for extending **d.** to the anti-derivative, using negative arguments. For example:

```
x=: i.6
```

```
f=: c&p.
```

```
f x
```

```
3 10 37 96 199 358
```

```
f d. _1 x
```

```
0 5.33333 26.6667 90 233.333 506.667
(0 adeco c) p. x
0 5.33333 26.6667 90 233.333 506.667
```

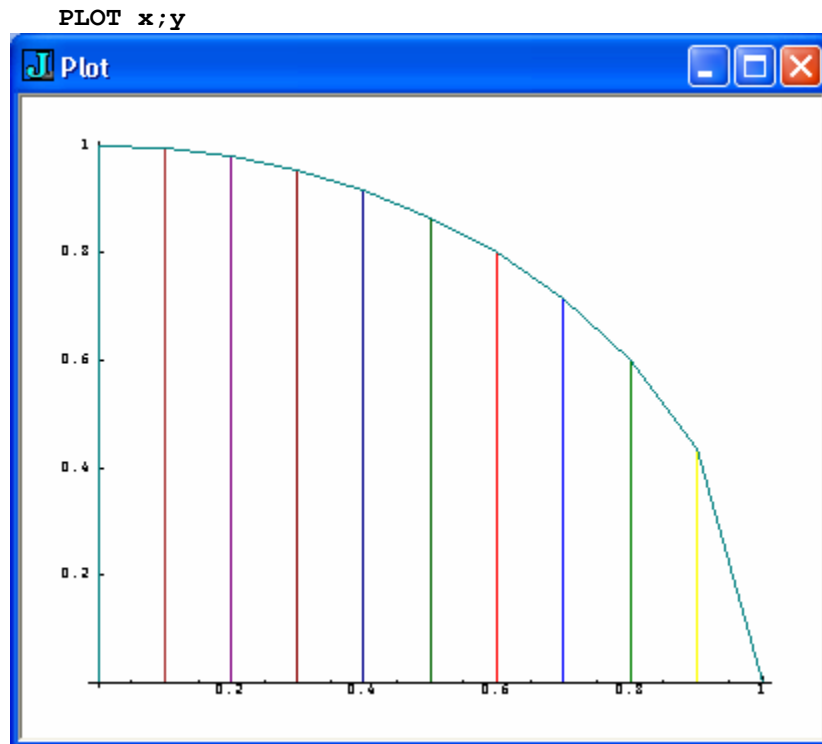
## N. Integral

The area under (bounded by) the graph of a function has many important interpretations and uses. For example, if `circle=: %: @ (1"0 - *:)`, then `circle x` gives the `y` coordinate of a point on a circle with radius 1. The first quadrant may then be plotted as follows:

```
circle=: %: @ (1"0 - *:)
```

    Square root of 1 minus the square

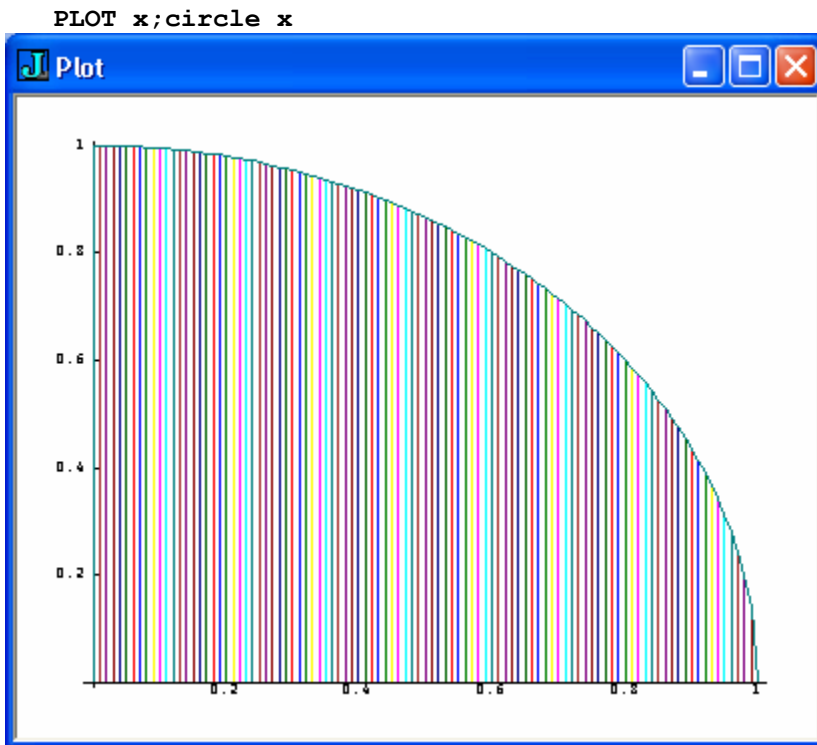
```
x=:0.1*i.11
y=:circle x
x,y
0      1
0.1 0.994987
0.2 0.979796
0.3 0.953939
0.4 0.916515
0.5 0.866025
0.6      0.8
0.7 0.714143
0.8      0.6
0.9 0.43589
1      0
```



The approximate area of the quadrant is given by the sum of the ten trapezoids, and (using `r=:0.1`) its *change* from `x` to `x+r` is `r` times the average height of the trapezoid, that is, the average of `circle x`, and `circle x+r`. Therefore, its rate-of change (derivative) at any argument value `x` is approximately the corresponding value of the circle function.

As the increment `r` approaches zero, the rate of change approaches the exact function value, as illustrated below for the value `r=:0.01`:

```
x=:0.01*i.101
```



In other words, the area under the curve is given by the anti-derivative.

# Chapter 3

## Vector Calculus

### A. Introduction

Applied to a list of three dimensions (length, width, height) of a box, the function `vol=:/` gives its volume. For example:

```
lwh=:4 3 2
vol=:/
vol lwh
24
```

Since `vol` is a function of a *vector*, or *list* (rank-1 array), the rank-0 derivative operator `d.` used in the differential calculus in Chapter 1 does not apply to it. But the derivative operator `D.` does apply, as illustrated below:

```
vol D.1 lwh
6 8 12
```

The last element of this result is the rate of change as the last element of the argument (height) changes or, as we say, the *derivative with respect to the last element* of the vector argument. Geometrically, this rate of change is the area given by the other two dimensions, that is, the length and width (whose product `12` is the area of the base).

Similarly, the other two elements of the result are the derivatives with respect to each of the further elements; for example, the second is the product of the length and height. The entire result is called the *gradient* of the function `vol`.

The function `vol` produces a rank-0 (called *scalar*, or *atomic*) result from a rank-1 (vector) argument, and is therefore said to have *form 0 1* or to *be a 0 1* function; its derivative produces a rank-1 result from a rank-1 argument, and has form `1 1`.

The product over the first two elements of `lwh` gives the "volume in two dimensions" (that is, the *area* of the base), and the product over the first element alone is the "volume in one dimension". All are given by the function `VOLS` as follows:

```
VOLS=:vol\
VOLS lwh
4 12 24
```

The function `VOLS` has form `1 1`, and *its* derivative has form `2 1`. For example:

**VOLS D.1 lwh**  
 1 0 0  
 3 4 0  
 6 8 12

This table merits attention. The last row is the gradient of the product over the entire argument, and therefore agrees with gradient of **vol** shown earlier. The second row is the gradient of the product over the first two elements (the base); its value does not depend at all on the height, and the derivative with respect to the height is therefore zero (as shown by the last element).

Strictly speaking, vector calculus concerns only functions of the forms **0 1** and **1 1**; other forms tend to be referred to as *tensor analysis*. Since the analysis remains the same for other forms, we will not restrict attention to the forms **0 1** and **1 1**. However, we will normally restrict attention to three-space (as in **vol 2 3 4** for the volume of a box) or two-space (as in **vol 3 4** for the area of a rectangle), although an arbitrary number of elements may be treated.

Because the result of a **1 1** function is a suitable argument for another of the same form, a sequence of them can be applied. We therefore reserve the term *vector function* for **1 1** functions, even though **0 1** and **2 1** functions are also vector functions in a more permissive sense.

We adopt the convention that a name ending in the digits **r** and **a** denotes an **r, a** function. For example, **F01** is a scalar function of a vector, **ABC11** is a vector function of a vector, and **G02** is a scalar function of a matrix (such as the determinant **det=:**  $-/$   $.$   $*$ ). The functions **vol** and **VOLS** might therefore be renamed **vol01** and **VOLS11**.

Although the function **vol** was completely defined by the expression **vol=:**  $*$  / our initial comments added the physical *interpretation* of the volume of a box of dimensions **lwh**. Such an interpretation can be exceedingly helpful in understanding the function and its rate of change, but it can also be harmful: to anyone familiar with finance and fearful of geometry, it might be better to use the interpretation **cost=:**  $*$  / applied to the argument **cip** (**c** crates of **i** items each, at the price **p**).

We will mainly allow the student to provide her own interpretation from some familiar topic, but will devote a separate Chapter (7) to the matter of interpretations. Chapter 7 may well be consulted at any point.

## B. Gradient

As illustrated above for the vector function **VOLS**, its first derivative produces a matrix result called the complete derivative or *gradient*. We will now use the conjunction **D.** to define an adverb **GRAD** for this purpose:

**GRAD=:D.1**  
**VOLS GRAD lwh**  
 1 0 0  
 3 4 0  
 6 8 12

We will illustrate its application to a number of functions:

<b>E01=:</b> $+/@:*$	Sum of squares
<b>F01=:</b> $\%:@E01$	Square root of sum of squares
<b>G01=:</b> $4p1"1 * *:@F01$	Four pi times square of <b>F01</b>

```

H01=: %@G01
p=: 1 2 3

(E01,F01,G01,H01) p
14 3.74166 175.929 0.00568411

E01 GRAD p
2 4 6
F01 GRAD p
0.267261 0.534522 0.801784

G01 GRAD p
25.1327 50.2655 75.3982

H01 GRAD p
_0.000812015 _0.00162403 _0.00243604

```

B1 Develop interpretations for each of the functions defined above.

ANSWERS:

**E01**  $p$  is the square of the distance (from the origin) to a point  $p$ .

**F01**  $p$  is the distance to a point  $p$ , or the radius of the sphere (with centre at the origin) through the point  $p$ .

**G01**  $p$  is the surface area of the sphere through the point  $p$ .

**H01** is the intensity of illumination at point  $p$  provided by a unit light source at the origin.

B2 Without using **GRAD**, provide definitions of functions equivalent to the derivatives of each of the functions defined above.

ANSWERS:

```

E11=: +:"1
F11=: -:@%@%:@E01 * E11
G11=: 4p1"0 * E11
H11=: -@%@*:@G01 * G11

```

Three important results (called the *Jacobian*, *Divergence*, and *Laplacian*) are obtained from the gradient by applying two elementary matrix functions. They are the *determinant*, familiar from high-school algebra, and the simpler but less familiar *trace*, defined as the sum of the diagonal. Thus:

```

det=:+/ . *
trace=:+/@((<0 1)&|:)

VOLS GRAD lwh
1 0 0
3 4 0
6 8 12

det VOLS GRAD lwh
48

trace VOLS GRAD lwh
17

```

We will also have occasion to use the corresponding adverbs `det@` and `trace@`. Thus:

```
DET=:det@
VOLS GRAD DET lwh
48
```

```
TRACE=:trace@
VOLS GRAD TRACE lwh
17
```

## C. Jacobian

The Jacobian is defined as the determinant of the gradient. Thus:

```
JAC=: GRAD DET

VOLS lwh
4 12 24

VOLS GRAD lwh
1 0 0
3 4 0
6 8 12

VOLS JAC lwh
48
```

The Jacobian may be interpreted as the *volume derivative*, or rate of change of volume produced by application of a function. This interpretation is most easily appreciated in the case of a linear function. We will begin with a linear function in 2-space, in which case the "volume" of a body is actually the area:

```
mp=: +/ . *      Matrix Product
]m=: 2 2$2 0 0 3
2 0
0 3

L11=: mp&m"1
]fig1=:>1 1;1 0;0 0;0 1
1 1
1 0
0 0
0 1

]fig2=: L11 fig1
2 3
2 0
0 0
0 3

L11 JAC 1 1
6
L11 JAC 1 0
6

L11 JAC fig1
6 6 6 6
```



The result of the Jacobian is indeed the ratio of the areas of **fig1** and **fig2**, as may be verified by plotting the two figures by hand. Moreover, for a linear function, the value of the Jacobian is the same at every point.

C1 Provide an interpretation for the function `K11 := (H11*) "1`.

[ The result of **K11** is the direction and magnitude of the repulsion of a negative electrical charge from a positive charge at the origin. The function `-@K11` may be interpreted as gravitational attraction. ]

C2 What is the relation between the Jacobian of the linear function **L11** and the determinant of the matrix **m** used in its definition?

C3 What is the relation between the Jacobians of two linear functions **LA11** and **LB11** and the Jacobian of `LC11 := LA11@LB11` (their composition).

[ `TEST := LA11@LB11 JAC |@- LA11 JAC * LB11 JAC` ]

C4 Define functions **LA11** and **LB11**, and test the comparison expressed in the solution to Exercise C3 by applying **TEST** to appropriate arguments.

C5 The Jacobian of the linear `LR11 := mp (>0 1;1 0) "1` is `_1`. State the significance of a negative Jacobian.

[ Plot figures **fig1** and **fig2**, and note that one can be moved smoothly onto the other "without crossing lines". Verify that this cannot be done with **fig1** and **LR11 fig1**; it is necessary to "lift the figure out of the plane and flip it over". A transformation whose Jacobian is negative is said to involve a "reflection". ]

C6 Enter, experiment with, and comment upon the following functions:

`RM2 := 2 2&$@(1 1 _1 1&*)@(2 1 1 2&o.) "0`

`R2 := (] mp RM2@[) "0 1`

[ **R2** is a linear function that produces a *rotation in 2-space*; the expression `a R2 fig` rotates a figure (such as **fig1** or **fig2**) about the origin through an angle of **a** radians in a counter-clockwise sense, without deforming the figure.]

C7 What is the value of the Jacobian of a rotation `a&R2 "1`?

C8 Enter an expression to define **FIG1** as an 8 by 3 table representing a cube, making sure that successive coordinates are adjacent, for example, `0 1 1` must not succeed `1 1 0`. Define 3-space linear functions to apply to **FIG1**, and use them together with **K11** to repeat Exercises 1-5 in 3-space.

C9 Enter, experiment with, and comment upon the functions

`RM3 := 1 0 0&,@(0&,.)@RM2`

`R30 := (] mp RM3@[) "0 1`

[ `a&R30 "1` produces a rotation through an angle **a** in the plane of the last two axes in 3-space (or *about* axis 0). Test the value of the Jacobian.]

C10 Define functions **R31** and **R32** that rotate about the other axes, and experiment with functions such as `a1&R31@(a2&R30) "1`.

[Experiment with the permutations `p := 2&A.` and `p := 5&A.` in the expression `p&. | :@p@RM3 o.%2`, and use the ideas in functions defined in terms of **R30**. ]

## D. Divergence and Laplacian

The *divergence* and *Laplacian* are defined and used as follows:

```

DIV=: GRAD TRACE

LAP=: GRAD DIV

f=: +/\ "1
f a
1 3 6

f GRAD a          f DIV a
1 0 0              3
1 1 0
1 1 1

g=: +/@ (] ^ >:@i.@#) "1
g a                g LAP a
32                 22.0268

```

It is difficult to provide a helpful interpretation of the divergence except in the context of an already-familiar physical application, and the reader may be best advised to seek interpretations in some familiar field. However, in his *Advanced Calculus* [8], F.S. Woods offers the following:

"The reason for the choice of the name *divergence* may be seen by interpreting  $\mathbf{F}$  as equal to  $r\mathbf{v}$ , where  $r$  is the density of a fluid and  $\mathbf{v}$  is its velocity. ... Applied to an infinitesimal volume it appears that  $\mathbf{div} \mathbf{F}$  represents the amount of fluid per unit time which streams or diverges from a point."

## E. Symmetry, Skew-Symmetry, and Orthogonality

A matrix that is equal to its transpose is said to be *symmetric*, and a matrix that equals the negative of its transpose is *skew-symmetric*. For example:

```

]m=: VOLS GRAD lwh  The gradient of the volumes function
1 0 0
3 4 0
6 8 12

| :m                The gradient is not symmetric
1 3 6
0 4 8
0 0 12

]ms=: (m+| :m) %2  The symmetric part of the gradient
1 1.5 3
1.5 4 4
3 4 12

]msk=: (m-| :m) %2  The skew-symmetric part
0 _1.5 _3

```

```
1.5  0  4
    3  4  0
```

```
    ms+msk          Sum of parts gives m
1 0  0
3 4  0
6 8 12
```

The determinant of any skew-symmetric matrix is 0, and its vectors therefore lie in a plane:

```
    det=-/ . *      The determinant function
    det msk        Shows that the vectors of msk lie in a plane
0
```

The axes of a rank-3 array can be "transposed" in several ways, by interchanging different pairs of axes. Such transposes are obtained by using `|:` with a left argument:

```
    ]a=:i.2 2 2
0 1
2 3

4 5
6 7

    0 2 1 |: a   Interchange last two axes
0 2
1 3

4 6
5 7

    1 0 2 |: a   Interchange first two axes
0 1
4 5

2 3
6 7
```

The permutation `0 2 1` is said to have *odd* parity because it can be brought to the normal order `0 1 2` by an odd number of interchanges of adjacent elements; `1 2 0` has *even* parity because it requires an even number of interchanges. The function `C.! .2` yields the parity of its argument, `1` if the argument has even parity, `_1` if odd, and `0` if it is not a permutation.

An array that is skew-symmetric under *any* interchange of axes is said to be *completely skew*. Such an array is useful in producing a vector that is *normal* (or *orthogonal* or *perpendicular*) to a plane. In particular, we will use it in a function called `norm` that produces the *curl* of a vector function, a vector normal to the plane of (the skew-symmetric part of) the gradient of the function.

We will generate a completely skew array by applying the parity function to the table of all indices of an array:

```
    indices={@(| # <@i.)

    indices 3
+-----+-----+-----+
|0 0 0|0 0 1|0 0 2|
```

```

+-----+-----+-----+
|0 1 0|0 1 1|0 1 2|
+-----+-----+-----+
|0 2 0|0 2 1|0 2 2|
+-----+-----+-----+

```

```

+-----+-----+-----+
|1 0 0|1 0 1|1 0 2|
+-----+-----+-----+
|1 1 0|1 1 1|1 1 2|
+-----+-----+-----+
|1 2 0|1 2 1|1 2 2|
+-----+-----+-----+

```

```

+-----+-----+-----+
|2 0 0|2 0 1|2 0 2|
+-----+-----+-----+
|2 1 0|2 1 1|2 1 2|
+-----+-----+-----+
|2 2 0|2 2 1|2 2 2|
+-----+-----+-----+

```

`e=:C.!2@>@indices`      Result is called an "e-system" by McConnell [4]

```

e 3
0 0 0
0 0 1
0 _1 0

0 0 _1
0 0 0
1 0 0

0 1 0
_1 0 0
0 0 0

```

<"2 e 4    Boxed for convenient viewing

```

+-----+-----+-----+-----+
|0 0 0 0 |0 0 0 0|0 0 0 0 |0 0 0 0|
|0 0 0 0 |0 0 0 0|0 0 0 0 |0 0 1 0|
|0 0 0 0 |0 0 0 1|0 0 0 0 |0 0 _1 0 0|
|0 0 0 0 |0 0 _1 0|0 1 0 0 |0 0 0 0|
+-----+-----+-----+-----+
|0 0 0 0 |0 0 0 0 |0 0 0 1|0 0 _1 0 0|
|0 0 0 0 |0 0 0 0 |0 0 0 0|0 0 0 0|
|0 0 0 _1|0 0 0 0 |0 0 0 0|1 0 0 0|
|0 0 1 0|0 0 0 0 |_1 0 0 0|0 0 0 0|
+-----+-----+-----+-----+
|0 0 0 0|0 0 0 0 |_1 0 0 0 0 |0 1 0 0|
|0 0 0 1|0 0 0 0 |0 0 0 0 0 |_1 0 0 0|
|0 0 0 0|0 0 0 0 |0 0 0 0 0 |0 0 0 0|
|0 _1 0 0|1 0 0 0 |0 0 0 0 0 |0 0 0 0|
+-----+-----+-----+-----+

```

```

|0 0 0 0| 0 0 1 0|0 _1 0 0|0 0 0 0 |
|0 0 _1 0| 0 0 0 0|1 0 0 0|0 0 0 0 |
|0 1 0 0|_1 0 0 0|0 0 0 0|0 0 0 0 |
|0 0 0 0|0 0 0 0|0 0 0 0|0 0 0 0 |
+-----+-----+-----+-----+

```

Finally, we will use `e` in the definition of the function `norm`, as follows:

```

norm=:+/^:(|`(#@$)`(* e@#)) % !@(# - #@$)

mp=:+/ . *           Matrix product

]m=:VOLS GRAD lwh    Gradient of the volumes function
1 0 0
3 4 0
6 8 12

]skm=(m-|:m)%2      Skew part
0 _1.5 _3
1.5 0 _4
3 4 0

]orth=:norm m        Result is perpendicular to plane of skm
_8 6 _3

orth mp skm          Test of perpendicularity
0 0 0

norm skm              Norm of skew part gives the same result
_8 6 _3

norm norm skm        Norm on a skew matrix is self-inverse
0 _1.5 _3
1.5 0 _4
3 4 0

```

These matters are discussed further in Chapter 6.

## F. Curl

The curl is the perpendicular to the grade, and is produced by the function `norm`. We will use the adverb form as follows:

```

NORM=:norm@

CURL=: GRAD NORM

VOLS CURL lwh
_8 6 _3

subtotals=:+/\
subtotals lwh
4 7 9

subtotals CURL lwh
_1 1 _1

```

Interpretation of the curl is perhaps even more intractable than the divergence. Again Woods offers some help:

The reason for the use of the word *curl* is hard to give without extended treatment of the subject of fluid motion. The student may obtain some help by noticing that if  $\mathbf{F}$  is the velocity of a liquid, then for velocity in what we have called irrotational motion,  $\text{curl } \mathbf{F} = \mathbf{0}$ , and for vortex motion,  $\text{curl } \mathbf{F} \neq \mathbf{0}$ .

It may be shown that if a spherical particle of fluid be considered, its motion in a time  $dt$  may be analyzed into a translation, a deformation, and a rotation about an instantaneous axis. The curl of the vector  $\mathbf{v}$  can be shown to have the direction of this axis and a magnitude equal to twice the instantaneous angular velocity.

In his *Div, Grad, Curl, and all that* [9], H.M. Schey makes an interesting attempt to introduce the concepts of the vector calculus in terms of a single topic. His first chapter begins with:

In this text the subject of the vector calculus is presented in the context of simple electrostatics. We follow this procedure for two reasons. First, much of vector calculus was invented for use in electromagnetic theory and is ideally suited to it. This presentation will therefore show what vector calculus is, and at the same time give you an idea of what it's for. Second, we have a deep-seated conviction that mathematics -in any case *some* mathematics- is best discussed in a context which is not exclusively mathematical.

Schey's treatment includes Maxwell's equations which, in Heaviside's elegant formulation, exhibit the powers of div, grad, and curl in joint use.

F1 Experiment with **GRAD**, **CURL**, **DIV**, and **JAC** on the functions in Exercise B2.

F2 Experiment with **GRAD**, **CURL**, **DIV**, and **JAC** on the following 1 1 functions:

```
q=: *:"1
r=: 4&A. @: q
s=: 1 1 _1&* @: r
t=: 3&A. @: ^ @: -
u=: ]% (+/@(*~)) ^ 3r2"0
```

F3 Enter the definitions  $\mathbf{x} =: 0\&\{$  and  $\mathbf{y} =: 1\&\{$  and  $\mathbf{z} =: 2\&\{$ , and use them to define the functions of the preceding exercise in a more conventional form.

```
[ as=: *:@z, *:@x, -@*:@y
at=: ^@-@y, ^@-@z, ^@-@x
au=: (x,y,z) % (*:@x + *:@y + *:@z) ^ 3r2"0 ]
```

F4 Experiment with **LAP** on various 0 1 functions.

F5 Express the cross product of Section 6G so as to show its relation to **CURL**. See Section 6H.

```
[ CR=: */ NORM          CURL=: GRAD NORM ]
```

# Chapter 4

## Difference Calculus

### A. Introduction

Although published some fifty years ago, Jordan's *Calculus of Finite Differences* [10] still provides an interesting treatment. In his introductory section on *Historical and Biographical Notes*, he contrasts the difference and differential (or *infinitesimal*) calculus:

Two sorts of functions are to be distinguished. First, functions in which the variable  $x$  may take every possible value in a given interval; that is, the variable is continuous. These functions belong to the domain of the *Infinitesimal Calculus*. Secondly, functions in which the variable takes only the given values  $x_0, x_1, x_2, \dots, x_n$ ; then the variable is discontinuous. To such functions the methods of Infinitesimal Calculus are not applicable, The *Calculus of Finite Differences* deals especially with such functions, but it may be applied to both categories.

The present brief treatment is restricted to three main ideas:

- 1) The development of a family of functions which behaves as simply under the *difference* (secant slope) adverb as does the family of power functions  $x^n$  under the derivative adverb.
- 2) The definition of a polynomial function in terms of this family of functions.
- 3) The development of a linear transformation from the coefficients of such a polynomial to the coefficients of an equivalent ordinary polynomial.

### B. Secant Slope Conjunctions

The slope of a line from the point  $\mathbf{x}, \mathbf{f}(\mathbf{x})$  to the point  $\mathbf{x} + \mathbf{r}, \mathbf{f}(\mathbf{x} + \mathbf{r})$  is said to be the *secant slope of  $\mathbf{f}$  for a run of  $\mathbf{r}$* , or the *r-slope* of  $\mathbf{f}$  at  $\mathbf{x}$ . Thus:

```

cube=: ^&3"0
x=: 1 2 3 4 5
r=: 0.1
((cube x+r) - (cube x))%r
3.31 12.61 27.91 49.21 76.51

```

The same result is given by the *secant-slope* conjunction **D**: as follows:

```

    r cube D: 1 x
3.31 12.61 27.91 49.21 76.51

    0.01 cube D: 1 x
3.0301 12.0601 27.0901 48.1201 75.1501

    0.0001 cube D: 1 x
3.0003 12.0006 27.0009 48.0012 75.0015

    cube d. 1 x
3 12 27 48 75
    3*x^2
3 12 27 48 75

```

In the foregoing sequence, smaller runs appear to be approaching a limiting value, a value given by the derivative. It is also equal to three times the square.

The alternate expression  $((\text{cube } x) - (\text{cube } x-r))\%r$  could also be used to define a slope, and it will prove more convenient in our further work. We therefore define an alternate conjunction for it as follows:

```

    SLOPE=:2 : (':' ; 'x. u."0 D: n. y.-x.')

    r cube SLOPE 1 x
2.71 11.41 26.11 46.81 73.51

    ((cube x)-(cube x-r))%r
2.71 11.41 26.11 46.81 73.51

    0.0001 cube SLOPE 1 x
2.9997 11.9994 26.9991 47.9988 74.9985

    cube d. 1 x
3 12 27 48 75

```

Much like the derivative, the slope conjunction can be used to give the slope of the slope, and so on. Thus:

```

    cube d.2 x
6 12 18 24 30
    r cube SLOPE 2 x
6 12 18 24 30

```

We will be particularly concerned with the "first" slope applied to scalar (rank-0) functions, and therefore define a corresponding adverb:

```

    S=:("0) SLOPE 1
    r cube S x
2.71 11.41 26.11 46.81 73.51

```

## C. Polynomials and Powers

In Chapter 3, the analysis of the power function  $x^n$  led to the result that the derivative of the polynomial  $c\&p$ . could be written as another polynomial:  $(\}.c*i.#c)\&p..$

This is an important property of the family of power functions, and we seek another family of functions that behaves similarly under the  $r$ -slope. We begin by adopting the



names  $p_0$  and  $p_1$  and  $p_2$ , etc., for the functions  $\wedge\&0$  and  $\wedge\&1$  and  $\wedge\&2$ , and by showing how each member of the family can be defined in terms of another. Thus:

$$p_4 = ] * p_3 = ] * p_2 = ] * p_1 = ] * p_0 = 1 : "0$$

The following expressions for the derivatives of sums and products of functions were derived in Chapter 1. The corresponding expressions for the  $r$ -slopes may be obtained by simple algebra:

$f + g$	Sum
$(r f S) + (r g S)$	$r$ -Slope
$(f d.1) + (g d.1)$	Derivative
$f * g$	
	Product
$(f * (r g S)) + ((r f S) * g) - (r^0 * (r f S) * (r g S))$	$r$ -Slope
$(f * g d.1) + (f d.1 * g)$	Derivative

For example:

$r = 0.1$	
$x = 1 \ 2 \ 3 \ 4 \ 5$	
$f = \wedge\&3$	
$g = \wedge\&2$	
$(f+g) \ x$	
2 12 36 80 150	
$r (f+g) \ S \ x$	Slope of sum
4.61 15.31 32.01 54.71 83.41	
$(r f S \ x) + (r g S \ x)$	Sum of slopes
4.61 15.31 32.01 54.71 83.41	
$(f+g) \ d. \ 1 \ x$	Derivative of sum
5 16 33 56 85	
$(f \ d.1 + g \ d.1) \ x$	
5 16 33 56 85	
$r (f*g) \ S \ x$	Slope of product
4.0951 72.3901 378.885 1217.58 3002.48	
$]t1 = (f \ x) * (r \ g \ S \ x)$	Terms for slope of product
1.9 31.2 159.3 505.6 1237.5	
$]t2 = (r \ f \ S \ x) * (g \ x)$	
2.71 45.64 234.99 748.96 1837.75	
$]t3 = r * (r \ f \ S \ x) * (r \ g \ S \ x)$	
0.5149 4.4499 15.4049 36.9799 72.7749	
$t1+t2-t3$	Sum and diff of terms gives slope
4.0951 72.3901 378.885 1217.58 3002.48	
$(f*g) \ d. \ 1 \ x$	Derivative of product
5 80 405 1280 3125	
$((f \ d.1 * g) + (f * g \ d.1)) \ x$	
5 80 405 1280 3125	

Since the derivative of the identity function  $J$  is the constant function  $1''0$ , expressions for the derivatives of the power functions can be derived using the expressions for the sum and product in informal proofs as follows:

$$\begin{array}{lll}
 p_0 \text{ d.1} & p_1 \text{ d.1} & p_2 \text{ d.1} \\
 1''0 \text{ d.1} & (J * p_0) \text{ d.1} & (J * p_1) \text{ d.1} \\
 0''0 & (J * p_0 \text{ d.1}) + (J \text{ d.1} * 1''0) & (J * p_1 \text{ d.1}) + (J \text{ d.1} * p_1) \\
 & (J * 0''0) + (1''0 * 1''0) & (J * 1''0) + (1''0 * p_1) \\
 & 1''0 & p_1 + p_1 \\
 & & 2''0 * p_1 \\
 \\
 p_3 \text{ d.1} & & p_4 \text{ d.1} \\
 (J * p_2) \text{ d.1} & & (J * p_3) \text{ d.1} \\
 (J * p_2 \text{ d.1}) + (J \text{ d.1} * p_2) & & (J * p_3 \text{ d.1}) + (J \text{ d.1} * p_3) \\
 (J * 2''0 * p_1) + (1''0 * p_2) & & (J * 3''0 * p_2) + (1''0 * p_3) \\
 (2''0 * p_2) + p_2 & & (3''0 * p_3) + p_3 \\
 3''0 * p_2 & & 4''0 * p_3
 \end{array}$$

Each of the expressions in the proofs may be tested by applying it to an argument such as  $x=: i. 6$ , first enclosing the entire expression in parentheses.

We will next introduce *stope* functions whose behavior under the slope operator is analogous to the behavior of the power function under the derivative.

## D. Stope Functions

The list  $x+r*i.n$  begins at  $x$  and changes in steps of size  $r$ , like the steps in a mine stope that follows a rising or falling vein of ore. We will call the product over such a list a *stope*:

$$\begin{array}{ll}
 x=:5 & \\
 r=:0.1 & \\
 n=:4 & \\
 \\
 x+r*i.n & \\
 5 \ 5.1 \ 5.2 \ 5.3 & \\
 \\
 */x+r*i.n & \\
 702.78 & \\
 \\
 */x+1*i.n & \text{Case } r=:1 \text{ is called a } \textit{rising factorial} \\
 1680 & \\
 \\
 */x+_1*i.n & \text{Falling factorial} \\
 120 & \\
 \\
 */x+0*i.n & \text{Case } r=:0 \text{ gives product over list of } n \text{ } x\text{'s} \\
 625 & \\
 x^n & \text{Equivalent to the power function} \\
 625 &
 \end{array}$$

The two final examples illustrate the fact that the case  $r=:0$  is equivalent to the power function. We therefore treat the stope as a *variant* of the power function, produced by the conjunction  $!$ . as follows:

$$x \text{ } ^{!.r} \text{ } n$$

702.78

```
x ^!.0 n
625
```

```
stope=: ^!. The stope adverb
x r stope n
702.78
```

We now define a set of stope functions analogous to the functions  $p_0=:^&0$  and  $p_1=:^&1$ , etc. used for successive powers. Thus:

```
q0=:r stope&0
```

```
q1=:r stope&1
```

```
q2=:r stope&2
```

```
q3=:r stope&3
```

```
q4=:r stope&4
```

```
x=:0 1 2 3 4
>(q0;q1;q2;q3;q4) x
1 1 1 1 1
0 1 2 3 4
0 1.1 4.2 9.3 16.4
0 1.32 9.24 29.76 68.88
0 1.716 21.252 98.208 296.184
```

## E. Slope of the Stope

We will now illustrate that the  $r$ -slope of  $r \text{ stope } n$  is  $n * r \text{ stope } (n-1)$ :

```
r q4 S x
0 5.28 36.96 119.04 275.52
```

```
4*q3 x
0 5.28 36.96 119.04 275.52
```

```
r q3 S x
0 3.3 12.6 27.9 49.2
```

```
3*q2 x
0 3.3 12.6 27.9 49.2
```

This behavior is analogous to that of the power functions  $p_4$ ,  $p_3$ , etc. under the derivative. Moreover, the stope functions can be defined as a sequence of products, in a manner similar to that used for defining the power functions. Thus (using  $R$  for a constant function):

```
R=:r"0
```

```
f4=: ([+3"0*R) *f3=: ([+2"0*R) *f2=: ([+1"0*R) *f1=: ([+0"0*R) *f0=:1"0
```

From these definitions, the foregoing property of the  $r$ -slopes of stopes can be obtained in the manner used for the derivative of powers, but using the expression:

$$(f'(r)g(S)) + ((r f(S))'g) - (R'(r f(S))'(r g(S)))$$

For the  $r$ -slope of the product of functions instead of the:

$$(f'g \text{ d.1}) + (f \text{ d.1}g)$$

used for the derivative.

## F. Slope Polynomials

The polynomial function  $p$  also possesses a variant  $p.^!.r$ , in which the terms are based upon the slope  $^!.r$  rather than upon the power  $^$ . For example:

$$\text{spr} = p.^!.r$$

$$c = :4 \ 3 \ 2 \ 1$$

$$c \&\text{spr } x$$

$$4 \ 10.52 \ 27.64 \ 61.36 \ 117.68$$

$$(4*x \ ^!.r \ 0) + (3*x \ ^!.r \ 1) + (2*x \ ^!.r \ 2) + (1*x \ ^!.r \ 3)$$

$$4 \ 10.52 \ 27.64 \ 61.36 \ 117.68$$

The  $r$ -slope of the *slope polynomial*  $c \&\text{spr}$  then behaves analogously to the derivative of the ordinary polynomial. Thus:

$$\text{deco} = :1 \}.] * i. \# \quad \text{Function for coefficients of derivative polynomial}$$

$$]d = : \text{deco } c$$

$$3 \ 4 \ 3$$

$$c \&p. \ x \quad \text{Ordinary polynomial with coefficients } c$$

$$4 \ 10 \ 26 \ 58 \ 112$$

$$c \&p. \ \text{d.1 } x \quad \text{Derivative of polynomial}$$

$$3 \ 10 \ 23 \ 42 \ 67$$

$$d \ p. \ x \quad \text{Agrees with polynomial with "derivative" coefficients}$$

$$3 \ 10 \ 23 \ 42 \ 67$$

$$\text{spr} = p.^!.r \quad \text{Slope polynomial for run } r$$

$$c \&\text{spr } x \quad \text{Slope polynomial with coefficients } c$$

$$4 \ 10.52 \ 27.64 \ 61.36 \ 117.68$$

$$r \ c \&\text{spr } S \ x \quad r\text{-slope of slope polynomial}$$

$$3 \ 10.3 \ 23.6 \ 42.9 \ 68.2$$

$$d \&\text{spr } x \quad \text{Agrees with slope polynomial with coefficients } d$$

$$3 \ 10.3 \ 23.6 \ 42.9 \ 68.2$$

We now define a slope polynomial adverb, whose argument specifies the run:

$$\text{SPA} = :1 \ : \ '[ \ p.^!.x. \ ]'$$

$$c \ 0 \ \text{SPA } x \quad \text{Zero gives ordinary polynomial}$$

$$4 \ 10 \ 26 \ 58 \ 112$$

$$c \ p. \ x$$

$$4 \ 10 \ 26 \ 58 \ 112$$

$$c \ r \ \text{SPA } x \quad \text{Slope with run } r$$

```
4 10.52 27.64 61.36 117.68
```

Integration behaves analogously:

```
adeco=: [ , ] % >:@i.@#[] The integral coefficient function

(0 adeco c)&spr x
0 6.959 25.773 70.342 160.566
```

## G. Coefficient Transformations

It is important to be able to express an ordinary polynomial as an equivalent stope polynomial, and vice versa. We will therefore show how to obtain the coefficients for an ordinary polynomial that is equivalent to a stope polynomial with given coefficients:

The expression `vm=:x ^/ i.#c` gives a table of powers of `x` that is called a *Vandermonde* matrix. If `mp=:+/. *` is the matrix product, then `vm mp c` gives weighted sums of these powers that are equivalent to the polynomial `c p. x`. For example:

```
x=:2 3 5 7 11
c=:3 1 4 2 1
c p. x
53 177 983 3293 17801

]vm=:x ^/ i.#c
1 2 4 8 16
1 3 9 27 81
1 5 25 125 625
1 7 49 343 2401
1 11 121 1331 14641

mp=:+/. *

]y=:vm mp c
53 177 983 3293 17801
```

If `x` has the same number of elements as `c`, and if the elements of `x` are all distinct, then the matrix `vm` is non-singular, and its inverse can be used to obtain the coefficients of a polynomial that gives any specified result. If the result is `y`, these coefficients are, of course, the original coefficients `c`. Thus:

```
(%.vm) mp y
3 1 4 2 1
```

The coefficients `c` used with a stope polynomial give a different result `y2`, to which we can apply the same technique to obtain coefficients `c2` for an equivalent ordinary polynomial:

```
r=:0.1
]y2=:c p.!r x
61.532 200.928 1077.98 3536.71 18690.4

]c2=(%.vm) mp y2
3 1.446 4.71 2.6 1

c2 p. x
61.532 200.928 1077.98 3536.71 18690.4
```

We now incorporate this method in a conjunction **FROM**, such that **r1 FROM r2** gives a function which, applied to coefficients **c**, yields **d** such that  $d \cdot p.!.r1 \cdot x$  is equivalent to  $c \cdot p.!.r2 \cdot x$ . Thus:

```

VM=:1 : '[ ^!.x./i.@#]'

FROM=: 2 : '((y. VM %. x. VM)~ @i.@#) mp ]'

]cr=:r FROM 0 c
3 0.619 3.47 1.4 1

cr p.!.r x
53 177 983 3293 17801
c p.!.0 x
53 177 983 3293 17801

```

A conjunction that yields the corresponding Vandermonde matrix rather than the coefficients can be obtained by removing the final matrix product from **FROM**. For the case of the falling factorial function (**r=:\_1**) this matrix gives results of general interest:

```

VMFROM=: 2 : '((y. VM %. x. VM)~ @i.@#)'

0 VMFROM _1 c
1 0 0 0 0
0 1 _1 2 _6
0 0 1 _3 11
0 0 0 1 _6
0 0 0 0 1

_1 VMFROM 0 c
1 0 0 0 0
0 1 1 1 1
0 0 1 3 7
0 0 0 1 6
0 0 0 0 1

```

The elements of the last of these tables are called *Stirling numbers of the second kind*, and the magnitudes of those of the first are *Stirling numbers of the first kind*.

- G1 Experiment with the adverb **VM**.  
 D2 Enter expressions to obtain the matrices **S1** and **S2** that are Stirling numbers of order 6 (that is, \$ **S1** is 6 6).

```
[c=: 6?9          S1=:0 FROM 1 c          S2=:1 FROM 0 c]
```

- D3 Test the assertion that **S1** is the inverse of **S2**.

## H. Slopes as Linear Functions

A linear function can be represented by a matrix bonded with the matrix product. For example, if **v** is a vector and **ag=: <:/~@i.@#**, then **sum=: ag v** is a *summation* or *aggregation* matrix; the linear function (**mp=: +/ . \***)&sum produces sums over prefixes of its argument. Thus:

```
]v=: ^&3 i. 6
0 1 8 27 64 125
```

```

mp=: +/ . *
ag=: <:/~@i.@#
sum=: ag v

sum          sum mp sum
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1
0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1

v mp sum          v mp (sum mp sum)
0 1 9 36 100 225 0 1 10 46 146 371

+/\v              +/\ +/\v
0 1 9 36 100 225 0 1 10 46 146 371

mp&sum v          mp&(sum mp sum) v
0 1 9 36 100 225 0 1 10 46 146 371

L1=: mp&sum       L2=:mp&(sum mp sum)

L1 v              L2 v
0 1 9 36 100 225 0 1 10 46 146 371

+/\v              +/\ +/\v
0 1 9 36 100 225 0 1 10 46 146 371

mp&sum v          mp&(sum mp sum) v
0 1 9 36 100 225 0 1 10 46 146 371

L1=: mp&sum       L2=:mp&(sum mp sum)

L1 v              L2 v
0 1 9 36 100 225 0 1 10 46 146 371

```

The results of `L1 v` are rough approximations to the areas under the graph of  $x^3$ , that is, to the integrals up to successive points. Similarly, the inverse matrix `dif=: %. sum` can define a linear function that produces differences between successive elements of its argument. For example:

```

dif=: %. sum

dif          dif mp dif
1 _1 0 0 0 0 1 _2 1 0 0 0
0 1 _1 0 0 0 0 1 _2 1 0 0
0 0 1 _1 0 0 0 0 1 _2 1 0
0 0 0 1 _1 0 0 0 0 1 _2 1
0 0 0 0 1 _1 0 0 0 0 1 _2
0 0 0 0 0 1 _1 0 0 0 0 0 1

LD1=: mp&dif      LD2=: mp&(dif mp dif)
LD1 v             LD2 v
0 1 7 19 37 61   0 1 6 12 18 24

```

These results may be compared with the 1-slopes of the cube function, noting that the first  $k$  elements of the  $k$ th slope are meaningless. The  $r$ -slopes of a function  $f$  can be obtained similarly, by applying `%r@LD1` to the results of  $f$  applied to arguments differing by  $r$ . For example:

```

]x=: r*i.6 [ r=: 0.1
0 0.1 0.2 0.3 0.4 0.5

%r@LD1 ^&3 x
0 0.01 0.07 0.19 0.37 0.61
r (^&3) S x
0.01 0.01 0.07 0.19 0.37 0.61

```

Because the results for the 1-slope are so easily extended to the case of a general  $r$ -slope, we will discuss only the 1-slope provided by the linear function `DIF=: mp&dif`.

Consider the successive applications of `DIF` to the identity matrix:

```

ID=: (i. =/ i.) 6
DIF=: mp&dif
ID          DIF ID          DIF DIF ID
1 0 0 0 0 0    1 _1 0 0 0 0    1 _2 1 0 0 0
0 1 0 0 0 0    0 1 _1 0 0 0    0 1 _2 1 0 0
0 0 1 0 0 0    0 0 1 _1 0 0    0 0 1 _2 1 0
0 0 0 1 0 0    0 0 0 1 _1 0    0 0 0 1 _2 1
0 0 0 0 1 0    0 0 0 0 1 _1    0 0 0 0 1 _2
0 0 0 0 0 1    0 0 0 0 0 1    0 0 0 0 0 1

```

```

2 3$ <"2@(DIF^:0 1 2 3 4 5) ID
+-----+-----+-----+
| 1 0 0 0 0 0 | 1 _1 0 0 0 0 | 1 _2 1 0 0 0 |
| 0 1 0 0 0 0 | 0 1 _1 0 0 0 | 0 1 _2 1 0 0 |
| 0 0 1 0 0 0 | 0 0 1 _1 0 0 | 0 0 1 _2 1 0 |
| 0 0 0 1 0 0 | 0 0 0 1 _1 0 | 0 0 0 1 _2 1 |
| 0 0 0 0 1 0 | 0 0 0 0 1 _1 | 0 0 0 0 1 _2 |
| 0 0 0 0 0 1 | 0 0 0 0 0 1 | 0 0 0 0 0 1 |
+-----+-----+-----+
| 1 _3 3 _1 0 0 | 1 _4 6 _4 1 0 | 1 _5 10 _10 5 _1 |
| 0 _1 3 _3 1 0 | 0 _1 4 _6 4 1 | 0 _1 5 _10 10 _5 |
| 0 0 1 3 _1 0 | 0 0 1 4 _6 4 | 0 0 1 5 _10 10 |
| 0 0 0 1 3 _1 | 0 0 0 1 4 _6 | 0 0 0 1 5 _10 |
| 0 0 0 0 1 3 | 0 0 0 0 1 4 | 0 0 0 0 1 5 |
| 0 0 0 0 0 1 | 0 0 0 0 0 1 | 0 0 0 0 0 1 |
+-----+-----+-----+

```

The foregoing results suggest that the  $k$ -th difference is a weighted sum of  $k+1$  elements in which the weights are the alternating binomial coefficients of order  $k$ . For example:

```

]v=: ^&3 i. 8
0 1 8 27 64 125 216 343

w=: mp & 1 _2 1
w 0 1 2{v
6
w 1 2 3{v          w 2 3 4{v          w 3 4 5{v
12                18                24

3 <\ v              Box applied to each 3-element window
+-----+-----+-----+

```



```
|0 1 8|1 8 27|8 27 64|27 64 125|64 125 216|125 216 343|
+-----+-----+-----+-----+-----+-----+-----+-----+
```

```
  3 w\ v          Weighting function applied to
6 12 18 24 30 36  each 3-element window
  4 (mp & _1 3 _3 1)\ v  The third difference of the cube
6 6 6 6 6         function is the constant !3
```

```
  6 (mp & _1 5 _10 10 _5 1)\ ^&5 i. 11
120 120 120 120 120 120
```

The binomial coefficients of order  $n$  are provided by the expression  $(i. n+1) ! n$ . For example:

```
(i.@>: ! ]) n=: 5
1 5 10 10 5 1
```

The alternating coefficients could be obtained by multiplying alternate elements by  $\_1$ . However, they are provided more directly by the extension of the function  $!$  to negative arguments, as may be seen in the following "bordered" function table:

```
]i=: i: 7
_7 _6 _5 _4 _3 _2 _1 0 1 2 3 4 5 6 7
```

```
  i ! table i
+-----+
| |  _7  _6  _5  _4  _3  _2  _1  0  1  2  3  4  5  6  7|
+-----+
|_7|  1  _6  15  _20  15  _6  1  0  0  0  0  0  0  0  0|
|_6|  0  1  _5  _10  _10  _5  _1  0  0  0  0  0  0  0  0|
|_5|  0  0  1  _4  _6  _4  1  0  0  0  0  0  0  0  0|
|_4|  0  0  0  1  _3  _3  1  0  0  0  0  0  0  0  0|
|_3|  0  0  0  0  1  _2  _1  0  0  0  0  0  0  0  0|
|_2|  0  0  0  0  0  1  _1  0  0  0  0  0  0  0  0|
|_1|  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0|
|_0|  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1|
| 1|  _7  _6  _5  _4  _3  _2  _1  0  1  2  3  4  5  6  7|
| 2|  28  21  15  10  6  3  1  0  0  1  3  6  10  15  21|
| 3|  84  56  35  20  10  4  1  0  0  0  1  4  10  20  35|
| 4|  210 126  70  35  15  5  1  0  0  0  0  1  5  15  35|
| 5|  462 252 126  56  21  6  1  0  0  0  0  0  1  6  21|
| 6|  924 462 210  84  28  7  1  0  0  0  0  0  0  1  7|
| 7| 1716 792 330 120  36  8  1  0  0  0  0  0  0  0  1|
+-----+
```

Except for a change of sign required for those of odd order, the required alternating binomial coefficients can be seen in the diagonals beginning in row 0 of the negative columns of the foregoing table. The required weights are therefore given by the following function:

```
w=: _1&^ * (i. ! i. - ])@>:"0
w 0 1 2 3 4
1 0 0 0 0
_1 1 0 0 0
 1 _2 1 0 0
_1 3 _3 1 0
 1 _4 6 _4 1
```

Differences may therefore be expressed as shown in the following examples:

$v = \sum_{i=0}^{\infty} \frac{1}{i!} x^i$   
0 1 8 27 64 125 216 343

$\frac{1}{2} \sum_{i=0}^{\infty} \frac{1}{i!} x^i$   
1 7 19 37 61 91 127

$\frac{1}{6} \sum_{i=0}^{\infty} \frac{1}{i!} x^i$   
6 12 18 24 30 36

$\frac{1}{120} \sum_{i=0}^{\infty} \frac{1}{i!} x^i$   
1560 3360 5880 9120 13080 17760

It may also be noted that the diagonals beginning in row 0 of the *non-negative* columns of the table contain the weights appropriate to successive integrations as, for example, in the diagonals beginning with 1 1 1 1 1 and 1 2 3 4 5 and 1 3 6 10 15. This fact can be used to unite the treatment of derivatives and integrals in what Oldham and Spanier call *differintegrals* in their *Fractional Calculus* [5]. Moreover, the fact that the function  $\Gamma$  is generalized to non-integer arguments will be used (in Chapter 5) to define fractional derivatives and integrals. For example:

$\Gamma(4) = 3!$   
1 4 6 4 1 0 00

$\Gamma(4) = (0.01 + i.7)!$  Formatted to four decimal places  
1.0210 4.0333 5.9998 3.9666 0.9793 0.0020 0.0003

# Chapter 5

## Fractional Calculus

### A. Introduction

The differential and the difference calculus of Chapters 2 and 4 concern derivatives and integrals of integer order. The fractional calculus treated in this chapter unites the derivative and the integral in a single *differintegral*, and extends its domain to non-integer orders.

Section H of Chapter 1 included a brief statement of the utility of the fractional calculus and a few examples of fractional derivatives and integrals. Section E of Chapter 4 concluded with the use of the alternating binomial coefficients produced by the *outof* function ! to compute differences of arbitrary integer order. The extension of the function ! to non-integer arguments was also cited as the basis for an analogous treatment of non-integer differences, and therefore as a basis for approximating non-integer differintegrals.

Our treatment of the fractional calculus will be based on Equation 3.2.1 on page 48 of OS (Oldham and Spanier [5]). Thus:

<code>f=: ^&amp;3</code>	Function treated
<code>q=: 2</code>	Order of differintegral
<code>N=: 100</code>	Number of points used in approximation
<code>a=: 0</code>	Starting point of integration
<code>x=: 3</code>	Argument
<code>OS=: '+/(s^-q)*(j!j-1+q)*f x-(s=:N%~x-a)*j=:i.N'</code> ". OS 17.82	Execute the Oldham Spanier expression to obtain the approximation to the second derivative of $f$ at $x$
<code>q=: 1</code> ". OS 26.7309	Approximation to the first derivative (the exact value is $3*x^2$ , that is, 27)
<code>q=: 0</code> ". OS	Zeroth derivative (the function itself)

27

```

q=: _1
". OS
20.657

```

The first integral (exact value is  $4\int x^4$ )

```

q=: _2
". OS
12.7677

```

The second integral (exact value is  $20\int x^5$ )

```

q=: 0.5
". OS
27.9682

```

Semi-derivative (exact value is  $28.1435$ )

We will use the expression OS to define a *fractional differintegral* conjunction fd such that  $q$  ( $a, N$ ) fd  $f$   $x$  produces an  $N$ -point approximation to the  $q$ -th derivative of the function  $f$  at  $x$  if  $q > 0$ , and the  $(|q|)$ -th integral from  $a$  to  $x$  if  $q$  is negative:

```

j=: ("_) (i.@}.@)
s=: (&(([] - 0: { [] % 1: { [])) (@])
m=: '[:+/(x.s^0:-[])*(x.j!x.j-1:+[])*[:y.]-x.s*x.j'
fd=: 2 : m

```

For example:

```

2 (0,100) fd (^&3) 3
17.82
2 (0,100) fd (^&3)"0 i. 4
_ . 5.94 11.88 17.82

```

An approximation to a derivative given by a set of  $N$  points will be better over shorter intervals. For example:

```

x=: 6
1 (0,100) fd f x
106.924
3*x^2
108
1 ((x-0.01),100) fd f x
107.998

```

Anyone wishing to study the OS formulation and discussion will need to appreciate the relation between the function ! used here, and the gamma function (G) used by OS. Although the gamma function was known to be a generalization of the factorial function on integer arguments, it was not defined to agree with it on integers. Instead,  $G n$  is equivalent to  $! n-1$ . Moreover, the dyadic case  $m!n$  is here defined as  $(!n) % (!m) * (!n-m)$ ; the three occurrences of the gamma function in Equation 3.2.1 of OS may therefore be written as  $j!j-1+q$ , as seen in the expression OS used above.

The related *complete beta* function is also used in OS, where it is defined (page 21) by  $B(p,q) = (G p) * (G q) % (G p+q)$ . This definition may be re-expressed so as to show its

relation to the binomial coefficients, by substituting  $m$  for  $p-1$  and  $n$  for  $p+q-1$ . The expression  $B(p,q)$  is then equivalent to  $(-1)^m \binom{-n-m}{-n}$ , or simply  $\binom{m}{n}$ .

## B. Table of Semi-Differintegrals

The differintegrals of the sum  $f+g$  and the difference  $f-g$  are easily seen to be the sums and differences of the corresponding differintegrals, and it might be expected that fractional derivatives satisfy further relationships analogous to those shown in Section 2K for the differential calculus. Such relations are developed by Oldham and Spanier, but most are too complex for treatment here.

We will confine attention to a few of their semi-differintegrals (of orders that are integral multiples of  $0.5$  and  $0.5$ ). We begin by defining a conjunction  $FD$  (similar to  $fd$ , but with the parameters  $a$  and  $N$  fixed at  $0$  and  $100$ ), and using it to define adverbs for approximating semi-derivatives and semi-integrals:

```
FD=: 2 : 'x."0 (0 100) fd y. ]' ("0
x=: 1 2 3 4 5
1 FD (^&3) x
2.9701 11.8804 26.7309 47.5216 74.2525
```

```
3*x^2          Exact expression
3 12 27 48 75
```

```
si=: _1r2 FD          _1r2 is the rational constant _1%2
sd=: 1r2 FD
s3i=: _3r2 FD
^&3 sd x          Semi-derivative of cube
1.79416 10.1493 27.9682 57.4131 100.296
```

```
sdc=: *:!@[*(4&*@) ^ [] % !@+:@[ *%:@o.@]
3 sdc x          Exact expression from OS[5] page 119
1.80541 10.2129 28.1435 57.773 100.925
```

```
^&3 si x          Semi-integral of cube
0.520349 5.88707 24.3343 66.6046 145.442
```

```
sic=: *:!@[*(4&*@)^+&0.5@[]%!@>:@+:@[*%:@o.@1:
3 sic x          Exact function from OS[5] page 119
0.51583 5.83596 24.123 66.0263 144.179
```

Although the conjunctions  $sd$  and  $si$  and  $s3i$  provide only rough approximations, we will use them in the following table to denote exact conjunctions for the semi-differintegrals. This makes it possible to use the expressions in computer experiments, remembering, of course, to wrap any fork in parentheses before applying it.

Function	Semi-derivative	Semi-integral
$f+g$	$f\ sd + g\ sd$	$f\ si + g\ si$
$f-g$	$f\ sd - g\ sd$	$f\ si - g\ si$
$] * g$	$(] * g\ sd) + -: (g\ si)$	$(] * g\ si) --: (g\ s3i)$
$c"0 * g$	$c"0 * g\ sd$	$c"0 * g\ si$
$c"0$	$c"0 % %:@o.$	$(2 * c)"0 * %:@(] %1p1"0)$
$1"0$	$%@%:@o.$	$+:@%:@%:@o.@%$
$]$	$+:@%:@%:@o.@%$	$4r3"0 * (^&3r2) %1p1r2"0$

<b>*</b> :	$8r^3 \cdot 0^{(3r^2)} \cdot 1p1r^2 \cdot 0$	$16r^{15} \cdot 0^{(5r^2)} \cdot 1p1r^2 \cdot 0$
<b>%</b> :	$1r^2 p1r^2 \cdot 0$	$- : @ (1 \cdot 1p1r^2 \cdot 0)$
<b>%@&gt;</b> :	$(% : @ > : - % : * \_5 \& o . @ % :) . % % : @ o . * > : ^3 r^2 \cdot 0$	$+ : @ ( \_5 \& o . @ % : ) % % : @ ( > : * 1p1 \cdot 0)$
<b>%@%</b> :	$0 \cdot 0$	$% : @ (1p1 \cdot 0)$
<b>% : @&gt;</b> :	$1p1r^2 \cdot 0 \% \sim % : @ \% + \_3 \& o . @ % :$	$1p1r^2 \cdot 0 \% \sim % : + > : * \_3 \& o . @ % :$
<b>%@% : @&gt;</b> :	$% @ ( > : * % : * 1p1r^2 \cdot 0)$	$+ : @ ( \_3 \& o . ) @ % : % 1p1r^2 \cdot 0$
<b>^&amp;p</b>	$% / @ ! @ ( (p-0 \ 1r^2) \cdot 0$ $) * ^ \& (p-1r^2)$	$% / @ ! @ ( (p+0 \ 1r^2) \cdot 0) * ^ \& (p+1r^2)$
<b>^&amp;n</b>	$* : @ ! @ (n \cdot 0) * ^ \& n @ 4 :$ $% ! @ + : @ (n \cdot 0) * % : @ o .$	$* : @ ! @ (n \cdot 0) * ^ \& (n+1r^2) @ 4 :$ $% ! @ > : @ + : @ (n \cdot 0) * 1p1r^2 \cdot 0$
<b>^&amp;(n+1r^2)</b>	$! @ > : @ + : @ (n \cdot 0) * 1p$ $1r^2 \cdot 0 * ^ \& n @ (1r^4 \&$ $*) % + : @ * : @ ! @ (n \cdot 0)$	$! @ > : @ > : @ + : @ (n \cdot 0) * 1p1$ $r^2 \cdot 0 * ^ \& (> : n) @ ( % \& 4)$ $% * : @ ! @ > : @ (n \cdot 0)$
<b>\_3&amp;o.@%</b> :	$- : @ % : @ (1p1 \cdot 0 \% > :)$	$1p1r^2 \cdot 0 * % : \& . > :$

**Notes:**

- f** Function
- g** Function
- n** Integer
- p** Constant greater than  $\_1$
- c** Constant

To experiment with entries in the foregoing table, first enter the definitions of **sd** and **si** and **s3i**, and definitions for **f** and **g** (such as **f** =  $\wedge 3$  and **g** =  $\wedge 2$ ). The first row would then be treated as:

```
(f+g) sd x=: 1 2 3 4
3.29303 14.3887 35.7565 69.404
```

```
(f sd + g sd) x
3.29303 14.3887 35.7565 69.404
```

```
(f+g) si x
1.12591 9.31267 33.7741 85.9827
```

```
(f si + g si) x
1.12591 9.31267 33.7741 85.9827
```

Entries in the table can be rendered more readable to anyone familiar only with conventional notation by a few assignments such as:

```
twice=: +:
sqrt=: %:
pitimes=: o.
reciprocal=: %
on=: @
```

The table entry for the semi-derivative of the identity function could then be expressed as follows:

```

] sd x
1.12697 1.59378 1.95197 2.25394

```

```

twice on sqrt on reciprocal on pimes on reciprocal x
1.12838 1.59577 1.95441 2.25676

```

Alternatively, it can be expressed using the *under* conjunction as follows:

```

under=: &.
twice on sqrt on (pimes under reciprocal) x
1.12838 1.59577 1.95441 2.25676

```





# Chapter 6

## Properties of Functions

### A. Introduction

In this chapter we will analyze relations among the functions developed in Chapter 2, and express them all as members of a single family. We will first attempt to discover interesting relations by experimentation, and then to construct proofs. In this section we will use the growth and decay functions to illustrate the process, and then devote separate sections to experimentation and to proof. We will use the adverb `D=: ("0) (D.1)`.

The reader is urged to try to develop her own experiments before reading Section B, and her own proofs before reading Section C.

In Sections E and F of Chapter 2, the functions `ec` and `eca` were developed to approximate growth and decay functions. Thus:

```

eca=: _1&^ * ec=: %@!
ec i.7
1 1 0.5 0.166667 0.0416667 0.00833333 0.00138889

eca i.7
1 _1 0.5 _0.166667 0.0416667 _0.00833333 0.00138889

```

We will now use the approximate functions to experiment with growth and decay:

```

GR=: (ec i.20) &p.

DE=: (eca i.20) &p.

```

It might be suspected that the decay function would be the reciprocal of the growth function, in other words that their product is *one*. We will test this conjecture in two ways, first by computing the product directly, and then by computing the coefficients of the corresponding product polynomial. Thus:

```

GR x=: 0 1 2 3 4
1 2.71828 7.38906 20.0855 54.5981

DE x
1 0.367879 0.135335 0.0497871 0.0183153
(GR x) * (DE x)

```

```
1 1 1 1 0.999979
```

```
(GR * DE) x
```

```
1 1 1 1 0.999979
```

```
PP=: +//.@(*)
```

```
1 2 1 PP 1 3 3 1
```

```
1 5 10 10 5 1
```

```
6{. (ec i.20) PP (eca i.20)
```

```
1 0 0 _2.77556e_17 6.93889e_18 _1.73472e_18
```

```
6{. (ec PP eca) i.20
```

```
1 0 0 _2.77556e_17 6.93889e_18 _1.73472e_18
```

```
((ec PP eca) i.20) p. x
```

```
1 1 1 1 0.999979
```

Since the growth and decay functions were defined only in terms of their derivatives, any proof of the foregoing conjecture must be based on these defining properties. We begin by determining the derivative of the product as follows:

```
(DE*GR) d.1
```

```
(DE*GR d.1)+(DE d.1 *GR) See Section 2K
```

```
(DE*GR)+(DE d.1 *GR) Definition of GR
```

```
(DE*GR)+(-@DE*GR) Definition of DE
```

```
(DE*GR)-(DE*GR)
```

```
0"0
```

Consequently, the derivative of  $DE*GR$  is zero;  $DE*GR$  is therefore a constant, whose value may be determined by evaluating the function at any point. At the argument 0, all terms of the defining polynomials are zero except the first. Hence the constant value of  $DE*GR$  is *one*, and it is defined by the function  $1"0$ . Thus:

```
(DE*GR) x
```

```
1"0 x
```

```
1 1 1 1 0.999979
```

```
1 1 1 1 1
```

A second experiment is suggested by the demonstration (in Section I of Chapter 2) that the derivative of the function  $f=: ^@(r&*)$  is  $r$  times  $f$ ; the case  $r=: _1$  should give the decay function:

```
r=: _1
```

```
DE x=: 0 1 2 3 4
```

```
1 0.367879 0.135335 0.0497871 0.0183153
```

```
^@(r&*) x
```

```
1 0.367879 0.135335 0.0497871 0.0183156
```

```
^ AM r x
```

```
1 0.367879 0.135335 0.0497871 0.0183156
```

The final expression uses the scaling conjunction of Section I of Chapter 2. We may now conclude that the function  $a^x$  describes growth at any rate, and that negative values of  $x$  subsume the case of decay.

In the foregoing discussion we have used simple observations (such as the probable reciprocity of growth and decay) to motivate experiments that led to the statement and proof of significant identities. To any reader already familiar with the exponential function these matters may seem so obvious as to require neither suggestion nor proof, and he may therefore miss the fact that all is based only on the bare definitions given in Sections 2E and 2F.

Similar remarks apply to the hyperbolic and circular functions treated in Sections 2G,H. The points might be better made by using featureless names such as  $f_1$ ,  $f_2$ , and  $f_3$  for the functions. However, it seems better to adopt commonly used names at the outset.

- A1 Test the proof of this section by entering each expression with an argument.
- A2 Make and display the table  $T$  whose (counter) diagonal sums form the product of the coefficients `ec i.7` and `eca i.7`.

```
[ T=: (ec */ eca) i.7 ]
```

- A3 Denoting the elements of the table `t=: 2 2 {.T` by  $t_{00}$ ,  $t_{01}$ ,  $t_{10}$ , and  $t_{11}$ , write explicit expressions for them. Then verify that  $t_{00}$  and  $t_{01}+t_{10}$  agree with the first two elements of the product polynomial given in the text.

```
[ t00 is 1*1    t01+t10 is (1*_1)+(1*1) ]
```

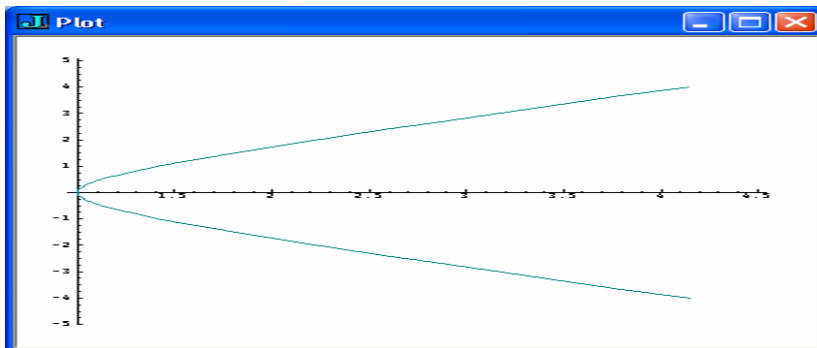
- A4 Use the scheme of A3 on larger subtables of  $T$  to check further elements of the polynomial product.
- A5 Repeat the exercises of this section for other relations between functions that might be known to you.

[Consider the functions `f=: ^^` and `g=: ^@+`: beginning by applying them to arguments such as `f"0 i.5` and `g"0 i. 5`]

## B. Experimentation

**Hyperbolics.** One hyperbolic may be plotted against the other as follows:

```
sinh=: 5&o.
cosh=: 6&o.
load'plot'
plot (cosh;sinh) 0.1*i:21
```



The resulting plot suggests a hyperbola satisfying the equation  $1 = (\text{sqr } x) - (\text{sqr } y)$ . Thus:

```
(*:@cosh - *:@sinh) i:10
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

Finally, each of the hyperbolics is the derivative of the other, and their second derivatives equal the original functions:

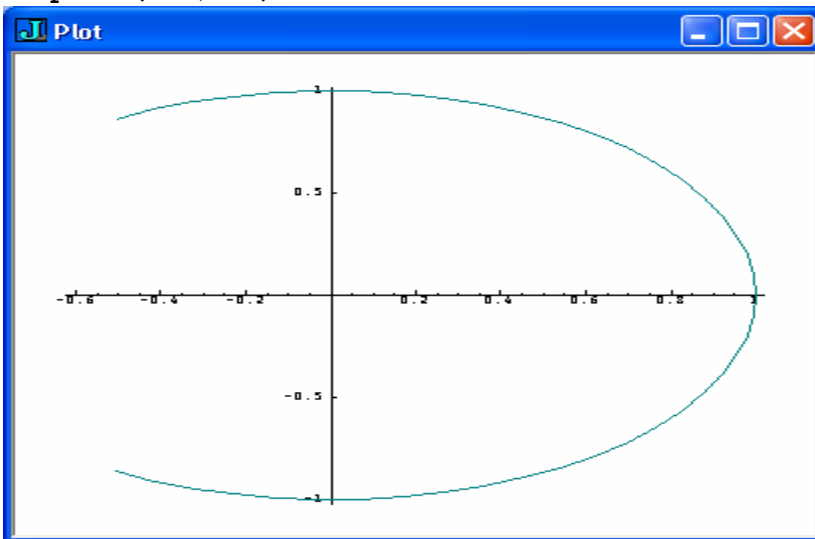
```
sinh x=: 0 1 2 3 4
0 1.1752 3.62686 10.0179 27.2899
cosh x
1 1.54308 3.7622 10.0677 27.3082
```

```
sinh d.1 x
1 1.54308 3.7622 10.0677 27.3082
cosh d.1 x
0 1.1752 3.62686 10.0179 27.2899
```

```
sinh d.2 x
0 1.1752 3.62686 10.0179 27.2899
cosh d.2 x
1 1.54308 3.7622 10.0677 27.3082
```

**Circulars.** The circular functions may be plotted similarly:

```
sin=: 1&o.
cos=: 2&o.
plot (cos;sin) 0.1*i:21
```



The resulting (partial) circle (flattened by scaling) suggests that the following sum of squares should give the result 1 :

```
(*:@cos + *:@sin) i:10
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

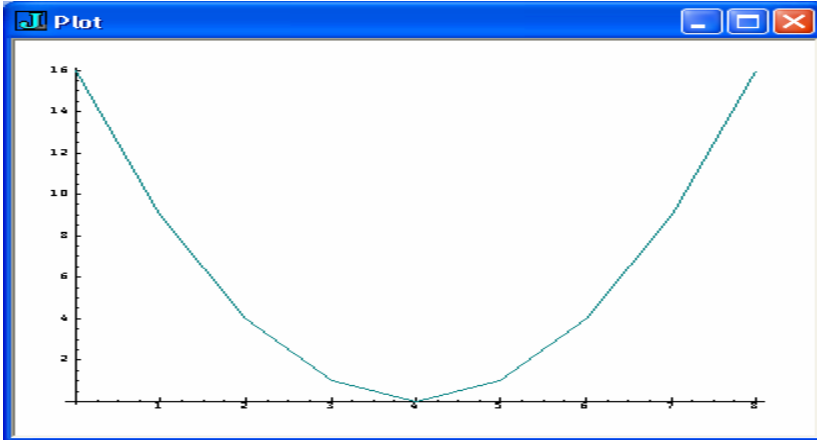
Finally, the derivative of `cos` is `-@sin` and `sin d.1` is `cos`.

**Parity.** If  $f -x$  equals  $f x$  for every value of  $x$ , then  $f$  is said to be *even*. Geometrically, this implies that the plot of  $f$  is reflected in the vertical axis. For example:

```

f=:^&2
x=:0 1 2 3 4
f x
0 1 4 9 16
f -x
0 1 4 9 16
plot f i:4

```

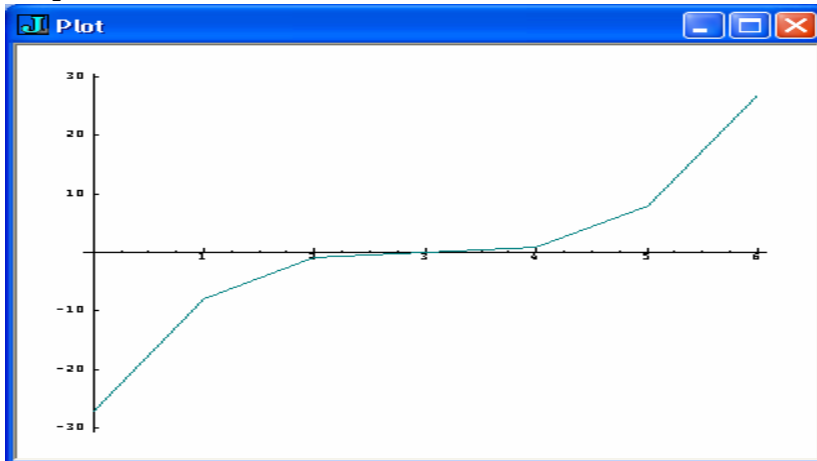


If  $f -x$  equals  $-f x$ , then  $f$  is said to be *odd*, and its plot is reflected in the origin:

```

f=:^&3
f x
0 1 8 27 64
f -x
0 _1 _8 _27 _64
plot f i:3

```



The adverbs:

```

EVEN=: .. -
ODD=: .: -

```

give the even and odd parts of a function to which they are applied; that is,  $f$  **EVEN** is an even function,  $f$  **ODD** is odd, and their sum is equal to  $f$ . For example:

```

^ x
0.0497871 0.135335 0.367879 1 2.71828 7.38906 20.0855

```

$$\begin{aligned} & \text{^ EVEN } x \\ 10.0677 \quad & 3.7622 \quad 1.54308 \quad 1 \quad 1.54308 \quad 3.7622 \quad 10.0677 \end{aligned}$$

$$\begin{aligned} & \text{^ ODD } x \\ \_10.0179 \quad & \_3.62686 \quad \_1.1752 \quad 0 \quad 1.1752 \quad 3.62686 \quad 10.0179 \end{aligned}$$

$$\begin{aligned} & (\text{^EVEN } x) + (\text{^ODD } x) \\ 0.0497871 \quad & 0.135335 \quad 0.367879 \quad 1 \quad 2.71828 \quad 7.38906 \quad 20.0855 \end{aligned}$$

Since the coefficients that define the hyperbolic and circular functions each have zeros in alternate positions, each is either odd or even. The following functions are all *tautologies*, that is, they yield 1 for any argument:

$$\begin{aligned} (\sinh = \sinh \text{ ODD}) & \quad (\sinh = \text{^ ODD}) \\ (\cosh = \cosh \text{ EVEN}) & \quad (\cosh = \text{^ EVEN}) \\ (\sin = \sin \text{ ODD}) & \quad (\cos = \cos \text{ EVEN}) \end{aligned}$$

B1 Repeat Exercises A2-A5 with modifications appropriate to the circular and hyperbolic functions.

## C. Proofs

We will now use the definitions of the hyperbolic and circular functions to establish the two main conjectures of Section B:

$$\begin{aligned} (*:@\cosh - *:@\sinh) \quad & \text{is} \quad 1 \\ (*:@\cos + *:@\sin) \quad & \text{is} \quad 1 \end{aligned}$$

See Section K of Chapter 2 for justification of the steps in the proof:

$$\begin{aligned} & (*:@\cosh - *:@\sinh) \text{ d.1} \\ & (*:@\cosh \text{ d.1} - *:@\sinh \text{ d.1}) \\ & ((*:\text{ d.1 } @\cosh * \sinh) - (*:\text{ d.1 } @\sinh * \cosh)) \\ & ((2^0 * \cosh * \sinh) - (2^0 * \sinh * \cosh)) \\ & (2^0 * ((\cosh * \sinh) - (\sinh * \cosh))) \\ & 0^0 \end{aligned}$$

The circular case differs only in the values for the derivatives:

$$\begin{aligned} \cos \text{ d.1} \quad & \text{is} \quad -\sin \\ \sin \text{ d.1} \quad & \text{is} \quad \cos \end{aligned}$$

C1 Write and test a proof of the fact that the sum of the squares of the functions 1&o. and 2&o. is 1.

## D. The Exponential Family

We have now shown how the growth, decay, and hyperbolic functions can be expressed in terms of the single exponential function  $e^x$ :

<code>^ AM r</code>	Growth at rate <code>r</code>
<code>^ EVEN</code>	Hyperbolic cosine
<code>^ ODD</code>	Hyperbolic sine

Complex numbers can be used to add the circular functions to the exponential family as follows:

<code>^@j. EVEN</code>	Cosine
<code>^@j. ODD</code>	Sine multiplied by <code>0j1</code>

For example:

```

^@j. EVEN x=: 0 1 2 3 4
1 0.540302 _0.416147 _0.989992 _0.653644

cos x
1 0.540302 _0.416147 _0.989992 _0.653644

^@j. ODD x
0 0j0.841471 0j0.909297 0j0.14112 0j_0.756802

j. sin x
0 0j0.841471 0j0.909297 0j0.14112 0j_0.756802

j.^:_1 ^@j. ODD x
0 0.841471 0.909297 0.14112 _0.756802

^ ODD &. j. x
0 0.841471 0.909297 0.14112 _0.756802

```

D1 Write and test tautologies involving `cosh` and `sinh`.

```
[ t=: cosh = sinh@j. and u=: sinh = cosh@j. ]
```

D2 Repeat D1 for `cos` and `sin`.

## E. Logarithm and Power

The inverse of the exponential is called the *logarithm*, or *natural logarithm*. It is denoted by `^.`; some of its properties are shown below:

```

I=: ^:_1                                Inverse adverb

^ I x=: 1 2 3 4 5
0 0.693147 1.09861 1.38629 1.60944

^ ^ I x
1 2 3 4 5

^. x                                      Natural log
0 0.693147 1.09861 1.38629 1.60944

```

```

^ . d.1 x
1 0.5 0.333333 0.25 0.2
% ^ . d.1 x
1 2 3 4 5

```

```

^ . x ^ b=: 3
0 2.07944 3.29584 4.15888 4.82831

```

```

b * ^ . x
0 2.07944 3.29584 4.15888 4.82831

```

The dyadic case of the logarithm  $\wedge.$  is defined in terms of the monadic as illustrated below:

```

(^.x) % (^b)
0 0.63093 1 1.26186 1.46497

```

```

b ^ . x
0 0.63093 1 1.26186 1.46497

```

```

b %&^~ x
0 0.63093 1 1.26186 1.46497

```

The dyadic case of  $\wedge$  is the power function; it has, like other familiar dyads (including  $+$   $-$   $*$   $\%$ ) been used without definition. We now define it in terms of the dyadic logarithm as illustrated below:

```

b&^ . I x          (Where I=: ^:_1 is the inverse adverb)
3 9 27 81 243

```

```

b ^ x
3 9 27 81 243

```

This definition extends the domain of the power function beyond the non-negative integer right arguments embraced in the definition of power as the product over repetitions of the left argument, as illustrated below:

```

m=: 1.5
n=: 4
n # m
1.5 1.5 1.5 1.5

```

```

*/ n # m
5.0625

```

```

m^n
5.0625

```

Moreover, the extended definition retains the familiar properties of the simple definition. For example:



$5^4 + 3$   
 78125  
 $(5^4) * (5^3)$   
 78125

E1 Comment on the question of whether the equivalence of  $a^n = n^a$  and  $m^n$  holds for the case  $n = 0$ .

## F. Trigonometric Functions

Just as a five-sided (or five-angled) figure may be characterized either as *pentagonal* or *pentangular*, so may a three-sided figure be characterized as *trigonal* or *triangular*. The first of these words suggests the etymology of *trigonometry*, the measurement of three-sided figures. This section concerns the equivalence of the functions **sin** and **cos** (that have been defined only by differential equations) and the corresponding trigonometric functions *sine* and *cosine*.

The sine and cosine are also called *circular* functions, because they can be defined in terms of the coordinates of a point on a *unit* circle (with radius 1 and centre at the origin) as functions of the length of arc to the point, measured counter-clockwise from the reference point with coordinates  $(1, 0)$ . As illustrated in Figure F1, the cosine of  $a$  is the horizontal (or  $x$ ) coordinate of the point whose arc is  $a$ , and the sine of  $a$  is the vertical coordinate.

The length of arc is also called the *angle*, and the ratio of the circumference of a circle to its diameter is called  $\pi$ , given by  $\pi = \frac{C}{d}$ , or by the constant  $\frac{1}{2}$ . The circular functions therefore have the period  $2\pi$ , that is *two pi*. Moreover, the coordinates of the end points of arcs of lengths  $\pi$  and  $0.5\pi$  are  $(-1, 0)$  and  $(0, 1)$ ; the *supplementary* angle  $\pi - a$  and the *complementary* angle  $0.5\pi - a$  are found by moving *clockwise* from these points.

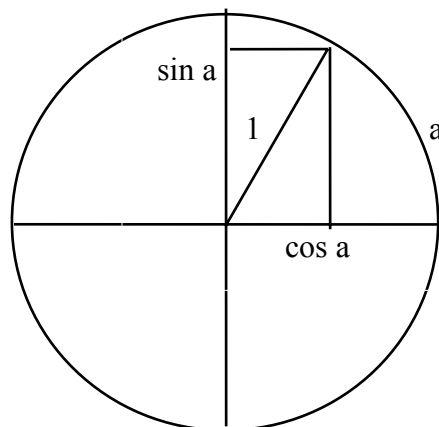


Figure F1

Taken together with these remarks, the properties of the circle make evident a number of useful properties of the sine and cosine. We will illustrate some of them below by tautologies, each of which can be tested by enclosing it in parentheses and applying it to an argument, as illustrated for the first of them:

$\sin^2 + \cos^2 = 1$ .

$C = 2 \cos$   
 $x = 1 \ 2 \ 3 \ 4 \ 5$   
 $(1^0 = \sin + \cos) \ x$   
 $1 \ 1 \ 1 \ 1 \ 1$

Theorem of Pythagoras

$\sin(-) = -\sin$  The sine is odd

$\sin \text{ ODD} = \sin$

$\cos(-) = \cos$  The cosine is even

$\sin(2\pi +) = \sin$  The period of the sine is twice pi

$\cos(2\pi +) = \cos$  The period of the cosine is twice pi

$\sin(\pi -) = \sin$  Supplementary angles

$\cos(\pi -) = -\cos$  "

$\sin(0.5\pi -) = \cos$  Complementary angles

$\cos(0.5\pi -) = \sin$  "

**Sum Formulas.** A function applied to a sum of arguments may be expressed equivalently in terms of the function applied to the individual arguments; the resulting relation is called a *sum formula*:

$a = 2 \ 3 \ 5 \ 7$   
 $b = 4 \ 3 \ 2 \ 1$   
 $+: a+b$   
 $12 \ 12 \ 14 \ 16$

$(+:a) + (+:b)$   
 $12 \ 12 \ 14 \ 16$

$*: a+b$   
 $36 \ 36 \ 49 \ 64$

$(*:a) + (+:a*b) + (*:b)$   
 $36 \ 36 \ 49 \ 64$

$^ a+b$   
 $403.429 \ 403.429 \ 1096.63 \ 2980.96$

$(^a) * (^b)$   
 $403.429 \ 403.429 \ 1096.63 \ 2980.96$

Sum formulas may also be expressed as tautologies:

$+:@+ = +:@[ + +:@]$   
 $a(+:@+ = +:@[ + +:@]) \ b$   
 $1 \ 1 \ 1 \ 1$

$*:@+ = *:@[ + +:@* + *:@]$   
 $^@+ = ^@[ * ^@]$

The following sum formulas for the sine and cosine are well-known in trigonometry:

$$\sin(\alpha + \beta) = (\sin \alpha \cos \beta) + (\cos \alpha \sin \beta)$$

$$\sin(\alpha - \beta) = (\sin \alpha \cos \beta) - (\cos \alpha \sin \beta)$$

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

$$\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$$

Since  $\sin 2\alpha$  is equivalent to (the monadic)  $\sin 2\alpha$ , we may obtain the following identities for the double angle:

$$\sin 2\alpha = 2 \sin \alpha \cos \alpha$$

$$\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha$$

The theorem of Pythagoras can be used to obtain two further forms of the identity for  $\cos 2\alpha$ :

$$\cos 2\alpha = 1 - 2 \sin^2 \alpha$$

$$\cos 2\alpha = 2 \cos^2 \alpha - 1$$

An identity for the sine of the half angle may be obtained as follows:

$$(\cos 2\alpha - 1) = 1 - 2 \sin^2 \alpha$$

$$(\cos 2\alpha - 1) = 1 - 2 \sin^2 \alpha$$

$$(2 \sin^2 \alpha - 1) = 1 - \cos 2\alpha$$

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

$$(\sin \alpha = \sqrt{\frac{1}{2} (1 - \cos 2\alpha)})$$

The last two tautologies above compare *magnitudes* ( $=\&|$ ) because the square root yields only the positive of the two possible roots. Similarly for the cosine:

$$(\cos 2\alpha - 1) = 2 \cos^2 \alpha - 1$$

$$(\cos 2\alpha - 1) = 2 \cos^2 \alpha - 1$$

$$(\cos 2\alpha - 1) = 2 \cos^2 \alpha - 1$$

Tautologies may be re-expressed in terms of arguments  $i$  and  $x$  as illustrated below for  $\sin 2\alpha$  and  $\cos 2\alpha$ :

$$i = 0.1$$

$$(\sin i+x) = ((\sin i) * (\cos x)) + ((\cos i) * (\sin x))$$

$$(\cos i+x) = ((\cos i) * (\cos x)) - ((\sin i) * (\sin x))$$

Derivatives. Using the results of Section 2A, we may express the secant slope of the sine function at the points  $x$  and  $i+x$  as follows:

$$((S_{i+x}) - (S_x)) / i$$

Using the sum formula for the sine we obtain the following equivalent expressions:

$$\begin{aligned} & ((S_i)(C_x) + (C_i)(S_x) - (S_x)) / i \\ & ((S_i)(C_x) + (S_x)(C_i)) / i \\ & ((S_i)(C_x)) / i - (S_x)((1-C_i)) / i \\ & ((C_x)(S_i)) / i - (S_x)((1-C_i)) / i \end{aligned}$$

To obtain the derivative of  $S$  from this secant slope, it will be necessary to obtain limiting values of the ratios  $(S_i) / i$  and  $(1-C_i) / i$ .

In the unit circle of Figure F2, the magnitude of the area of the sector with arc length (angle in radians)  $i$  lies between the areas of the triangles OSC and OST. Moreover, the lengths of the relevant sides are as shown below:

OC	CS	OS	ST
$C_i$	$S_i$	1	$(S/C)_i$

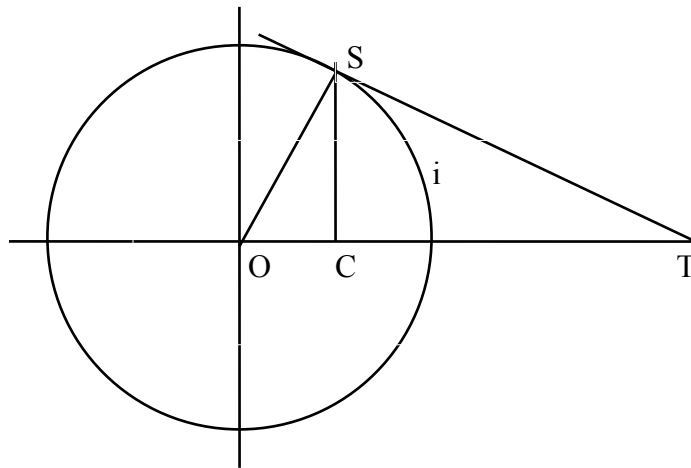


Figure F2

ST is the tangent to the circle, and its length is called the *tangent* of  $i$ . Its value  $(S/C)_i$  follows from the ratios in the similar triangles.

The values of the cited areas are therefore  $-\frac{1}{2}(S \cdot C)_i$  and  $-\frac{1}{2}i$  and  $-\frac{1}{2}(S/C)_i$ . Multiplying by 2 and dividing by  $S_i$  gives the relative sizes  $C_i$  and  $i/S_i$  and  $(S/C)_i$ . Hence, the ratio  $i/S_i$  lies between  $C_i$  and  $(S/C)_i$ , both of which are 1 if  $i=0$ . Finally, the desired limiting ratio  $(S_i) / i$  is the reciprocal, also 1.

The limiting value of  $(1-C_i) / i$  is given by the identity  $\frac{d}{dx}(1-\cos x) = \sin x$ , for

$$\begin{aligned} & (1-C_i) / i \\ & (\frac{d}{dx}(1-\cos x)) / \frac{dx}{dx} \\ & (\sin x) / 1 \end{aligned}$$

$$((S \ i^2)\%i^2) * (S \ i^2)$$

The limit of the first factor has been shown to be 1, and the limit of  $S \ i^2$  is 0; hence the limit of  $(1-C \ i)\%i$  is their product, that is, 0.

Substituting these limiting values in the expression for the secant slope  $((C \ x)*((S \ i)\%i)) - (S \ x)*((1-C \ i)\%i)$  we obtain the expression for the derivative of the sine, namely:

$$\frac{(C \ x)*(1) - (S \ x)*(0)}{C \ x}$$

Similar analysis shows that the derivative of  $C$  is  $-S$ , and we see that the relations between  $S$  and  $C$  and their derivatives are the same as those between  $\sin$  and  $\cos$  and their derivatives. Moreover, the values of  $S$  and  $\sin$  and of  $C$  and  $\cos$  agree at the argument 0.

F1 Define  $f =: \sin @ (+/) = \text{perm} @: sc$  and  $sc =: 1 \ 2 \&o. "0$  and  $\text{perm} =: +/ \ . \ *$  and  $\sin =: 1 \&o.$  and  $\cos =: 2 \&o.$ ; then evaluate  $f \ a, b$  for various scalar values of  $a$  and  $b$  and comment on the results.

[  $f$  is a tautology recognizable as  
 $(\sin(a+b)) = ((\sin a) * (\cos b)) + ((\cos a) * (\sin b))$  ]

F2 Define other tautologies known from trigonometry in the form used in F1.

[ Consider the use of  $\text{det} =: -/ \ . \ *$  ]

## G. Dot and Cross Products

As illustrated in Section 3E, the vector derivative of the function  $*/\backslash$  yields a matrix result; the vectors in this matrix lie in a plane, and the vector *perpendicular* or *normal* to this plane is an important derivative called the *curl* of the vector function. We will now present a number of results needed in its definition, including the *dot* or *scalar* product and the *cross* or *vector* product.

The angle between two rays from the origin is defined as the length of arc between their intersections with a circle of unit radius centred at the origin. The angle between two vectors is defined analogously. For example, the angle between the vectors  $3 \ 3$  and  $0 \ 2$  is  $1r4p1$  (that is, one-fourth of  $\pi$ ) radians, or 45 degrees. If the angle between two vectors is  $1r2p1$  radians (90 degrees), they are said to be *perpendicular* or *normal*.

Similar notions apply in three dimensions, and a vector  $\mathbf{r}$  that is normal to each of two vectors  $\mathbf{p}$  and  $\mathbf{q}$  is said to be normal to the plane defined by them, in the sense that it is normal to every vector of the form  $(\mathbf{a} * \mathbf{p}) + (\mathbf{b} * \mathbf{q})$ , where  $\mathbf{a}$  and  $\mathbf{b}$  are scalars.

The remainder of this section defines the dot and cross products, and illustrates their properties. Proofs of these properties may be found in high-school level texts as, for example, in Sections 6.7, 6.8, and 6.12 of Coleman et al [11]. Again we will leave interpretations to the reader, and will defer comment on them to exercises.

The dot product may be defined by  $+/@*$  or, somewhat more generally, by  $+/ \ . \ *$ . Thus:

$$\mathbf{a} =: 1 \ 2 \ 3 \quad [ \quad \mathbf{b} =: 4 \ 3 \ 2 \\ +/\mathbf{a} * \mathbf{b}$$

```

dot=: +/ . *
a dot b
16
The product of the cosine of the angle between a
and b with the product of their lengths

dot~ a
14
Squared length of a

L=: %:@(dot~)"1
a,:b
1 2 3
4 3 2
Length function

L a,:b
3.74166 5.38516

*/ L a,:b
20.1494
Product of lengths

(a dot b) % */L a,:b
0.794067

cos=:dot % */@(L@,:)
a cos b
0.794067
Re-definition of cos (not of 2&o.)

0 0 1 cos 0 1 0
0
Perpendicular or normal vectors

0 0 1 cos 0 1 1
0.707107

2 o. 1r4p1
0.707107
Cosine of 45 degrees

```

The following expressions lead to a definition of the cross product and to a definition of the sine of the angle between two vectors:

```

rot=: |."0 1
1 _1 rot a
2 3 1
3 1 2
Rotation of vectors

(1 _1 rot a) * (_1 1 rot b)
4 12 3
9 2 8

]c=:-(1 _1 rot a)*(_1 1 rot b)
_5 10 _5
Cross product

a dot c
0
The vectors are each normal to
their cross product

b dot c
0

```

```

cross=: -/@(1 _1&rot@[ * _1 1&rot@])
a cross b
_5 10 _5

```

```

(a,:b) dot a cross b
0 0

```

```

b cross a
5 _10 5

```

The cross product is not commutative

```

L a cross b
12.2474

```

The product of the sine of the angle between the vectors with the product of their lengths

```

(L a cross b) % */ L a,:b
0.607831

```

The sine of the angle

```

sin=: L@cross % */@(L@,: )
a sin b
0.607831

```

The sine function

```

a +/ @: * : @ (sin , cos) b
1

```

The following expressions suggest interpretations of the dot and cross products that will be pursued in exercises:

```

c=: 4 1 2
c dot a cross b
_20

```

```

m=: c,a,:b
m
4 1 2
1 2 3
4 3 2

```

```

- / . * m
_20

```

- G1 Experiment with the dot and cross products, beginning with vectors in 2-space (that is with two elements) for which the results are obvious. Continue with other vectors in 2-space and in 3-space. Sketch the rays defined by the vectors, showing their intersection with the unit circle (or sphere).

## H. Normals

We now use the function `e` introduced in Section 3E to define a function `norm` that is a generalization of the cross product; it applies to arrays other than vectors, and produces a result that is normal to its argument. Moreover, when applied to skew arrays of odd order (having an odd number of items) it is self-inverse. Thus:

```

indices=: { @ ( ) # < @ i . )
e=: C.!.2 @ > @ indices

```

Result is called an "e-system" by McConnell [4]

```

]skm=: * : . : | : i . 3 3
0 _4 _16
4 0 _12
16 12 0

```

A skew matrix

```

]v=: -: +/ +/ skm * e #skm
_12 16 _4

```

```
v +/ . * skm
0 0 0
```

Test of orthogonality

```
+/ v * e #v
0 _4 _16
4 0 _12
16 12 0
```

Inverse transformation

```
norm=: +/^: (]`(#@$)`(* e@#)) % !@(#-#@$)
]m=: (a=: 1 2 3) */ (b=: 4 3 2)
4 3 2
8 6 4
12 9 6
```

```
n=: norm ^:
0 n m
4 3 2
8 6 4
12 9 6
```

An adverb for powers of `norm`

```
1 n m
_5 10 _5
```

```
a cross b
_5 10 _5
```

```
2 n m
0 _2.5 _5
2.5 0 _2.5
5 2.5 0
3 n m
_5 10 _5
```

Skew part of `m`

```
1 n a
0 1.5 _1
_1.5 0 0.5
1 _0.5 0
```

```
2 n a
1 2 3
```

Self-inverse for odd dimension

```
mp=: +/ . *
a mp 1 n a*/b
0
```

```
b mp 1 n a*/b
0
```

```
x=: 1 2
1 n x
_2 1
```

```
x mp 1 n x
0
```

```
2 n x
_1 _2
```

For even orders `2 n` is inverse only up to sign change



$$\begin{vmatrix} 4 & n & x \\ 1 & 2 & \end{vmatrix}$$

$$\begin{vmatrix} 2 & n & y \\ 1 & 2 & 3 & 4 \\ \_1 & \_2 & \_3 & \_4 \end{vmatrix}$$

$$\begin{vmatrix} 4 & n & y \\ 1 & 2 & 3 & 4 \end{vmatrix} \quad \begin{vmatrix} 2 & n & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix}$$

$$\begin{vmatrix} \mathbf{cr} \\ \mathbf{a} & \mathbf{cr} & \mathbf{b} \\ \_5 & 10 & \_5 \end{vmatrix}$$

Alternative definition of cross product

$$\begin{vmatrix} \mathbf{a} & \mathbf{cross} & \mathbf{b} \\ \_5 & 10 & \_5 \end{vmatrix}$$

- H1 Experiment with the expressions of this section.
- H2 Using the display of  $\mathbf{e}_3$  shown in Section 3E, and using  $\mathbf{a}_0$ ,  $\mathbf{a}_1$ , and  $\mathbf{a}_2$  to denote the elements of a vector  $\mathbf{a}$  in 3-space, show in detail that  $\mathbf{norm}(\ast/)$  is indeed an alternative definition of the cross product.
- H3 Show in detail that  $\ast/(\mathbf{e}_i \# \ast/)$  is an alternative definition of the determinant.



# Chapter 7

## Interpretations and Applications

### A. Introduction

As remarked in Section 3A, various interpretations of a particular function definition are possible (as in `vol=: */` and `cost=: */`), and any one of them may be either helpful or confusing, depending upon the background of the reader. A helpful interpretation may also be misleading, either by suggesting too little or too much. We will illustrate this point by three examples.

**Example 1.** The sentences:

```
S=: 2 : '%&x. @ (] -&y. -&x.)'
f=: ^
h=: 1e_8
sf=: h S f
sf x=: 1
2.71828
```

define and use the function `sf`. Moreover, `sf` can be helpfully interpreted as the secant slope of the exponential with spacing `h`, and (because `h` is small) as an approximation to the tangent slope of the exponential.

However, for the case of the discontinuous *integer part* function `<`, this interpretation would be misleading because its "tangent slope" at the point `1` is infinite. Thus:

```
h S <. x
1e8
```

**Example 2.** If the spacing `h` is complex, the function `h S ^` has the behaviour expected of a secant slope:

```
^ y=: 2j3
_7.31511j1.04274

h=: 1e_6j1e_8
h S ^ y
```

```
_7.31511j1.04274
```

```
(r=: 1e_6j0) s ^ y
_7.31511j1.04274
```

```
(i=: 0j1e_8) s ^ y
_7.31511j1.04274
```

Again the interpretation of the function  $h \ s \ f$  as an approximation to the tangent slope is valid. However, the (continuous) *conjugate* function  $+$  shows unusual behaviour:

```
h s + y          r s + y          i s + y
0.9998j_0.019998  1                _1
```

The problem arises because the conjugate is not an *analytic* function. A clear and simple discussion of this matter may be found in Churchill [12].

### Example 3.

Section 2D interprets the integral of a function  $f$  as a function that gives the area under the graph of  $f$  from a point  $a$  (that is, the point  $a, f \ a$  on the graph of  $f$ ) to a second point  $b$ . This interpretation is helpful for real-valued functions, but how should we visualize the area under a function that gives a complex result?

It is, of course, possible to interpret the integral as a complex result whose real and imaginary parts are the areas under the real and imaginary parts of  $f$ , respectively. However, the beginning and end points may themselves be complex, and although there is a clearly defined "path" through real numbers between a pair of real numbers  $a$  and  $b$ , there are an infinity of different paths through complex numbers from complex  $a$  to complex  $b$ .

This observation leads to the more difficult, but highly useful, notion of integration along a prescribed path (called a *line* or *contour* integral), a notion not hinted at by the interpretation of integration as the area under a curve.

## B. Applications and Word Problems

What we have treated as *interpretations* of functions may also be viewed as *applications* of math, or as *word problems* in math. For example, if  $\cos=:2&o.$  and  $\sin=:1&o.$ , then the function:

```
f=:0.1&path=: (cos, sin)@*^0
```

may be *interpreted* as the "Position of a car ... moving on a circular path at an angular velocity of 0.1 radians per second". Conversely, the expression in quotes could be considered as an *application* of the circular functions, and could be posed as a *word problem* requiring as its solution a definition of the function  $f$ .

Similarly, the phrase  $f \ D.1$  may be interpreted as the velocity of the car whose position is prescribed by  $f$ . Because the phrase involved a derivative, the corresponding word problem would be considered as an application of the calculus.

Just as a reader's background will determine whether a given interpretation is helpful or harmful in grasping new concepts in the calculus, so will it determine the utility of word problems. We will limit our treatment of interpretations and applications to a few examples, and encourage the reader to choose further applications from any field of interest, or from other calculus texts.

## C. Extrema and Inflection Points

If  $\mathbf{f} = (\mathbf{c} = 0 \ 1 \ 2.5 \ 2 \ 0.25) \&\mathbf{p}.$ , then  $\mathbf{p}.$  is a polynomial in terms of coefficients, and  $\mathbf{f}$  is a specific polynomial whose (tightly) formatted results:

```
(fmt=: 5.1&" :) f x=: 0.1*>:i.6 10
 0.1 0.3 0.5 0.7 0.9 1.1 1.3 1.5 1.6 1.8
 1.8 1.9 1.8 1.8 1.6 1.4 1.2 0.9 0.5 0.0
_0.5 _1.1 _1.8 _2.6 _3.4 _4.2 _5.2 _6.1 _7.2 _8.2
_9.4 _10.5 _11.7 _12.9 _14.1 _15.3 _16.5 _17.7 _18.9 _20.0
_21.1 _22.1 _23.0 _23.9 _24.6 _25.2 _25.7 _26.1 _26.3 _26.2
_26.0 _25.6 _25.0 _24.1 _22.9 _21.4 _19.6 _17.4 _14.9 _12.0
```

suggest that it has a (local) maximum (of 1.9) near 1.2 and a minimum near 4.9. Moreover, a graph of the function over the interval from 0 to 4 shows their location more precisely.

A graph of the derivative  $\mathbf{f} \ \mathbf{d}.1$  over the same interval illustrates the obvious fact that the derivative is zero at an extremum (minimum or maximum):

```
fmt f d.1 x
 1.4 1.8 2.0 2.1 2.1 2.1 1.9 1.7 1.4 1.0
 0.6 0.1 _0.4 _1.0 _1.6 _2.3 _2.9 _3.6 _4.3 _5.0
_5.7 _6.4 _7.1 _7.7 _8.4 _9.0 _9.6 _10.1 _10.6 _11.0
_11.4 _11.7 _11.9 _12.1 _12.1 _12.1 _12.0 _11.8 _11.4 _11.0
_10.4 _9.8 _8.9 _8.0 _6.9 _5.6 _4.2 _2.6 _0.9 1.0
 3.1 5.4 7.8 10.5 13.4 16.5 19.8 23.3 27.0 31.0
```

We may therefore determine the location of an extremum by determining the *roots* (arguments where the function value is zero) of the derivative function. Since we are concerned only with real roots we will define a simple adverb for determining the value of a root in a specified interval, where the function values at the ends of the interval must differ in sign. The method used is sometimes called the *bisection* method; the interval is repeatedly halved in length by using the midpoint (that is, the mean) together with that endpoint for which the function value differs in sign. Thus:

```
m=: +/ % #
bis=: 1 : '2&{.@(m , ] #~ m ~: &(*@x.) ])'
f y=: 1 4                                Interval that bounds a root of f
1.75 _20

f bis y                                    One step of the bisection method
2.5 1
```

```

f f bis y
_3.35938 1.75

```

Resulting interval still bounds a root

```

f bis^:0 1 2 3 4 y
  1      4
 2.5    1
 1.75   2.5
 2.125  1.75
 1.9375 2.125

```

Successive bisections

```

f bis^:_ y
2 2

```

Limit of bisection

```

]root=: m f bis^:_ y
2

```

Root is mean of final interval

```

f root
_3.55271e_14

```

A root of the derivative of  $f$  identifies an extremum of  $f$ :

```

f d.1 z=: 0.5 1.5
2.125 _1.625

```

```

]droot=: m f d.1 bis^:_ z
1.21718

```

```

f d.1 droot
_9.52571e_14

```

When the derivative of  $f$  is increasing, the graph of  $f$  bends upward; when the derivative is decreasing, it bends downward. At a maximum (or minimum) point of the derivative, the graph of  $f$  therefore changes its curvature, and the graph crosses its own tangent. Such a point is called a *point of inflection*.

Since an extremum of the derivative occurs at a zero of *its* derivative, an inflection point of  $f$  occurs at a zero of  $f$  d.2 . Thus:

```

fmt f d.2 x
 3.8  2.7  1.7  0.7  _0.3  _1.1  _1.9  _2.7  _3.4  _4.0
_4.6  _5.1  _5.5  _5.9  _6.3  _6.5  _6.7  _6.9  _7.0  _7.0
_7.0  _6.9  _6.7  _6.5  _6.2  _5.9  _5.5  _5.1  _4.6  _4.0
_3.4  _2.7  _1.9  _1.1  _0.2  0.7  1.7  2.7  3.8  5.0
_6.2  7.5  8.9 10.3 11.8 13.3 14.9 16.5 18.2 20.0
21.8 23.7 25.7 27.7 29.8 31.9 34.1 36.3 38.6 41.0

```

```

* f (d.2) 0 1
1 _1

```

```

]infl=: m f d.2 bis^:_ (0 1)
0.472475

```

```

f d.2 infl
_6.83897e_14

```

A graph of  $f$  will show that the curve crosses its tangent at the point `infl`.

C1 Test the assertion that `droot` is a local minimum of `f`.

```
[ f droot + _0.0001 0 0.0001
```

It is not a minimum, but a maximum. ]

C2 What is the purpose of `2&{.e}` in the definition of `bis`?

```
[ Remove the phrase and try f bis 1 3 ]
```

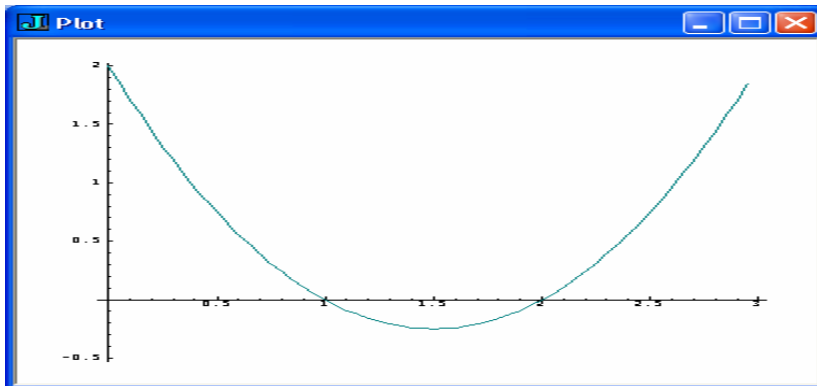
C3 For various coefficients `c`, make tables or graphs of the derivative `c&p`. `D` to determine intervals bounding roots, and use them with `bis` to determine extrema of the polynomial `c&p`.

## D. Newton's Method

Although the bisection method is certain to converge to a root when applied to an interval for which the function values at the endpoints differ in sign, this convergence is normally very slow. The derivative of the function can be used in a method that normally converges much faster, although convergence is assured only if the initial guess is "sufficiently near" the root.

The function `g=: (]-1:)* (]-2:)` has roots at 1 and 2, as shown by its graph:

```
plot y;g y=: 1r20*i.60
```



Draw a tangent at the point `x,g` `x=: 3` intersecting the axis at a point `nx,0` and note that `nx` is a much better approximation to the nearby root at 2 than is `x`. The length `x-nx` is the *run* that produces the *rise* `g x` with the *slope* `d.1 x`. As a consequence, `nx=:x-(g x) % (g d.1 x)` is a better approximation to the root at 2. Thus:

```
x=:3
g=: (]-1:)* (]-2:)
g x
2
]nx=:x-(g x) % (g d.1 x)
2.33333
g nx
0.444444
```

A root can be determined by repeated application of this process, using an adverb `N` as follows:

```
N=: (1 : ']- x. % x. d.1') (^:_)
```

```
f=: (c=: 0 1 2.5 _2 0.25) &p.
```

Used in Section C

```
f N 6
6.31662
```

```
f f N 6
7.01286e_16
```

Test if  $f N 6$  is a root of  $f$

```
f N x=: i. 7
0 _0.316625 2 2 2 6.31662 6.31662
```

Different starts converge  
to different roots

```
f f N x
0 0 0 0 0 7.01286e_16 7.01286e_16
```

This use of the derivative to find a root is called *Newton's method*. Although it converges rapidly near a single root, the method may not converge to the root nearest the initial guess, and may not converge at all. The initial guess `droot` determined in the preceding section as a maximum point of  $f$  illustrates the matter; the derivative at the point is approximately 0, and division by it yields a very large value as the next guess:

```
f N droot
6.31662
```

Since the derivative of a polynomial function `c&p.` can be computed directly using the coefficients `.c*i.#c`, it is possible to define a version of Newton's method that does not make explicit use of the derivative adverb. Thus:

```
dc=: 1 : '].@[ * i.@#)@(x."_) p. ]' ("0)
NP=: 1 : ']-x.&p. % x. dc' ("0) (^:_
c NP x
0 _0.316625 2 2 2 6.31662 6.31662
```

```
c&p. N x
0 _0.316625 2 2 2 6.31662 6.31662
```

The following utilities are convenient for experimenting with polynomials and their roots:

```
pir=:<@[ p. ]
_1 _1 _1 pir x
1 8 27 64 125 216 343
```

Polynomial in terms of roots

```
1 3 3 1 p. x
1 8 27 64 125 216 343
```

```
pp=: +//.@(*/)
1 2 1 pp 1 3 3 1
1 5 10 10 5 1
```

Polynomial product

```
(1 2 1 pp 1 3 3 1) p. x
1 32 243 1024 3125 7776 16807
(1 2 1 p. x) * (1 3 3 1 p. x)
1 32 243 1024 3125 7776 16807
```

```
cfr=: pp/@(- ,. 1:)
cfr _1 _1 _1
```

Coefficients from roots



1 3 3 1

D1 Use Newton's method to determine the roots for which the bisection method was used in Section C.

## E. Kerner's Method

Kerner's method for the roots of a polynomial is a generalization of Newton's method; at each step it treats an  $n$ -element list as an approximation to all of the  $<:\#c$  roots of the polynomial  $c\&p.$ , and produces an "improved" approximation. We will first define and illustrate the use of an adverb  $\kappa$  such that  $c \kappa b$  yields the  $<:\#c$  (or  $\#b$ ) roots of the polynomial with coefficients  $c$ :

```

k=: 1 : ']-x.&p. % (<0 1)&|:@((1&(*/\."1))@(-/~))'
K=: k (^:_ )
b=: 1 2 3 4
]c=: cfr b+0.5
59.0625 _93 51.5 _12 1

```

Coefficients of polynomial  
with roots at  $b+0.5$

```

c k b
2.09375 2.46875 3.28125 4.15625

```

Single step of Kerner

```

c K b
1.5 2.5 3.5 4.5

```

Limit of Kerner  
Roots of  $c\&p.$

```

c k ^: (i.7) b
      1      2      3      4
2.09375 2.46875 3.28125 4.15625
1.20508 2.59209 3.7207 4.48213
1.45763 2.53321 3.50503 4.50413
1.49854 2.50154 3.49996 4.49997
      1.5      2.5      3.5      4.5
      1.5      2.5      3.5      4.5

```

Six steps of Kerner

```

]rb=: 4?.20
17 4 9 7

```

Random starting value

```

c K rb
1.5 3.5 4.5 2.5

```

The adverb  $\kappa$  applies only to a *normalized* coefficient  $c$ , that is, one whose last non-zero element (for the highest order term) is 1. Thus:

```

norm=: (] % {:)@(>./\.@:|@:* # ])
norm 1 2 0 3 4 0 0
0.25 0.5 0 0.75 1

```

The polynomials  $c\&p.$  and  $(norm\ c)\&p.$  have the same roots, and  $norm\ c$  is a suitable argument to the adverb  $\kappa$ .

Kerner's method applies to polynomials with complex roots; however it will not converge to complex roots if the beginning guess is completely real: **begin** provides a suitable beginning argument:

```
(begin=: %:@-@i.@<:@#) 1 3 3 1
0 0j1 0j1.41421
```

For example, the coefficients `d=: cfr 1 2 2j3 4 2j_4` define a polynomial with two complex roots. Thus:

```
d=: cfr 1 2 2j3 4 2j_4
]roots=: (norm d) K begin d
4 2j3 2j_4 2 1

/:~roots                               Sorted roots
1 2j3 2j_4 2 4
```

The definition of the adverb `k` (for a single step of Kerner) can be revised to give an alternative equivalent adverb by replacing the division (`%`) by matrix division (`%.`), and removing the phrase `<0 1)&|:@` that extracts the diagonal of the matrix produced by the subsequent phrase. Thus:

```
ak=: 1 : ']-x.&p. % . ((1&(*\."1))@(-/~))'
c ak b
2.09375 2.46875 3.28125 4.15625
```

In this form it is clear that the vector of *residuals* produced by `x.&p.` (the values of the function applied to the putative roots, which must all be reduced to zero) is divided by the matrix produced by the expression to the right of `%.`. This expression produces the vector derivative with respect to each of the approximate roots; like the analogous case of the direct calculation of the derivative in the adverb `NP` it is a direct calculation of the derivative without explicit use of the vector derivative adverb `VD=: ("1) (D. 1)`. These matters are left for exploration by the reader.

- E1 Find all roots of the functions used in Section C.
- E2 Define some polynomials that have complex roots, and use Kerner's method to find all roots.

## F. Determinant and Permanent

The function `-/ . *` yields the *determinant* of a square matrix argument. For example:

```
det=: -/ . *
]m=: >3 1 4;2 7 8;5 1 6
3 1 4
2 7 8
5 1 6

det m
_2
```

The determinant is a function of rank 2 that produces a rank 0 result; its derivative is therefore a rank 2 function that produces a rank 2 result. For example:

```
MD=: ("2) (D.1)
```

```

det MD m
34 28 _33
_2 _2 _2
_20 _16 19

```

This result can be checked by examining the evaluation of the determinant as the alternating sum of the elements of any one column, each weighted by the determinant of its respective *complementary minor*, the matrix occupying the remaining rows and columns; the derivative with respect to any given element is its weighting factor. For example, the complementary minor of the leading element of **m** is the matrix **m00=: 7 8 ,:1 6**, whose determinant is **34**, agreeing with the leading element of the derivative.

Corresponding results can be obtained for the *permanent*, defined by the function **+/. \***. For example:

```

(per=: +/. *) m
350
per MD m
50 52 37
10 38 8
36 32 23

```

- F1 Read the following sentences and try to state the meanings of the functions defined and the exact results they produce. Then enter the expressions (and any related expressions that you might find helpful) and again try to state their meanings and results.

```

alph=: 4 4$ 'abcdefghijklmnop'
m=: i. 4 4
box=: <"2
minors=: 1&(|:\.)"2 ^:2
box minors m
box minors alph
box^:2 minors^:2 alph

```

[The function **minors** produces the *complementary minors* of its argument; the complementary minor of any element of a matrix is the matrix obtained by deleting the row and column in which the element lies.]

- F2 Enter and then comment upon the following sentences:

```

sqm=: *:m
det minors sqm
det D.1 sqm
(det D.1 sqm) % (det minors sqm)
((+/. *D.1)%+/. *@minors)sqm

```

## G. Matrix Inverse

The matrix inverse is a rank 2 function that produces a rank 2 result; its derivative is therefore a rank 2 function that produces a rank 4 result. For example:

```

m=: >3 1 4;2 7 8;5 1 6
MD=: ("2) (D. 1)

<"2 (7.1) ": (miv=: %.) MD m
+-----+
| 289.0 238.0 280.5| 17.0 17.0 17.0| 170.0 136.0 161.5|
| 17.0 14.0 16.5| 1.0 1.0 1.0| 10.0 8.0 9.5|
| 170.0 140.0 165.0| 10.0 10.0 10.0| 100.0 80.0 95.0|
+-----+
| 238.0 196.0 231.0| 14.0 14.0 14.0| 140.0 112.0 133.0|
| 17.0 14.0 16.5| 1.0 1.0 1.0| 10.0 8.0 9.5|
| 136.0 112.0 132.0| 8.0 8.0 8.0| 80.0 64.0 76.0|
+-----+
| 280.5 231.0 272.2| 16.5 16.5 16.5| 165.0 132.0 156.7|
| 17.0 14.0 16.5| 1.0 1.0 1.0| 10.0 8.0 9.5|
| 161.5 133.0 156.8| 9.5 9.5 9.5| 95.0 76.0 90.3|
+-----+

```

## H. Linear Functions and Operators

As discussed in Section 1K, a linear function distributes over addition, and any rank 1 linear function can be represented in the form  $\mathbf{mp} \&\mathbf{m}^1$ , where  $\mathbf{m}$  is a matrix, and  $\mathbf{mp}$  is the matrix product. For example:

```

r=: |."1
a=: 3 1 4 [ b=: 7 5 3
r a
4 1 3
r b
3 5 7
(r a)+(r b)
7 6 10
r (a+b)
7 6 10

mp=: +/ . *
]m=: i. 3 3
0 1 2
3 4 5
6 7 8
L=: m&mp
L a
9 33 57
L b
11 56 101
L a+b
20 89 158

VD=: ("1) (D. 1)
L VD a

```

Rank 1 reversal

Reversal is linear.

A linear function

The derivative of a linear

```

0 1 2
3 4 5
6 7 8
  =/~a
1 0 0
0 1 0
0 0 1

```

function yields the matrix  
that represents it.

An identity matrix

```

  L =/~a
0 1 2
3 4 5
6 7 8

```

A linear function applied to  
the identity matrix also yields  
the matrix that represents it.

```

  r VD a
0 0 1
0 1 0
1 0 0

```

The matrix that represents the  
linear function *reverse*

```

  r =/~a
0 0 1
0 1 0
1 0 0

```

```
perm=: 2&A.
```

A permutation is linear.

```
perm a
1 3 4
```

```
perm VD a
0 1 0
1 0 0
0 0 1
```

A function such as  $(\wedge 0 \ 1 \ 2) \ "0$  can be considered as a *family* of component functions.  
For example:

```

F=: (\wedge 0 1 2) "0
  F 3
1 3 9
  F y=: 3 4 5
1 3 9
1 4 16
1 5 25

```

The function  $L@F$  provides *weighted sums* or *linear combinations* of the members of the family  $F$ , and the adverb  $L@$  is called a *linear operator*. Thus:

```

L @ F y
21 60 99
36 99 162
55 148 241

```

The linear function  $F$  applied  
to the results of the family of  
functions  $F$

```
LO=: L@
F LO y
```

A linear operator

```
21 60 99
36 99 162
55 148 241
```

```
  C=: 2&o.@(*&0 1 2)"0      Family of cosines (harmonics)
  C y
1 _0.989992  0.96017
1 _0.653644  _0.1455
1  0.283662  _0.839072
```

```
  C LO y      A Fourier series
  0.930348    3.84088  6.75141
_0.944644    _0.342075 0.260494
_1.39448    _0.0607089 1.27306
```

H1 Enter and experiment with the expressions of this section.

## I. Linear Differential Equations

```
If f=: 2&o. and:
  F=: (f d.0)^(f d.1)^(f d.2)`:0"0
  L=: mp&c=: 1 0 1
```

then  $F$  is a family of derivatives of  $f$ . If the function  $L@F$  is identically zero, then the function  $f$  is said to be a solution of the *linear differential equation* defined by the linear function  $L$ . In the present example,  $f$  was chosen to be such a solution:

```
L@F y=: 0.1*i.4
0 0 0 0
```

The solution of such a differential equation is not necessarily unique; in the present instance the sine function is also a solution:

```
f=: 1&o.
L@F y=: 0.1*i.4
0 0 0 0
```

In general, the basic solutions of a linear differential equation defined by the linear function  $L=: mp&c$  are  $f=: ^@(*&sr)$ , where  $sr$  is any one root of the polynomial  $c&p$ . In the present instance:

```
f=: s=: ^@(*&0j1)

L@F y
0 0 0 0
c=: 1 0 1
  c K begin c      Roots of c&p. using Kerner's method
0j_1 0j1

f=: t=: ^@(*&0j_1)
L@F y
0 0 0 0
```

Moreover, any linear combination of the basic solutions  $\mathbf{s}$  and  $\mathbf{t}$  is also a solution. In particular, the following are solutions:

$\mathbf{u} = (\mathbf{s} + \mathbf{t})/2$  The cosine function  $\cos x$ .

$\mathbf{v} = (\mathbf{s} - \mathbf{t})/2j$  The sine function  $\sin x$ .

Since  $\mathbf{u}$  is equivalent to the cosine function, this agrees with the solution  $\mathbf{f}$  used at the outset.

11 Enter the expressions of this section, and experiment with similar differential equations.

## J. Differential Geometry

The differential geometry of curves and surfaces, as developed by Eisenhart in his book of that title [13], provides interpretations of the vector calculus that should prove understandable to anyone with an elementary knowledge of coordinate geometry. We will provide a glimpse of his development, beginning with a function which Eisenhart calls a *circular helix*.

The following defines a circular helix in terms of an argument in degrees, with a rise of 4 units per revolution:

```

CH = (1*cos(%&180p_1), 2*cos(%&180p_1), *&4r360)"0
CH 0 1 90 180 360
      0      1      0
0.0174524 0.999848 0.0111111
      1      0      1
      0      -1     2
_2.44921e_16      1      4

D = ("0) (D. 1)
x =: 0 1 2 3 4

CH D x
0.0174533      0 0.0111111
0.0174506 _0.000304602 0.0111111
0.0174427 _0.000609111 0.0111111
0.0174294 _0.000913435 0.0111111
0.0174108 _0.00121748 0.0111111
CH D D x
      0      0 0
_5.3163e_6 _0.000304571 0
_1.0631e_5 _0.000304432 0
_1.59424e_5 _0.0003042 0
_2.1249e_5 _0.000303875 0

```

The derivatives produced by  $\mathbf{CH} \ \mathbf{D}$  in the expression above are the directions of the tangents to the helix; *their* derivatives produced by  $\mathbf{CH} \ \mathbf{D} \ \mathbf{D}$  are the directions of the *binormals*. The binormal is perpendicular to the tangent, and indeed to the *osculating* (kissing) *plane* that touches the helix at the point given by  $\mathbf{CH}$ .

These matters may be made more concrete by drawing the helix on a mailing tube or other circular cylinder. An accurate rendering of a helix can be made by drawing a sloping straight line on a sheet of paper and rolling it on the tube. A drawing to scale can be made by marking the point of overlap on the paper, unrolling it, and drawing the straight line with a rise of 4 units and a run of the length of the circumference. Finally, the use of a sheet of transparent plastic will make visible successive laps of the helix. Then proceed as follows:

1. Use a nail or knitting needle to approximate the tangent at one of the points where its directions have been computed, and compare with the computed results.
2. Puncture the tube to hold the needle in the direction of the binormal, and again compare with the computed results.
3. Puncture a thin sheet of flat cardboard and hang it on the binormal needle to approximate the osculating plane.
4. Hold a third needle in the direction of the *principal* normal, which lies in the osculating plane perpendicular to the tangent.

To compute the directions of the principal normal we must determine a vector perpendicular to two other vectors. For this we can use the skew array used in Section 6I, or the following simpler *vector product* function:

```

vp=: (1&|.@[ * _1&|.@[) - (_1&|.@[ * 1&|.@[)
a=: 1 2 3 [ b=: 7 5 2
]q=: a vp b
_11 19 _9

a +/ . * q      b +/ . * q
0                0

```

Although we used degree arguments for the function **CH** we could have used radians, and it is clear that the choice of the argument to describe a curve is rather arbitrary. As Eisenhart points out, it is possible to choose an argument that is intrinsic to the curve, namely the length along its path. In the case of the helix defined by **CH**, it is easy to determine the relation between the path length and the degree argument. From the foregoing discussion of the paper tube model it is clear that the length of the helix corresponding to 360 degrees is the length of the hypotenuse of the triangle with sides 360 and 4. Consequently the definition of a function **df1** to give degrees from length is given by:

```
df1=: %&((%: +/ *: 4 360) % 360)
```

and the function **CH@df1** defines the helix in terms of its own length.

It is possible to modify the definition of the function **CH** to produce more complex curves, all of which can be modelled by a paper tube. For example:

1. Replace the constant multiple function for the last component by other functions, such as the square root, square, and exponential.
2. Multiply the functions for the first two elements by constants **a** and **b** respectively, to produce a helix on an elliptical cylinder. This can be



modelled by removing the cardboard core from the cylinder and flattening it somewhat to form an approximate ellipse.

## K. Approximate Integrals

Section M of Chapter 2 developed a method for obtaining the *integral* or *anti-derivative* of a polynomial, and Section N outlined a method for approximating the integral of any function by summing the function values over a grid of points to approximate the area under the graph of the function. Better approximations to the integral can be obtained by weighting the function values, leading to methods known by names such as *Simpson's Rule*.

We will here develop methods for producing these weights, and use them in the definition of an adverb (to be called  $\mathbf{I}$ ) such that  $\mathbf{f} \mathbf{I} \mathbf{x}$  yields the area under the graph of  $\mathbf{f}$  from  $0$  to  $\mathbf{x}$ .

The fact that the derivative of  $\mathbf{f} \mathbf{I}$  equals  $\mathbf{f}$  can be seen in Figure C1; since the difference  $(\mathbf{f} \mathbf{I} \mathbf{x+h}) - (\mathbf{f} \mathbf{I} \mathbf{x})$  is approximately the area of the rectangle with base  $\mathbf{h}$  and altitude  $\mathbf{f} \mathbf{x}$ , the secant slope of the function  $\mathbf{f} \mathbf{I}$  is approximately  $\mathbf{f}$ . Moreover, the approximation approaches equality for small  $\mathbf{h}$ .

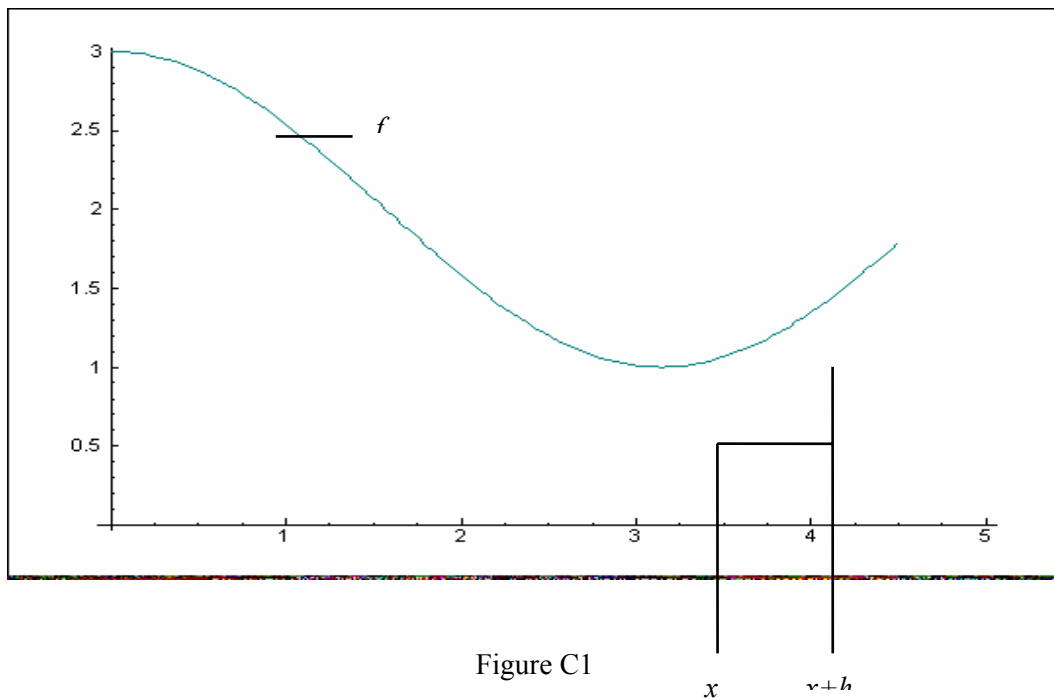


Figure C1

Figure C1 can also be used to suggest a way of approximating the function  $\mathbf{AREA} =: \mathbf{f} \mathbf{I}$ ; if the area under the curve is broken into  $\mathbf{n}$  rectangles each of width  $\mathbf{x} \div \mathbf{n}$ , then the area is approximately the sum of the areas of the rectangles with the common base  $\mathbf{h}$  and the altitudes  $\mathbf{f} \mathbf{h} \cdot \mathbf{i} \cdot \mathbf{n}$ . For example:

```

h=: y % n=: 10 [ y=: 2
cube=: ^&3
cube h*i.n
0 0.008 0.064 0.216 0.512 1 1.728 2.744 4.096 5.832

+/h*cube h*i.n          (4: %~ ^&4) y

```

3.24

4

The approximation can be improved by taking a larger number of points, but it can also be improved by using the areas of the *trapezoids* of altitudes  $f_{h^*k}$  and  $f_{h^*k+1}$  (and including the point  $h^*n$ ). Since the area of each trapezoid is its base times the average of its altitudes, and since each altitude other than the first and last enter into two trapezoids, this is equivalent to multiplying the altitudes by the *weights*  $w = (0.5, (1 \# \sim n-1), 0.5)$ . Thus:

```
]w=: 0.5, (1 #~ n-1), 0.5
0.5 1 1 1 1 1 1 1 1 0.5
```

```
+/h*w*cube h*i. n+1
4.04
```

The trapezoids provide, in effect, linear approximations to the function between grid points; much better approximations to the integral can be obtained by using groups of  $1+2*k$  points, each group being fitted by a polynomial of degree  $2*k$ . For example, the case  $k = 1$  provides fitting by a polynomial of degree 2 (a parabola) and a consequent weighting of  $3 \# \sim 1 \ 4 \ 1$  for the three points. If the function to be fitted is itself a polynomial of degree two or less, the integration produced is exact. For example:

```
w=: 3#~1 4 1
h=: (x=: 5)%(n=:2)
]grid=: h*i. n+1
0 2.5 5
f=: ^&2
w*f grid
0 8.33333 8.33333
```

```
+/h*w* f grid          Exact integral of ^&2
41.6667
```

```
+/h*w* ^&4 grid          Exact result is 625
651.042
```

Better approximations are given by several groups of three points, resulting in weights of the form  $3 \# \sim 1 \ 4 \ 2 \ 4 \ 2 \ 4 \ 2 \ 4 \ 1$ . For example, using  $g$  groups of  $1+2*k$  points each:

```
n=: (g=: 4) * 2 * (k=: 1)
]h=: n %~ x=: 5
0.625
```

```
]grid=: h*i. n+1
0 0.625 1.25 1.875 2.5 3.125 3.75 4.375 5
1, (4 2 $~ <: 2*g), 1
1 4 2 4 2 4 2 4 1
```

```
w=: 3#~ 1, (4 2 $~ <: 2*g), 1
+/h*w*^&2 grid          +/h*w*^&4 grid
41.6667                625.102
```

This case of fitting by parabolas ( $k=1$ ) is commonly used for approximate integration, and is called *Simpson's Rule*. The weights  $3/4, 1$  used in Simpson's rule will now be derived by a general method that applies equally for higher values of  $k$ , that is, for any odd number of points. Elementary algebra can be used to determine the coefficients  $c$  of a polynomial of degree 2 that passes through any three points on the graph of a function  $f$ . The integral of this polynomial (that is,  $(0, c/1, 2, 3) \cdot p$ ) can be used to determine the exact area under the parabola, and therefore the approximate area under the graph of  $f$ .

The appropriate weights are given by the function  $w$ , whose definition is presented below, after some examples of its use:

```

W 1
0.333333 1.333333 0.333333

W 2
0.311111 1.42222 0.533333 1.42222 0.311111

3*W 1          45*W 2
1 4 1          14 64 24 64 14

```

The derivation of the definition of  $w$  is sketched below:

```

vm=: ^~/~@i=: i.@>:@+      (Transposed)
vm 2                        Vandermonde of i. k
                             (for k=: 1+2* n)
1 1 1 1 1
0 1 2 3 4
0 1 4 9 16
0 1 8 27 64
0 1 16 81 256

%. vm 2                      Inverse of Vandermonde
1 _2.08333 1.45833 _0.416667 0.0416667
0          4 _4.33333 1.5 _0.166667
0          3 4.75 2 0.25
0 1.33333 2.33333 1.16667 _0.166667
0 _0.25 0.458333 _0.25 0.0416667

integ=: (0: , .%.@(^~/~)%"1>: )@i      Rows are integrals
integ 2                                of rows of inverse Vm
0 1 _1.04167 0.486111 _0.104167 0.00833333
0 0          2 _1.44444 0.375 _0.0333333
0 0          1.5 _1.58333 0.5 0.05
0 0 0.666667 _0.777778 0.291667 _0.0333333
0 0 _0.125 0.152778 _0.0625 0.00833333

W=: integ p. +:                  Polynomial at double argument
3*W 1          45*W 2
1 4 1          14 64 24 64 14

```

The results produced by  $w$  may be compared with those derived in more conventional notation, as in Hildebrand [7], p 60 ff. Finally, we apply the adverb  $f.$  to fix the

definition of  $w$  (by replacing each function used in its definition by *its* definition in terms of primitives:

```

W f.
(0: ,. %.@(^~/~) %"1 >:.)@(i.@>:@+:) p. +:

W=: (0: ,. %.@(^~/~) %"1 >:.)@(i.@>:@+:) p. +:
W 1
0.333333 1.333333 0.333333

```

A result of the function  $x$ : is said to be in *extended precision*, because a function applied to its result will be computed in extended precision, giving its results as rationals (as in  $1r3$  for the result of  $1\%3$ ). Thus:

```

! x:20                               Factorial 20 to complete precision
2432902008176640000

```

```

1 2 3 4 5 6 % x:3
1r3 2r3 1 4r3 5r3 2

```

```

W x:1
1r3 4r3 1r3
3*W x:1
1 4 1

```

```

W x:3
41r140 54r35 27r140 68r35 27r140 54r35 41r140
140*W x:3
41 216 27 272 27 216 41

```

We now define a function  $EW$  for extended weights, such that  $g EW k$  yields the weights for  $g$  groups of fits for  $1+2*k$  points:

```

ew=: ;@(#<) +/;.1~ 0: ~: #@[ | 1: >. i.@(*#)
EW=: ew W
2 EW x:1
1r3 4r3 2r3 4r3 1r3

```

```

3*2 EW x:1
1 4 2 4 1
45*2 EW x:2
14 64 24 64 28 64 24 64 14

```

Finally, we define a conjunction  $ai$  such that  $w ai f x$  gives the approximate integral of the function  $f$  to the point  $x$ , using the weights  $w$ :

```

ai=: 2 : '+/@(x.&space * x.&[ * y.@(x.&grid))"0'
grid=: space * i.@#[[
space=: ] % <:@#[[
3*w=: 1 EW 1
1 4 1

```

```

w ai *: x=: 1 2 3 4                               Weights for Simpson's rule (gives
0.333333 2.66667 9 21.3333                          exact results for the square function)

```

```

(x^3)%3

```

```
0.333333 2.66667 9 21.3333
```

```
(1 EW 2) ai (^&4) x
0.2 6.4 48.6 204.8
```

Weights give exact results for  
integral of fourth power

```
(x^5)%5
0.2 6.4 48.6 204.8
```

```
(cir=:0&o.)0 0.5 1
1 0.866025 0
```

0&o. is %:@(1"0-\*:) and cir 0.866025  
is the altitude of a unit circle

```
(2 EW 2) ai cir 1
0.780924
```

Approximation to area under cir  
(area of quadrant)

```
4 * (2 EW 2) ai cir 1
3.1237
```

Approximation to pi

```
4*(20 EW 3) ai (0&o.) 1
3.14132
```

```
o.1
3.14159
```

For use in exercises and in the treatment of interpretations in Section L, we will define the adverb **I** in terms of the weights **4 EW 4**, that is, four groups of a polynomial approximation of order eight:

```
I=: (4 EW 4) ai
^&9 I x=: 1 2 3 4
0.0999966 102.397 5904.7 104854
(x^10) % 10
0.1 102.4 5904.9 104858
^&9 d._1 x
0.1 102.4 5904.9 104858
```

K1 Use the integral adverb **I** to determine the area under the square root function up to various points.

K2 Since the graphs of the square and the square root intersect at **0** and **1**, they enclose an area. Determine its size.

```
[ (%:I-*:I) 1 or (%:-*:)I 1 ]
```

K3 Experiment with the expression  $(f - f I D) x$  for various functions **f** and arguments **x**.

## L. Areas and Volumes

The integral of a function may be interpreted as the area under its graph. To approximate integrals, we will use the adverb **I** defined in the preceding section. For example:

```
(0&o.) I 1
```

Approximate area of quadrant of circle

0.784908

```
4 * (0&o.) I 1
3.13963
```

Approximation to pi

```
*: I x=: 1 2 3 4
0.333317 2.66654 8.99956 21.3323
(^&3 % 3"0) x
0.333333 2.66667 9 21.3333
```

The foregoing integral of the square function can be interpreted as the area under its graph. Alternatively, it can be interpreted as the *volume* of a three-dimensional solid as illustrated in Figure L1; that is, as the volume of a pyramid. In particular, the equivalent function  $x^3 \div 3$  is a well-known expression for the volume of a pyramid.

Similarly for a function that defines the area of a circle in terms of its radius:

```
ca=: o.@*:@] " 0
ca x
3.14159 12.5664 28.2743 50.2655
ca I x
1.04715 8.37717 28.273 67.0174
```

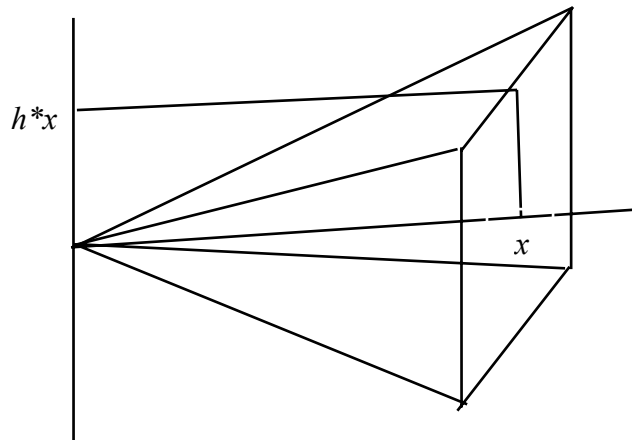


Figure L1

By drawing a figure analogous to Figure L1, it may be seen that the cone whose volume is determined by  $\pi x^2$  can be generated by revolving the 45-degree line through the origin about the axis. The volume is therefore called a *volume of revolution*.

Functions other than  $y = x$  (the 45-degree line) can be used to generate volumes of revolution. For example:

```
cade=: ca@^@-
cade x
0.425168 0.0575403 0.00778723 0.00105389
```

Area of circle whose radius is the decaying exponential

```
cade I x
1.3583 1.5423 1.56746 1.57123
```

Volume of revolution of the decaying exponential

Because the expression  $f \int y$  applies the function  $f$  to points ranging from 0 to  $y$ , the area approximated is the area over the same interval from 0 to  $y$ . The area under  $f$  from  $a$  to  $b$  can be determined as a simple difference. For example:

$$\begin{array}{ll} f =: x^3 & \\ f \int b =: 4 & f \int a =: 2 \\ 63.9965 & 3.99978 \\ \\ (f \int b) - (f \int a) & -/f \int b, a \\ 59.9967 & 59.9967 \end{array}$$

However, this approach will not work for a function such as  $\frac{1}{x}$ , whose value at 0 is infinite. In such a case we may use the related function  $\frac{1}{x+a}$ , whose value at 0 is  $\frac{1}{a}$ , and whose value at  $b-a$  is  $\frac{1}{b}$ . Thus:

$$\begin{array}{ll} g =: \frac{1}{x+a} & \\ g \int 0 & g \int b-a \\ 0.5 & 0.25 \end{array}$$

$$\begin{array}{ll} g \int b-a & \text{The integral of the reciprocal from 2 to 4} \\ 0.693163 & \end{array}$$

$$\begin{array}{ll} \ln 2 & \text{The natural log of 2} \\ 0.693147 & \end{array}$$

- L1 Use integration to determine the areas and volumes of various geometrical figures, including cones and other volumes of revolution.

## M. Physical Experiments

Simple experiments, or mere observation of everyday phenomena, can provide a host of problems for which simple application of the calculus provides solutions and significant insights. The reason is that phenomena are commonly governed by simple relations between the functions that describe them, and their rates of change (that is, derivatives).

For example, the position of a body as a function of time is related to its first derivative (velocity), its second derivative (acceleration), and its third derivative (jerk). More specifically, if  $p \int t$  gives the position at time  $t$  of a body suspended on a spring or rubber band, then the acceleration of the body ( $p \int d.2$ ) is proportional to the force exerted by the spring, which is itself a simple linear function of the position  $p$ .

If position is measured from the rest position (where the body rests after motion stops) this linear function is simply multiplication by a constant function  $c$  determined by the elasticity of the spring, and  $c \cdot p$  must be equal and opposite to  $m \cdot p \int d.2$ , where the constant function  $m$  is the mass of the body. In other words,  $(c \cdot p) - (m \cdot p \int d.2)$  must be zero.

This relation can be simplified to  $0 = p - c2 \cdot p \int d.2$ , where  $c2$  is the constant function defined by  $c2 =: m \cdot c$ . The function  $p$  is therefore (as seen in Section I) the sine function, or, more generally,  $p =: (a \cdot \sin) + (b \cdot \cos)$ , where  $a$  and  $b$  are constant functions.

This result is only an approximation, since a body oscillating in this manner will finally come to rest, unlike the sine and cosine functions which continue with undiminished amplitude. The difference is due to resistance (from friction with the air and internal friction in the rubber band) which is approximately proportional to the velocity. In other words, the differential equation:

$$0 = (d \cdot p) + (e \cdot p \cdot D. 1) + (f \cdot p \cdot D. 2)$$

provides a more accurate relation.

As seen in Section I, a solution of such a linear differential equation is given by  $e^{\alpha x}$ , where  $\alpha$  is a (usually complex) root of the polynomial  $(d, e, f) \cdot p$ . If  $\alpha = \alpha_1 + j \cdot \alpha_2$ , then  $e^{\alpha x}$  may also be written as  $(e^{\alpha_1 x}) \cdot (e^{j \cdot \alpha_2 x})$ , showing that the position function is a product of a decay function ( $e^{\alpha_1 x}$ ) and a periodic function ( $e^{j \cdot \alpha_2 x}$ ) like the solution to the simpler case in which the (resistance) constant  $e$  was zero.

Because oscillations similar to those described above are such a familiar sight, most of us could perform the corresponding "thought experiment" and so avoid the effort of an actual experiment. However, the performance of actual experiments is salutary, because it commonly leads to the consideration of interesting related problems.

For example, direct observations of the effect of greater damping can result from immersing the suspended body in a pail of water. The use of a heavier fluid would increase the damping, and raise the following question: Could the body be completely damped, coming to rest with no oscillation whatever?

The answer is that no value of the decay factor  $e^{\alpha_1 x}$  could completely mask the oscillatory factor  $e^{j \cdot \alpha_2 x}$ . However, a positive value of the factor  $f$  (the coefficient of  $p \cdot D. 2$ ) will provide real roots  $\alpha$ , resulting in non-oscillating solutions in terms of the hyperbolic functions  $\sinh$  and  $\cosh$ . Such a positive factor cannot, of course, be realized in the experiment described.

The performance of actual experiments might also lead one to watch for other phenomena governed by differential equations of the same form. For example, if the function  $q$  describes the quantity of electrical charge in a capacitor whose terminals are connected through a resistor and a coil, then  $q \cdot D. 1$  is the *current* (whose value determines the voltage drop across the resistor), and  $q \cdot D. 2$  is *its* rate of change (which determines the voltage drop across the coil). In other words, the charge  $q$  satisfies the same form of differential equation that describes mechanical vibrations, and enjoys the same form of *electrical* oscillation.

Other systems concerning motion suggest themselves for actual or thought experiments:

- \* The voltage generated by a coil rotating in a magnetic field.
- \* The amount of water remaining in a can at a time  $t$  following the puncture of its bottom by a nail.
- \* The amount of electrical charge remaining in a capacitor draining through a resistor (used in circuits for introducing a time delay).

Coordinate geometry also provides problems amenable to the calculus. For example,  $c = (1 \& o. , 2 \& o. ) \cdot 0$  is a rank 1 0 function that gives the coordinates of a circle, and the gradient  $c \cdot D. 1$  gives the slope of its tangent. Similarly,  $e = (a \cdot 1 \& o. ) , (b \cdot 2 \& o. )$  gives the coordinates of an ellipse.



If we are indeed surrounded by phenomena so clearly and simply described by the calculus, why is it that so many students forced into calculus fail to see any point to the study? This is an important question, for which we will now essay some answers:

1. Emphasis on rigorous analysis of limits in an introductory course tends to obscure the many interesting aspects of the calculus which can be enjoyed and applied without it.
2. On the other hand, a superficial treatment that does not lead the student far enough to actually produce significant new results is likely to leave her uninterested. Textbook pictures of suspension bridges with encouraging but unhelpful remarks that calculus can be used to analyze the form assumed by the cables, are more likely to discourage than stimulate a student.
3. The use of scalar notation makes it difficult to reach the interesting results of the vector calculus in an introductory course.
4. Although the brief treatments of mechanical and electrical vibrations given here may provide significant insights into their solutions, they would prove unsatisfactory in a text devoted to physics: they ignore the matter of relating the coefficients in the differential equations to the actual physical measurements (Does *mass* mean the same as *weight*? In what system of units are they expressed?); they ignore questions concerning the goodness of the approximation to the actual physical system; and they ignore the practicality of the computations required.

The treatment of such matters, although essential in a physics text, would make difficult its use by a student in some other discipline looking only for guidance in calculus.



# Chapter 8

## Analysis

### A. Introduction

To a math student conversant only with high-school algebra and trigonometry, the arguments used in Section 1E to determine the exact derivative of the cube (dividing the rise in the function value by the *run*  $\mathbf{x}$ , and then setting  $\mathbf{x}$  to zero in the resulting expression) might appear not only persuasive but conclusive. Moreover, the fact that the derivative so determined leads to consistent and powerful results would only tend to confirm a faith in the validity of the arguments.

On the other hand, a more mature student familiar with the use of rigorous axiomatic and deductive methods would, like Newton's colleagues at the time of his development of what came to be the calculus, have serious qualms about the validity of assuming a quantity  $\mathbf{x}$  to be non-zero and then, at a convenient point in the argument, asserting it to be zero.

Should a student interested primarily in the practical results of the calculus dismiss such qualms as pedantic “logic-chopping”, or are there important lessons to be learned from the centuries-long effort to put the calculus on a “firm” foundation? If so, what are they, and how may they be approached?

The important lesson is to appreciate the limitations of the methods employed, and to learn the techniques for assuring that they are being properly observed. As Morris Kline says in the preface to his *Mathematics: The Loss of Certainty* [14]:

But intellectually oriented people must be fully aware of the powers of the tools at their disposal. Recognition of the limitations, as well as the capabilities, of reason is far more beneficial than blind trust, which can lead to false ideologies and even to destruction.

Concerning “This history of the illogical development [of the calculus] ...”, Kline states (page 167):

But there is a deeper reason. A subtle change in the nature of mathematics had been unconsciously made by the masters. Up to about 1500, the concepts of mathematics were immediate realizations of or abstractions from experience. ... In other words, mathematicians were [now] contributing concepts rather than abstracting ideas from the real world.

Chapter VII of Kline provides a brief and readable overview of ingenious attempts to put the calculus on a firm basis, and equally ingenious refutations. Students are urged to read it in full, and perhaps to supplement it with Lakatos' equally readable account of the interplay between proof and refutation in mathematics. In particular, a student should be aware of the fact that weird and difficult functions sometimes brought into presentations of the calculus are included primarily because of their historical role as refutations. The words of Poincare (quoted by Kline on page 194) are worth remembering:

When earlier, new functions were introduced, the purpose was to apply them. Today, on the contrary, one constructs functions to contradict the conclusions of our predecessors and one will never be able to apply them for any other purpose.

The central concept required to analyze derivatives is the *limit*; it is introduced in Section B, and applied to *series* in Section D.

## B. Limits

The function  $h = (x - 9) / (x - 3)$  applied to the argument  $a = 3$  yields the meaningless result of zero divided by zero. On the other hand, a list of arguments that differ from  $a$  by successively smaller amounts appear to be approaching the limiting value  $g = 6$ . Thus:

```

g = 6
h = (x - 9) / (x - 3)
a = 3
h a
0

]i = , (+, -) (10^-i.5)
1 1 0.1 0.1 0.01 0.01 0.001 0.001 0.0001 0.0001

a+i
4 2 3.1 2.9 3.01 2.99 3.001 2.999 3.0001 2.9999

h a+i
7 5 6.1 5.9 6.01 5.99 6.001 5.999 6.0001 5.9999

|(g-h) a+i
1 1 0.1 0.1 0.01 0.01 0.001 0.001 0.0001 0.0001

```

We might therefore say that  $h(x)$  approaches a limiting value, or *limit*, as  $x$  approaches  $a$ , even though it differs from  $h(a)$ . In this case the limit is the constant function  $6$ .

We make a more precise definition of limit as follows: The function  $h$  has the limit  $g$  at  $a$  if there is a *frame* function  $f_r$  such that for any positive value of  $e$ , the expression  $e > |(g - h)(y)|$  is true for any  $y$  such that  $(|y - a|) < f_r(e)$ . In other words, for any positive value  $e$ , however small, there is a value  $d = f_r(e)$  such that  $h(y)$  differs from  $g$  by no more than  $e$ , provided that  $y$  differs from  $a$  by no more than  $d$ .

Figure B1 provides a graphic picture of the role of the frame function:  $d = f_r(e)$  specifies the half-width of a frame such that the horizontal boundary lines at  $e$  and  $-e$  are not crossed by the graph of  $g - h$  within the frame.

As illustrated at the beginning of this section, the function  $g = 6x^3$  is the apparent limit of the function  $h = (x-9)^3 - 3^3$  at the point  $a = 3$ . The simple frame function  $f(x) = |x-3|$  suffices, as illustrated (and later proved) below:

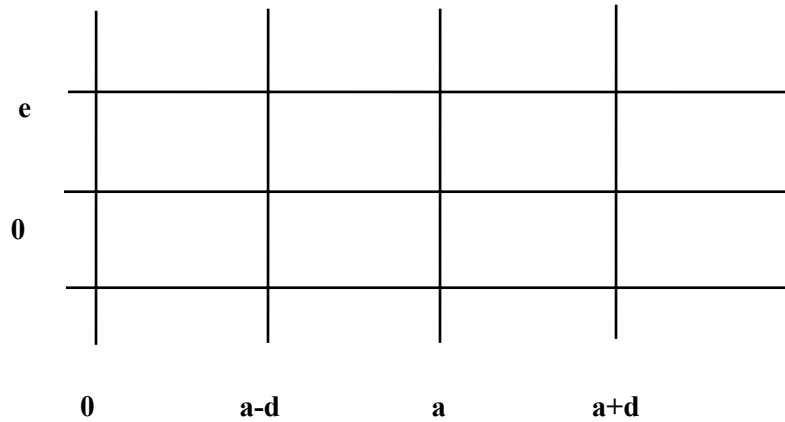


Figure B1

```

fr=: |
a=: 3
e=: 0.2
]d=: a fr e
0.2

]i=: , (+,-)"0,5%~>:i.5
0.2 _0.2 0.4 _0.4 0.6 _0.6 0.8 _0.8 1 _1

]j=: d*i
0.04 _0.04 0.08 _0.08 0.12 _0.12 0.16 _0.16 0.2 _0.2

|(g-h) a+j
0.04 0.04 0.08 0.08 0.12 0.12 0.16 0.16 0.2 0.2

e>:|(g-h) a+j
1 1 1 1 1 1 1 1 1 1

```

We now offer a proof that  $f(x) = |x-3|$  suffices, by examining the difference function  $g-h$  in a series of simple algebraic steps as follows:

```

g-h
6x^3 - (x-9)^3 - 3^3      Definitions of g and h
6x^3 + (x-9)^3 - 3^3
((6x^3-3^3)+(x-9)^3)-3^3
((18x^3-6x^3)+(x-9)^3)-3^3
((9x^3-6x^3)+x^3)-3^3
((3x^3-3^3)+x^3)-3^3
3x^3-3^3                  Cancel terms, but the domain now excludes 3

```

To recapitulate: for the limit point  $a = 3$  we require a frame function  $f(x)$  such that the magnitude of the difference  $(g-h)$  at the point  $a + a f(x) e$  shall not exceed  $e$ . We have just shown that the difference function  $(g-h)$  is equivalent to  $(3x^3-3^3)$ . Hence:

```
|(g-h) a + a fr e
```

```

| (3^0-1) 3+3 fr e           Definition of (g-h) and limit point
|3-(3+3 fr e)
|-3 fr e
|3 fr e

```

Consequently, the simple function  $f(x) = 3$  will suffice.

In the preceding example, the limiting function was a constant. We will now examine a more general case of the limit of the secant slope (that is, the derivative) of the fourth-power function. Thus:

```

f=: ^&4
h=: [ %~ ] -&f -~
x=: 0 1 2 3 4
]a=: 10^->:i. 6
0.1 0.01 0.001 0.0001 1e_5 1e_6

a h"0/ x
_0.001  3.439  29.679 102.719 246.559
_1e_6   3.9404 31.7608 107.461 255.042
_1e_9   3.994  31.976 107.946 255.904
_1e_12  3.9994 31.9976 107.995 255.99
_1e_15  3.99994 31.9998 107.999 255.999
_1e_18  3.99999      32      108      256

```

The last row of the foregoing result suggests the function  $4x^3$  as the limit. Thus:

```

g=: 4:*^&3
g x
0 4 32 108 256

a=: 1e_6
(g-a&h) x
1e_18 5.99986e_6 2.4003e_5 5.39897e_5 9.59728e_5

```

In simplifying the expression for the difference  $(g-a&h) x$  we will use functions for the polynomial and for weighted binomial coefficients as illustrated below:

```

w=: (]^i.@->:@[] * i.@>:@[ ! [
x=: 0 1 2 3 4 5
a=: 0.1
(x-a)^4
0.0001 0.6561 13.0321 70.7281 231.344 576.48

(4 w -a) p. x
0.0001 0.6561 13.0321 70.7281 231.344 576.48
4 w -a
0.0001 _0.004 0.06 _0.4 1

```

The following expressions for the difference can each be entered so that their results may be compared:

```

(g-a&h) x
(4*x^3)-a %~ (f x) - (f x-a)

```

```

(4*x^3)-a %~ (x^4) - (x-a)^4
(0 0 0 4 0 p. x)-a%~(0 0 0 0 1 p. x)-(4 w -a)p. x
a%~((a*0 0 0 4 0)p.x)-(0 0 0 0 1 p.x)-(4 w -a)p.x
a%~(1 _4 6 * a^ 4 3 2) p. x
(1 _4 6 * a^3 2 1) p. x

```

We will now obtain a simple upper bound for the magnitude of the difference (that is,  $|(g-a&h) \ x|$ ), beginning with the final expression above, and continuing with a sequence of expressions that are greater than or equal to it: (If the expressions are to be entered,  $x$  should be set to a scalar value, as in  $x=: 5$ , to avoid length problems)

```

x=:5
|(1 _4 6 * a^3 2 1) p. x           Magnitude of (g-a&h) x
| +/1 _4 6*(a^3 2 1)*x^i.3         Polynomial as sum of terms
+/(|1 _4 6)*(|a^3 2 1)*(|x^i.3)   Sum of mags>: mag of sum
+/1 4 6*(a^3 2 1)*|x^0 1 2       a is non-negative
+/6*(a^3 2 1)*|x^0 1 2
+/6*a*|x^0 1 2                    For a<1, the largest term is a^1
6*a+|x^0 1 2
a* (6+|x^0 1 2)

```

The final expression provides the basis for a frame function: if  $a=: e \% (6+|x^0 1 2)$ , then the magnitude of the difference  $|(g-a&h) \ x|$  will not exceed  $e$ . For example:

```

e=: 0.001
a=: e \% (6+|x^0 1 2)
|(g-a&h) x
0.000806451

```

## C. Continuity

Informally we say that a function  $f$  is continuous in an interval if its graph over the interval can be drawn without lifting the pen. Formally, we define a function  $f$  to be continuous in an interval if it possesses a limit at every point in the interval.

For example, the integer part function  $\lfloor \cdot \rfloor$  is continuous in the interval from  $0.1$  to  $0.9$ , but not in an interval that contains integers.

## D. Convergence of Series

The exponential coefficients function  $ec=: \%!$ , generates coefficients for a polynomial that approximates its own derivative, and the growth function (exponential) is defined as the limiting value for an infinite number of terms. Since the coefficients produced by  $ec$  decrease rapidly in magnitude (the 20th element is  $\%!19$ , approximately  $8e_{18}$ ), it seemed reasonable to assume that the polynomial  $(ec \ i. n) \&p.$  would converge to a

limit for large  $n$  even when applied to large arguments. We will now examine more carefully the conditions under which a sum of such a series approaches a limit.

It might seem that the sum of a series whose successive terms approach zero would necessarily approach a limiting value. However, the series  $\sum_{i=0}^n r^i$  provides a counter example, since (by considering sums over successive groups of  $2^i$  elements) it is easy to show that its sum can be made as large as desired.

If at a given term  $t$  in a series the remaining terms are decreasing in such a manner that the magnitudes of the ratios between each pair of successive terms are all less than some value  $r$  less than 1, then the magnitude of the sum of the terms after  $t$  is less than the magnitude of  $t(1-r)$ ; if this quantity can be shown to approach 0, the sum of the entire series therefore approaches a limit.

This can be illustrated by the series  $r^i$ , which has a fixed ratio  $r$ , and has a sum equal to  $(1-r^{n+1}) / (1-r)$ . For example:

```
S=: [ ^ i.@]
T=: (1"0-^)% (1"0-[]
r=: 3
n=: 10
r S n
1 3 9 27 81 243 729 2187 6561 19683

+ / r S n
29524
r T n
29524
```

A proof of the equivalence of  $T$  and the sum over  $S$  can be based on the patterns observed in the following:

```
(1,-r) */ r S n
1 3 9 27 81 243 729 2187 6561 19683
_3 _9 _27 _81 _243 _729 _2187 _6561 _19683 _59049

]dsums=:+//.(1,-r) */ r S n
1 0 0 0 0 0 0 0 0 0 _59049

-r^10          +/dsums          (1-r) * r T n
_59049         _59048           _59048
```

If the magnitude of  $r$  is less than 1, the value of  $r^n$  in the numerator of  $r T n$  approaches zero for large  $n$ , and the numerator itself therefore approaches 1; consequently, the result of  $r T n$  approaches  $(1-r)$  for large  $n$ .

The expression  $e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!}$  gives a pair of successive coefficients of the polynomial approximation to the exponential, and  $\frac{e^{x+1}}{e^x} = \sum_{j=0}^{\infty} \frac{x^j}{j!} = e^x$  gives their ratio. For example:

```
ec=: %@!
j=: 4
ec j-0 1
0.0416667 0.166667
```



```

% / e^x j - 0 1
0.25
% j
0.25

```

The ratio of the corresponding terms of the polynomial  $(e^x - 1) / x$  applied to  $x$  is  $x$  times this, namely,  $x^j$ . At some point this ratio becomes less than 1, and the series for the exponential therefore converges. Similar proofs of convergence can be made for the series for the circular and hyperbolic sines and cosines, after removing the alternate zero coefficients.

Another generally useful proof of convergence can be made for certain series by establishing upper and lower bounds for the series. This method applies if the elements alternate in sign and decrease in magnitude.

We will illustrate this by first developing a series approximation to the arctangent, that is, the inverse tangent  $\arctan x$ . The development proceeds in the following steps:

1. Derivative of the tangent
2. Derivative of the inverse tangent
3. Express the derivative as a polynomial in the tangent
4. Express the derivative as the limit of a polynomial
5. Integrate the polynomial
6. Apply the polynomial to the argument 1 to get a series whose sum approximates the arctangent of 1 (that is, one-quarter pi):

```

]x=: 1,1r6p1,1r4p1,1r3p1
1 0.523599 0.785398 1.0472

```

```

'`sin cos tan arctan'=: (1&o.)` (2&o.)` (3&o.)` (_3&o.)

```

```

sin x
0.841471 0.5 0.707107 0.866025

```

```

cos x
0.540302 0.866025 0.707107 0.5

```

```

tan x
1.55741 0.57735 1 1.73205
(sin % cos) x
1.55741 0.57735 1 1.73205

```

```

INV=: ^: _1
tan INV tan x
1 0.523599 0.785398 1.0472

```

```

D=: ("0) (D.1)
tan D

```

```

(sin % cos) D
(sin%cos)*(sin D%sin)-(cos D%cos)
tan*(cos%sin)-(-@sin%cos)
tan * %@tan +tan

```

Definition of tan  
 $\frac{d}{dx} \tan^{-1} x = \frac{1}{1+x^2}$   
 $\tan^{-1} x = \int \frac{1}{1+x^2} dx$

```

1"0 + tan * tan
1"0 + *:@tan
Derivative of tangent QED

tan INV D
1"0 % tan D @(tan INV)
1"0 % (1"0 + *:@tan) @ (tan INV)
1"0 % (1"0@(tan INV)) + *:@tan@(tan INV)
1"0 % 1"0 + *:@]
1"0 % 1"0 + *:
%@(1"0+*:)
Derivative of inverse tan QED

c=: 1 0 1
% c&p. x
0.5 0.784833 0.618486 0.476958
Derivative of inverse tangent as
reciprocal of a polynomial

b=: 1 0 _1 0 1 0 _1 0 1 0 _1
c */ b
1 0 _1 0 1 0 _1 0 1 0 _1
0 0 _0 0 0 0 _0 0 0 0 _0
1 0 _1 0 1 0 _1 0 1 0 _1

+//. c */ b
1 0 0 0 0 0 0 0 0 0 0 _1
Product polynomial shows that
b&p. is approx reciprocal of c&p.

%@(1:+*:) x
0.5 0.784833 0.618486 0.476958

b&p. x
0 0.7845 0.584414 _0.352555
Better approx needs more terms of b

int=: 0: , ] % 1: + i.&#
a=: int b
a&p. x
0.744012 0.482334 0.6636 0.736276
The fn a&p. is the integral of b&p.
Approximation to arctangent

tan INV x
0.785398 0.482348 0.665774 0.808449

7.3 ": 8{. a
0.000 1.000 0.000 _0.333 0.000 0.200 0.000 _0.143
1r4p1 , a p. 1
0.785398 0.744012
Coeffs for arctan are reciprocals of odds
Arctan 1 is one-quarter pi

+/a
0.744012
Polynomial on 1 is sum of coefficients

gaor=: _1&^@i. * 1: % 1: + 2: * i.
gaor 6
1 _0.333333 0.2 _0.142857 0.111111 _0.0909091
Generate alternating odd reciprocals

+/\gaor 6
1 0.666667 0.866667 0.72381 0.834921 0.744012

7 2 $ +/\ gaor 14
1 0.666667
First column (sums of odd number

```

0.866667 0.72381  
 0.834921 0.744012  
 0.820935 0.754268  
 0.813091 0.76046  
 0.808079 0.764601  
 0.804601 0.767564

of terms) are decreasing upper  
 bounds of limit. Second column  
 (sums of even number of terms)  
 are increasing lower bounds of limit.

`1r4p1 , +/gaor 1000`  
 0.785398 0.785148

D1 Test the derivations in this section by enclosing a sentence in parens and applying it to an argument, as in `(1: + *:tan) x`

D2 Prove that a decreasing alternating series can be bounded as illustrated.

[Group pairs of successive elements to form a sum of positive or negative terms]



## Appendix

### Topics in Elementary Math

#### A. Polynomials

An atomic constant multiplied by an integer power (as in  $a \cdot 0 \cdot x^n$ ) is called a *monomial*, and a sum of monomials is called a *polynomial*. We now define a polynomial function, the items of its list left argument being called the *coefficients* of the polynomial:

```
pol=: +/@([ * ] ^ i.@#@[) " 1 0
```

For example:

```

c=: 1 2 3 [ x=: 0 1 2 3 4
c pol x
1 6 17 34 57
          1 3 3 1 pol x
1 8 27 64 125

```

The polynomial may therefore be viewed as a weighted sum of powers, the weights being specified by the coefficients. It is important enough to be treated as a *primitive*, denoted by  $p$ .

It is important for many reasons. In particular, it is easily expressed in terms of sums, products, and integral powers; it can be used to approximate almost any function of practical interest; and it is *closed* under a number of operations; that is, the sums, products, derivatives, and integrals of polynomials are themselves polynomials. For example:

```

x=: 0 1 2 3 4 [ b=: 1 2 1 [ c=: 1 3 3 1
(b p. x) + (c p. x)          Sum of polynomials
2 12 36 80 150

b +/@,: c                    "Sum" of coefficients
2 5 4 1

(b +/@,: c) p. x            Sum polynomial
2 12 36 80 150

(b p. x) * (c p. x)         Product of polynomials
1 32 243 1024 3125

b +//.@(*/) c              "Product" of coefficients
1 5 10 10 5 1

(b +//.@(*/) c) p. x       Product polynomial

```

```
1 32 243 1024 3125
```

```
  c&p. d.1 x          Derivative of polynomial
3 12 27 48 75
```

```
  c&p. d._1 x         Integral of polynomial
0 3.75 20 63.75 156
```

```
  derc=: }.@[ ] * i.@#
  derc c              "Derivative" coefficients
3 6 3
```

```
  (derc c) p. x      Derivative polynomial
3 12 27 48 75
```

```
  intc=: 0: , ] % >:@i.@#
  intc c              "Integral" coefficient
```

```
0 1 1.5 1 0.25
  (intc c) &p. x
0 3.75 20 63.75 156
```

A polynomial is "linear in its coefficients" in that  $(c+d) p. x$  is  $(c p. x) + (d p. x)$ . This linearity can be made clear by expressing  $c p. x$  as  $m \mp c$ , where  $m$  is the *Vandermonde* matrix obtained as a function of  $x$  and  $c$ . Thus:

```
  vm=: [ ^/ i.@#@[
  x=: 0 1 2 3 4 5
  c=: 1 3 3 1
  x vm c              (x vm c) mp c              c p. x
1 0 0 0              1 8 27 64 125 216          1 8 27 64 125
216
1 1 1 1
1 2 4 8
1 3 9 27
1 4 16 64
1 5 25 125
```

The expression  $c=: (f x) \% . x^/i.n$  yields an  $n$ -element list of coefficients such that  $c p. x$  is the best least-squares approximation to the values of the function  $f$  applied to the list  $x$ . In other words, the value of  $+/\text{sqr} (f x) - c p. x$  is the least achievable for an  $n$ -element list of coefficients  $c$ .

We now define a conjunction **FIT** such that  $a \text{ FIT } f x$  produces the coefficients for the best polynomial fit of  $a$  elements:

```
  FIT=: 2 : 'y. \% . ^/&(i. x.)'
  ]c=: 5 FIT ! x=: 0 1 2 3 4
1 _2.08333 3.625 _1.91667 0.375

  c p. x              !x
1 1 2 6 24           1 1 2 6 24
  ]c=: 4 FIT ! x
0.871429 3.27381 _3.71429 1.08333
```

```

c p. x
0.871429 1.51429 1.22857 6.51429 23.8714

```

## B. Binomial Coefficients

$m!n$  is the number of ways that  $m$  things can be chosen out of  $n$ ; for example  $2!3$  is 3, and  $3!5$  is 10. The expression  $c=: (i. n+1)!n$  yields the *binomial coefficients of order n*, and  $c p. x$  is equivalent to  $(x+1)^n$ . For example:

```

]c=: (i. n+1)!n=: 3
1 3 3 1

```

```

c p. x=: 0 1 2 3 4 5
1 8 27 64 125 216

```

```

(x+1) ^ n
1 8 27 64 125 216

```

```

<@(i.@>: ! ])"0 i. 6

```

1	1 1	1 2 1	1 3 3 1	1 4 6 4 1	1 5 10 10 5 1
---	-----	-------	---------	-----------	---------------

## C. Complex Numbers

Just as subtraction and division applied to the *counting* numbers (positive integers) introduce new classes of numbers (called *negative* numbers and *rational* numbers), so does the square root applied to negative numbers introduce a new class called *imaginary* numbers. For example:

```

a=: 1 2 3 4 5 6
]b=: -a
_1 _2 _3 _4 _5 _6

```

Negative numbers

```

% a
1 0.5 0.333333 0.25 0.2 0.166667

```

Rational numbers

```

%: b
0j1 0j1.41421 0j1.73205 0j2 0j2.23607 0j2.44949

```

Imaginary numbers

Arithmetic functions are extended systematically to this new class of numbers to produce *complex* numbers, which are represented by two real numbers, a *real* part and an *imaginary* part, separated by the letter  $j$ . Thus:

```

a+%:b
1j1 2j1.41421 3j1.73205 4j2 5j2.23607 6j2.44949

```

Complex numbers

```

j. a
j1 0j2 0j3 0j4 0j5 0j6
]d=: a+j. 5 4 3 2 1 0
1j5 2j4 3j3 4j2 5j1 6

```

The function  $j.$  multiplies its argument by  $0j1$   
The monad  $+$  is the *conjugate* function; it reverses the

`+d` sign of the imaginary part  
`1j_5 2j_4 3j_3 4j_2 5j_1 6`

`d*d` Product with the conjugate  
`26 20 18 20 26 36` produces a real number

`abs(d)` Magnitude of a  
`5.09902 4.47214 4.24264 4.47214 5.09902 6` complex number

`abs(d)`  
`5.09902 4.47214 4.24264 4.47214 5.09902 6`

## D. Circular and Hyperbolic Functions.

<code>sin=: 1&amp;o.</code>	<code>sinh=: 5&amp;o.</code>	
<code>cos=: 2&amp;o.</code>	<code>cosh=: 6&amp;o.</code>	
<code>tan=: 3&amp;o.</code>	<code>tanh=: 7&amp;o.</code>	
<code>SIN=: sin@rfd</code>		Sine in degrees
<code>COS=: cos@rfd</code>		
<code>TAN=: tan@rfd</code>		
<code>rfd=: o.@(%&amp;180)</code>		Radians from degrees

## E. Matrix Product and Linear Functions

The dot conjunction applied to the sum and product functions yields a function commonly referred to as the *dot* or *matrix* product. Thus:

```

mp=: +/ . *
]m=: i. 3 3           ]n=: i. 4 3
0 1 2                 0 1 2
3 4 5                 3 4 5
6 7 8                 6 7 8
                      9 10 11

n mp m               3 2 1 mp m
15 18 21             12 18 24
42 54 66             1 4 6 mp m
69 90 111            48 59 70
96 126 156

```

Left and right bonds of the matrix product distribute over addition; that is, `a&mp c+d` is `(a&mp c)+(a&mp d)`, and `mp&b c+d` is `(mp&b c)+(mp&b d)`. For example:

```

mp&m 3 2 1 + 1 4 6
60 77 94
(mp&m 3 2 1) + (mp&m 1 4 6)
60 77 94

```

A function that distributes over addition is said to be *linear*; the name reflects the fact that a linear function applied to the coordinates of collinear points produces collinear points. For example:



```

]line=: 3 _7 1,:2 2 4      ]a=: 3 1,:_4 2
3 _7 1                    3 1
2 2 4                     _4 2

a&mp line
11 _19 7
_8 32 4

mp& 3 1 _2 line          mp&3 1 _2 a &mp line
0 0                      0 0

```

## F. Inverse, Reciprocal, And Parity

We will now define and illustrate the use of four further adverbs:

<code>I=: ^: _1</code>	Inverse adverb
<code>R=: %@</code>	Reciprocal adverb
<code>ODD=: .: -</code>	Odd adverb
<code>EVEN=: .. -</code>	Even adverb
<code>*: I x=: 0 1 2 3 4 5</code>	Inverse of the square,
<code>0 1 1.41421 1.73205 2 2.23607</code>	that is, the square root
<code>*: R x</code>	Reciprocal of the square,
<code>_ 1 0.25 0.111111 0.0625 0.04</code>	that is, <code>%*:</code> , or <code>^&amp;_2</code>
<code>c=: 4 3 2 1</code>	
<code>even=: c&amp;p. EVEN</code>	Even part of polynomial <code>c&amp;p.</code>
<code>odd=: c&amp;p. ODD</code>	Odd part of polynomial
<code>even x</code>	Even function applied to <code>x</code>
<code>4 6 12 22 36 54</code>	
<code>odd x</code>	Odd function applied to <code>x</code>
<code>0 4 14 36 76 140</code>	
<code>(even + odd) x</code>	Sum of even and odd parts
<code>4 10 26 58 112 194</code>	is equal to the original
<code>c&amp;p. x</code>	function <code>c&amp;p.</code>
<code>4 10 26 58 112 194</code>	
<code>4 0 2 0 p. x</code>	Even part is a polynomial with non-
<code>4 6 12 22 36 54</code>	zero coefficients for even powers
<code>0 3 0 1 p. x</code>	Odd part is a polynomial with non-
<code>0 4 14 36 76 140</code>	zero coefficients for odd powers

For an even function,  $f(-y)$  equals  $f(y)$ ; for an odd function,  $f(-y)$  equals  $-f(y)$ . Plots of even and odd functions show their graphic properties: the graph of an even function is "reflected" in the vertical axis, and the odd part in the origin.

### Exercises

AP1 Enter the expressions of this section, and verify that the results agree with those given in the text.

AP2 Predict the results of each of the following sentences, and then enter them to validate your predictions:

`D=: ("1) (D.1)`

`x=: 1 2 3 4 5`

`|. D x`

`2&|. D x`

`3 1 0 2 &{ D x`

`+/\ D x`

`+/\. D x`

AP3 Define `show=: {&'.*'` and use it to display the results of Exercises G2, as in `show |. D x`.

AP4 Define a function `rFd` to produce radians from degrees, and compare `rFd 90 180` with

`o. 0.5 1.`

`[ rFd=: %&180@o. ]`

AP5 Define a function `AREA` such that `AREA v` yields the area of a triangle with two sides of lengths `0{v` and `1{v` and with an angle of `2{v` degrees between them. Test it on triangles such as `2 3 90` and `2 3 30`, whose areas are easily computed.

`[AREA=: -:@(0&{ * 1&{ * 1&o.@rFd@{:})"1]`

AP6 Experiment with the vector derivative of the triangle area function of Exercise G5, using `VD=: ("1) (D.1)`.

`[AREA VD 2 3 90]`

AP7 Heron's formula for the area of a triangle is the square root of the product of the semiperimeter with itself less *zero* and less each of the three sides. Define a function `hat` to give Heron's area of a triangle, and experiment with its vector derivative `hat VD`. In particular, try the case `hat VD 3 4 5`, and explain the (near) zero result in the final element.

`[ hat=: %:@(*)@(-:@(+/-) - 0: , )"1 ]`

AP8 Define a function `bc` such that `bc n` yields the binomial coefficients of order `n`, a function `tbc` such that `tbc n` yields a table of all binomial coefficients up to

order  $n$ , and a function `tabc` for the corresponding alternating binomial coefficients.

```
[ bc=: i.@>: ! ]
tabc=: !/~ @ (i.@>:)
tabc=: %.@tabc ]
```

AP9 Test the assertion that  $(bc\ n)\ p.\ x=: i.\ 4$  is equivalent to  $x^{n+1}$  for various values of  $n$ .

AP10 Write an expression to yield the matrix  $m$  such that  $mp\&m$  is equivalent to a given linear function  $L$ . Test it on the linear functions  $L=: |.\ "1$  and  $L=: 3\&A.\ "1$ , using the argument  $x=: 3\ 1\ 4\ 1\ 6$

```
[ L = i. # x ]
```

AP11 Experiment with the use of various functions on imaginary and complex numbers, including the exponential, the sine, cosine, hyperbolic sine and hyperbolic cosine. Also experiment with matrices of complex numbers and with the use of the matrix inverse and matrix product functions upon them.



---

## References

1. Iverson, Kenneth E., *Arithmetic*, ISI 1991
2. Lakatos, Imre, *Proofs and Refutations: the logic of mathematical discovery*, Cambridge University Press.
3. Lanczos, Cornelius, *Applied Analysis*, Prentice Hall, 1956.
4. McConnell, A.J., *Applications of the Absolute Differential Calculus*, Blackie and Son, Limited, London and Glasgow, 1931.
5. Oldham, Keith B., and Jerome Spanier, *The Fractional Calculus*, Academic Press, 1974.
6. Johnson, Richard E., and Fred L. Kiokemeister, *Calculus with analytic geometry*, Allyn and Bacon, 1957.
7. Hildebrand, F.B., *Introduction to Numerical Analysis*, McGraw-Hill, 1956.
8. Woods, Frederick S., *Advanced Calculus*, Ginn and Company, 1926.
9. Schey, H.M., *Div, Grad, Curl, and All That*, W.W. Norton, 1973.
10. Jordan, Charles, *Calculus of Finite Differences*, Chelsea, 1947.
11. Coleman, A.J. et al, *Algebra*, Gage, 1979.
12. Churchill, Ruel V., *Modern Operational Mathematics in Engineering*, McGraw-Hill, 1944.
13. Eisenhart, Luther Pfahler, *A Treatise on the Differential Geometry of Curves and Surfaces*, Ginn, 1909.
14. Kline, Morris, *Mathematics: The loss of certainty*, Oxford, 1980

# Index

- acceleration*, 10, 28, 29, 105
- adverb, 11, 12, 15, 25, 32, 49, 67, 73, 74, 82, 87, 90, 91, 92, 96, 99, 103, 104, 123
- adverbs*, 11, 12, 63, 123
- aggregation*, 57
- alternating binomial coefficients, 59, 125
- alternating sum, 14, 93
- ambivalent, 12
- Analysis, 109
- angle, 41, 42, 75, 77, 78, 79, 80, 81, 124
- anti-derivative*, 15
- Applications, 86
- AREA**, 100, 124
- AREAS, 104
- Argument Transformations, 31
- atop*, 30, 31
- Atop, 30
- axes, 42
- beta* function, 63
- binomial coefficients, 59, 60, 61, 63, 112, 121, 125
- Binomial Coefficients, 121
- binormals*, 98
- bisection* method, 87, 91
- bold brackets, 13
- Calculus, 7
- Calculus of Differences, 21
- Calculus of Finite Differences*, 49
- Celsius, 31
- chain* rule, 15
- circle, 70
- circular*, 26, 30, 31, 69, 70, 72, 73, 75, 86, 97, 98, 115
- Circular, 122
- Circulars, 29
- closed*, 119
- Coefficient Transformations, 55
- Coefficients, 28, 91
- comments, 13
- complementary minor*, 93
- complex numbers, 86, 121, 125
- Complex Numbers, 121
- complex roots, 92
- computer, 10, 11, 13, 14, 15, 22, 63
- COMPUTER, 15
- conjugate*, 86, 122
- conjunction, 12, 13, 15, 30, 31, 62, 63, 65, 69, 103, 120, 122
- conjunctions*, 11, 12, 31, 63
- constant function, 33, 106, 110
- Continuity, 113
- continuous, 26, 49, 86, 113

## 2 Calculus

---

- contour* integral, 86
- conventional notation, 10
- CONVERGENCE OF SERIES, 114
- copula*, 11, 12
- cos**, 30, 31, 70, 72, 73, 75, 79, 80, 81, 86, 106, 116, 122
- cosh*, 29, 69, 70, 72, 73, 106, 122
- cosine, 29, 30, 31, 73, 75, 76, 77, 80, 97, 106, 125
- cross*, 79
- cross product, 47, 80, 81, 82, 83
- Cross Products, 79
- curl*, 46, 79
- curves, 97
- cylinder, 98
- de Morgan, 15
- decay, 28, 67, 68, 69, 73, 106
- Decay, 27
- degrees*, 30, 79, 80, 97, 98, 122, 124
- derivative, 9, 10, 15, 16, 17, 18, 22, 25, 26, 27, 28, 29, 32, 33, 37, 40, 49, 52, 61, 62, 63, 65, 68, 70, 78, 79, 86, 87, 88, 89, 90, 92, 93, 94, 95, 99, 105, 109, 112, 114, 115, 124
- Derivative, 15, 16
- Derivative of polynomial, 120
- derivative operator*, 10
- derivatives, 15, 16, 21, 26, 30, 31, 32, 39, 51, 52, 60, 61, 63, 68, 70, 72, 79, 96, 98, 105, 110, 119
- derived* function, 15
- determinant, 38, 41, 83, 92, 93
- Determinant, 92
- diagonal sums, 69
- difference*, 49
- difference* calculus, 16, 61
- Difference Calculus, 49
- Differential Calculus, 23
- differential equation, 96
- Differential Equations, 25
- Differential Geometry, 97
- differintegral*, 61, 62
- differintegrals*, 60
- direction, 41
- discontinuous, 85
- displayed, 15
- divergence*, 42, 46
- Divergence, 42
- division, 90, 92, 121
- dot, 122
- DOT, 79
- electrical system, 29
- Elementary Math**, 119
- ellipse, 99, 107
- Even part, 123
- executable, 10, 11, 22
- executed*, 14
- EXERCISES**, 13, 124
- experimentation, 10, 15, 22, 67
- Experimentation, 69
- experiments, 11, 63, 67, 69, 105, 106, 107
- explore, 11
- exponential, 12, 16, 26, 27, 28, 29, 69, 73, 85, 99, 105, 114, 115, 125
- Exponential Family, 73
- exponentially*, 26

- 
- extrema, 89
- Extrema, 87
- $\epsilon$  ., 16, 28, 47, 86, 96, 99, 101, 102
- factorial function, 15, 62
- Fahrenheit, 31
- Family of cosines, 96
- first derivative, 29, 32, 61, 105
- foreign* conjunction, 15
- fork*, 33, 52, 63
- Fourier series, 96
- Fractional Calculus, 61
- Fractional* derivatives, 21
- function, 12
- functions*, 7
- Functions, 11, 32, 105
- gamma function, 62
- gamma function and imaginary numbers., 15
- Gradient, 38
- growth, 7, 16, 26, 27, 28, 67, 68, 69, 73, 114
- Growth, 26
- harmonics, 96
- heaviside, 10
- Heaviside's, 46
- helix*, 97
- Heron's area, 124
- hierarchy, 12
- high-school algebra, 12, 109
- hyperbola, 29, 70
- hyperbolic*, 26, 29, 69, 72, 73, 106, 115, 125
- Hyperbolic Functions, 122
- Hyperbolics, 28
- identity, 33, 52, 58, 65, 77, 79, 95
- imaginary numbers, 15, 121
- imaginary* part, 121
- induction*, 33
- infinitesimal*, 49
- Infinitesimal Calculus*, 49
- Inflection Points, 87
- informal proofs, 13, 22, 52
- initial guess, 89
- insert, 11
- integer* part, 85
- integral*, 15, 21, 22, 55, 61, 62, 63, 86, 99, 100, 101, 103, 104, 105, 116, 119
- Integral, 15, 16
- integration, 61, 86, 100, 101, 105
- Interpretations, 18, 85
- Inverse, 123
- inverse matrix, 57
- irrotational, 46
- items, 12, 82, 119
- Jacobian, 40, 41, 42
- jerk, 105
- Jordan, 49
- Kerner's method, 97
- KERNER'S METHOD, 91
- Kline, 109
- Lakatos, 21
- Laplacian, 42
- leibniz, 10
- Less than, 11



- Lesser of, 11
- limit*, 10, 22, 79
- Limits, 110
- line*, 86
- linear*, 123
- linear combinations*, 96
- Linear Differential Equations, 96
- linear form*, 15
- linear function, 40, 41, 57, 58, 94, 95, 96, 106, 123, 125
- linear functions, 41
- Linear Functions, 94
- LINEAR FUNCTIONS, 122
- linear operator*, 96
- lists*, 11, 14, 32
- local, 87
- local* behaviour, 16
- local minimum, 89
- logarithm*, 73, 74
- Logarithm, 73
- Loss of Certainty*, 109
- lower bounds, 115
- magnitude, 41
- Magnitude, 122
- matrices, 11, 56, 125
- MATRIX INVERSE, 94
- MATRIX PRODUCT, 122
- maximum, 87, 88, 89, 90
- Maxwell's, 46
- mechanical system, 29
- minimum*, 11, 15, 87, 88, 89
- minors**, 93
- modern, 10
- multiplication table, 12
- natural logarithm*, 73
- negation, 12, 27
- negative* numbers, 121
- newton, 10
- Newton's Method, 89
- normal*, 79
- normalized coefficient*, 91
- Normals, 82
- notation, 10, 11, 15, 22, 64, 102, 107
- Notation, 11
- NOTATION, 15
- nouns*, 11, 12
- number of items*, 82
- numerator, 115
- Odd part, 123
- operators, 10, 11
- Operators, 94
- oscillations, 29, 106
- osculating*, 98
- outof*, 61
- Parentheses, 12
- Parity**, 71, 123
- Partial* derivatives, 21
- periodic* functions, 29
- Permanent, 92
- permutation, 95
- permutations, 42

- 
- perpendicular*, 79
- Physical Experiments, 105
- pi, 103, 104
- plane, 80
- point of inflection*, 88
- polynomial, 26, 27, 30, 31, 49, 50, 55, 67, 69, 87, 89, 90, 91, 92, 96, 100, 101, 103, 106, 112, 114, 115, 116, 119, 120, 123, 124
- polynomials, 26, 28, 68, 90, 92, 119
- Polynomials, 119
- positive integers, 121
- power*, 10, 15, 17, 21, 49, 51, 52, 74, 103, 112, 119
- Power, 73
- precedence, 12
- primes, 14
- principal normal*, 98
- Product of polynomials, 119
- pronouns*, 11
- proof*, 14, 17, 67, 68, 69, 72, 110, 111, 114, 115
- proofs, 13, 21, 22, 52, 67, 115
- Proofs, 72, 80
- Proofs and Refutations*, 21
- proverbs*, 11
- punctuation, 12
- pyramid, 104
- Pythagoras, 76
- quotes, 14, 86
- radians*, 30, 41, 78, 79, 86, 124
- Random starting value, 91
- rank, 12, 21, 37, 93, 94, 107
- rank conjunction, 12
- rank-0, 37
- rate of change*, 7
- rational constant, 63
- rational numbers*, 121
- real part*, 121
- Reciprocal, 123
- residuals*, 92
- rise*, 89
- roots*, 87
- rotation*, 41, 42, 46
- Rotation, 80
- run*, 89
- scalar product*, 79
- scalars*, 21, 80
- Scaling, 30
- Secant Slope, 15, 16
- secant slopes*, 16
- second derivative*, 10
- Semi-Differintegrals, 63
- series, 115
- Simpson's Rule*, 101
- sin**, 13, 30, 31, 70, 72, 73, 75, 79, 81, 86, 106, 116, 122
- sine, 29, 30, 73, 75, 76, 77, 78, 79, 80, 81, 96, 97, 106, 125
- Sine, 13, 73, 122
- sinh*, 29, 69, 70, 72, 73, 106, 122
- Skew part, 82
- slope*, 89
- Slopes As Linear Functions, 57
- Stirling numbers*, 56
- stope polynomial, 55

## 6 Calculus

---

subtraction, 12, 121

**Sum Formulas**, 76

Sum of polynomials, 119

*summation*, 57

surfaces, 97

*tables*, 11, 89

tangent, 15, 22, 78, 85, 86, 88, 89, 98, 107, 115, 116

*tautologies*, 72, 73, 76, 77, 79

Tautologies, 78

tautology, 33, 79

*tensor analysis*, 38

Terminology, 11

third derivative, 105

*trapezoids*, 100

*trigonometric*, 26, 75

Trigonometric Functions, 75

trigonometry, 30, 75, 77, 79, 109

*under*, 15, 49, 51, 57, 65, 86, 99, 100, 101, 103, 104, 105, 114, 119

upper, 115

Vandermonde, 101, 120

*vector calculus*, 38, 46, 97, 107

Vector Calculus, 37

vector derivative, 92

*vector product*, 79, 98

vectors, 10, 11, 79, 80, 81, 82, 98

Vectors, 11

*velocity*, 7

*verbs*, 11, 12

vocabulary, 15

*volume derivative*, 40

*volume of revolution*, 105

VOLUMES, 104

*weighted sums*, 96

Word Problems, 86



