FORMS V AND VI

FIRST, SECOND AND THIRD LEVELS

Curr. Coll.

LB
1617
.A82
.N43

BOARD OF SENIOR SCHOOL STUDIES

MATHEMATICS SYLLABUS

AND

NOTES TO THE SYLLABUS

FORM V AND FORM VI

4 UNIT COURSE

(*Unit value of* 4 *Unit Course in Fifth Form is* 3 *Units*)

(*Approved by the Board on* 3rd *October*, 1973)

Issued by
THE NEW SOUTH WALES
DEPARTMENT OF EDUCATION
for the
BOARD OF SENIOR SCHOOL STUDIES

c 48081 *a*

MACQUARIE
UNIVERSITY
LIBRARY

MATHEMATICS SYLLABUS

4 UNIT COURSE

The Board recognizes that the aims and objectives of the syllabus may be achieved in a variety of ways and by the application of many different techniques. Success in the achievement of these aims and objectives is the concern of the Board which does not, however, either stipulate or evaluate specific teaching methods.

INTRODUCTION

This syllabus is designed for students with a special interest in mathematics who have shown that they possess special aptitude for the subject. It represents a distinctly high level in school mathematics involving the development of considerable manipulative skill and a high degree of abstract formulation of the fundamental ideas of arithmetic, algebra and calculus. These topics are treated in some depth. Thus the course provides a sufficient basis for a wide range of useful applications of mathematics and an adequate foundation for the further study of the subject.

AIMS AND OBJECTIVES

The general aim is to present mathematics as a living art which is intellectually exciting, aesthetically satisfying, and relevant to a great variety of practical situations.

Specific aims of the course are:

(a) To give an understanding of the nature and role of deductive reasoning.

(b) To give an understanding of basic mathematical ideas such as variable, function, limit, mathematical induction, etc.

(c) To provide the mathematical background and techniques necessary for concurrent study of science, or for further studies in mathematics.

(d) To offer a programme which will be of interest and value to pupils with the highest levels of mathematical ability at the stage of the Higher School Certificate and which will present some challenge to such pupils.

(e) To develop useful and important mathematical techniques appropriate to these levels of ability.

(f) To introduce the beginnings of some relatively modern topics in mathematics so as to give some idea of the new fields of mathematics studied at the tertiary level.

For the achievement of these aims, the following points are important:

(i) Understanding of the basic ideas and precise use of language must be emphasized.

(ii) A clear distinction must be made between results which are proved, and results which are merely stated or made plausible.

(iii) Where proofs are given, they should be carefully developed, with emphasis on the deductive processes used.

c 48081—1

SYLLABUS

This course is made up of two parts, the first (I) being the whole of the syllabus for the 3 Unit Course and the second (II) being those additional items listed.

PART I

† Students are *not* required to reproduce proofs of results contained in items preceded by this symbol.

1. THE REAL NUMBER SYSTEM

(a) Review of rational numbers.

(b) Irrational numbers. Surds. Rationalization of denominator of fractions in surd form.

† (c) Ordering of real numbers. Review of inequalities. Absolute values. Inequalities involving absolute values.

2. THE CONCEPT OF A FUNCTION, AND ITS REPRESENTATION IN ANALYTICAL GEOMETRY

(a) Dependent and independent variable. Functional notation. Range and domain.

(b) The graph of a function. Simple examples.

(c) Locus and equation. Region and inequality. Simple examples.

3. TRIGONOMETRIC RATIOS: REVIEW AND SOME PRELIMINARY RESULTS

† (a) Review of the trigonometric ratios, using the unit circle.

† (b) Trigonometric ratios of $-\theta$, $90° - \theta$, $180° \pm \theta$, $360° \pm \theta$.

† (c) Sine and cosine rules for a triangle. Area of a triangle, given two sides and the included angle.

4. THE LINEAR EXPRESSION AND THE STRAIGHT LINE

(a) Review: Linear expression; linear equation $ax + b = 0$ and its root. Graph of $y = ax + b$.

(b) Slope of a line. Equation of a line (i) passing through a given point with given slope, (ii) passing through two given points. Parallel lines. General equation $Ax + By + C = 0$.

(c) Review: intersection of two lines and solution of two linear equations in two unknowns. Regions determined by two lines and sets determined by two linear inequalities.

† (d) Perpendicular lines. Perpendicular distance of a point from a line.

† (e) Midpoint formula; ratio formula.

(f) Proofs of geometrical results involving lines.

5. SERIES, SEQUENCES, AND THE PRINCIPLE OF MATHEMATICAL INDUCTION

(a) Review of index laws for rational indices. Arithmetic and geometric sequences and series.

(b) The principle of mathematical induction.

(c) Intuitive statement of the limit of the sum of a geometric series as the number of terms increases indefinitely.

(d) Sequences. Intuitive idea of the limit of a sequence.

† (e) Formal definition of the limit of a sequence (to be treated lightly without drill).

(f) Proof of the identity, for positive integral n

$$x^n - c^n = (x - c)(x^{n-1} + cx^{n-2} + \; . \; . \; . + c^{n-1}).$$

6. TANGENT AND DERIVATIVE

(a) Informal discussion of continuity.

(b) The notion of the limit of a function; (†) formal definition of limit and continuity. Limits of $f + g$, $f - g$, fg.

(c) Gradient of a secant.

(d) Tangent as the limiting position of a secant.

(e) Formal definition of the derivative of $f(x)$ at the point where $x = c$.

(f) The gradient function. Notations $f'(x)$, $\frac{dy}{dx}$, $\frac{d}{dx}$.

(g) Differentiation of x^n for positive integral n. The tangent to $y = x^n$.

† (h) Differentiation of $x^{\frac{1}{2}}$ from first principles. For the two functions u and v, differentiation of cu (c constant), $u + v$, $u - v$, uv. The composite function rule.

† (i) Differentiation of: general polynomial; x^n for n rational; $1/f(x)$; $[f(x)]^{1/n}$; quotient of two functions.

7. THE QUADRATIC POLYNOMIAL AND THE PARABOLA

(a) Quadratic polynomial $ax^2 + bx + c$. Graph of a quadratic function. Roots of a quadratic equation.

(b) General theory of quadratic equations. Relation between roots and coefficients. The discriminant.

(c) Classification of quadratic expressions; identity of two quadratic expressions.

(d) Equations reducible to quadratics.

(e) The parabola defined as a locus. The equation $x^2 = 4Ay$. Tangents and normals.

(f) The parabola — parametric equations. Chord of contact. Simple geometric properties of the parabola. Simple locus problems.

8. CHANGE OF CO-ORDINATE SYSTEMS. TRANSFORMATIONS

(a) Change of origin without change of direction of axes.

(b) Transformation of $y = ax^2 + bx + c$ to the form $Y = aX^2$.

† (c) Concept of invariance under a transformation.

9. GEOMETRICAL APPLICATIONS OF DIFFERENTIATION

(a) Significance of the sign of the derivative.

(b) Stationary points on curves.

(c) The second derivative $f''(x)$, $\dfrac{d^2y}{dx^2}$.

(d) Geometrical significance of the second derivative.

(e) Use of the calculus in sketching the curve $y = ax^2 + bx + c$.

(f) The sketching of simple curves. Simple problems on maxima and minima.

(g) Tangents and normals to curves.

(h) The primitive function and its geometrical interpretation.

10. INTEGRATION

(a) The definite integral.

† (b) Approximate methods: mid-ordinate rule and Simpson's rule.

† (c) The relation between the integral and the primitive function.

(d) Applications of definite integration to areas and to volumes of solids of revolution.

11. THE LOGARITHMIC AND EXPONENTIAL FUNCTIONS

† (a) Definition and properties of the natural logarithmic function.

(b) General discussion of inverse functions. The relation $\dfrac{dy}{dx}\dfrac{dx}{dy} = 1$.

† (c) The exponential function.

12. THE TRIGONOMETRIC FUNCTIONS

(a) Circular measure of angles. Angle, arc, sector.

(b) The trigonometric functions and their graphs.

(c) Periodicity and other simple properties of the trigonometric functions.

† (d) Expressions for $\sin(x \pm y)$, $\cos(x \pm y)$, $\tan(x \pm y)$.

Deductions of $2 \sin A \cos B = \sin(A + B) + \sin(A - B)$, etc.

And $\sin x + \sin y = 2 \sin \frac{x + y}{2} \cos \frac{x - y}{2}$, etc.

Formal proofs should be given.

(e) The angle between two lines.

(f) Approximations to $\sin x$, $\cos x$, $\tan x$, when x is small.

The result $\lim_{x \to 0} \frac{\sin x}{x} = 1$.

(g) Differentiation of $\sin x$, $\cos x$, $\tan x$.

(h) Primitive functions of $\sin(ax)$, $\cos(ax)$.

13. THE INVERSE TRIGONOMETRIC FUNCTIONS

(a) The inverse trigonometric functions.

(b) The graphs of $\sin^{-1}x$, $\cos^{-1}x$, $\tan^{-1}x$.

(c) Simple properties of the inverse trigonometric functions.

(d) The derivatives of $\sin^{-1}x$, $\cos^{-1}x$, $\tan^{-1}x$.

(e) Differentiation of $\sin^{-1}(x/a)$ and $\tan^{-1}(x/a)$, and the corresponding integrations. Simple illustrative examples.

14. ANALYTICAL GEOMETRY IN THREE DIMENSIONS

(a) Review of the elements of solid geometry. Points, planes, and surfaces. Relations between points, lines, and planes.

(b) Points and lines treated analytically. Distances; direction cosines; angle between two lines through the origin; perpendicular lines; equation of a line.

(c) Simple notions of surfaces. The sphere. Cylinders.

(d) The plane; canonical and general form of its equation. Parallel and perpendicular planes. Regions specified by linear inequalities.

15. APPLICATIONS OF THE CALCULUS TO THE PHYSICAL WORLD

(a) Velocity and acceleration as derivatives.

(b) Determination of velocity and position, given the acceleration and initial conditions, for the cases:

(i) $d^2x/dt^2 = \text{constant}$;

(ii) $d^2x/dt^2 = f(t)$;

(iii) $d^2x/dt^2 = g(x)$.

(c) Simple harmonic motion from $d^2x/dt^2 = -n^2x$.

(d) Motion of a particle under Hooke's law.

(e) Parabolic motion under gravity. Parametric and Cartesian equations of the path.

16. FURTHER TRIGONOMETRY AND CALCULUS

(a) Expressions for sin θ, cos θ, tan θ, in terms of tan $\frac{1}{2}\theta$.

(b) *Simple* trigonometric equations.

(c) Derivatives of cot θ, sec θ, and cosec θ.

17. FURTHER POLYNOMIALS

(a) Definitions of polynomial, degree, monic polynomial. Graphs of simple polynomials.

(b) The remainder and factor theorems.

(c) The roots and coefficients of a polynomial equation.

(d) Approximation methods for determining roots of equations:

(i) Halving the interval;

(ii) Newton's method.

18. THE BINOMIAL THEOREM

(a) Expansion of $(1 + x)^n$ for $n = 2, 3, 4$. Pascal triangle. Proof of the Pascal triangle relations. Extension to expansion of $(a + x)^n$.

† (b) Proof by mathematical induction of the formula for ${}^nC_k$ (also denoted by $\binom{n}{k}$ or ${}_nC_k$).

(c) Combinatorial interpretation of the binomial coefficients.

19. THEORY OF PROBABILITY

(a) Statistical regularity.

(b) Random experiments. Finite sample space. Simple, composite, and mutually exclusive events. Algebra of events. The opposite (complementary) event. Probability of an event. Theorem of total probability.

(c) Two stage random experiments. Independent events. The product rule.

(d) Systematic enumeration in a finite sample space. Definitions of ${}^nP_r$, ${}^nC_r$.

(e) Binomial probabilities and the binomial distribution.

(f) The general notion of a random variable and probability distribution, illustrated mainly in connection with the binomial distribution. Expectation. The expectation of a binomial variable.

PART II

1. THEORETICAL ARITHMETIC

(*a*) *The Integers or Whole Numbers*

Review of the fundamental properties of the set of integers. Factorisation and divisibility. Primes and composite numbers. The division transformation. Highest common factor. Euclid's algorithm and deductions from it. The unique factorisation into primes.

(*b*) *The Rational Numbers*

Review of the origin of the rational numbers. Comparison of the set of rationals with the set of integers. Decimal representation of rationals.

(*c*) *The Real Numbers*

Set of all decimals as an ordered set in which arithmetical operations may be defined so as to give this set the same essential structure as that of the set of rationals. Comparison of the set of real numbers with the set of rational numbers. The monotonic principle of convergence. The definition of a^x for $a > 0$.

2. ALGEBRA

(*a*) *Polynomials*

The general notion. The "indeterminate", coefficients, degree. The influence of the coefficient set on the structure of the theory. Addition and multiplication of polynomials. Factorisation and divisibility. Prime and composite polynomials. Highest common factor. The division transformation. Euclid's algorithm and deductions from it. The unique factorisation into primes.

(*b*) *Polynomials as Functions*

Polynomial equations. Roots. The remainder and factor theorems. Relation between the number of roots and the degree. The identity of two polynomials. Relations between the roots and the coefficients in the case of a polynomial which is completely reducible to linear factors.

(*c*) *Rational Functions*

Partial fractions.

3. ELEMENTARY DYNAMICS OF A PARTICLE

Rectilinear motion of a particle described by a functional relation $x = f(t)$. Velocity and acceleration as differential coefficients. Kinematical formulae for uniformly accelerated motion. Newton's laws of motion—first and second laws. Differential equations of motion in one dimension. Momentum, work and energy. Kinetic and potential energy. Equation of energy. Resisted motion, terminal velocity. Simple harmonic motion. Projectiles. Motion in two dimensions under gravity.

Only TWO of the following items 4, 5, 6 need be covered.

4A. SEQUENCIES AND SERIES

Notion of a limit in connection with sequences. The monotonic principles of convergence. Convergence of series, including a discussion of the series $\sum \frac{1}{n^p}$ with proof. The comparison test. Conditional and absolute convergence.

4B. COMPLEX NUMBERS

Origin of the idea in connection with the solution of quadratic equations. Exploration of the possibilities by introduction of the imaginary i, $i^2 = -1$, and recognition of the formal theory as a calculus of ordered pairs of real numbers. Modulus, argument, conjugate. Geometric representation of addition and multiplication of complex numbers in the Argand diagram. The relations

$$|z_1 + z_2| \leqslant |z_1| + |z_2| \text{ , } |z_1 z_2| = |z_1| . |z_2|,$$

$$\arg z_1 z_2 = \arg z_1 + \arg z_2.$$

Polynomials over the complex field.

5. CALCULUS

(*a*) *The Function Concept*

The notion of related variables. Dependent and independent variables. Functions as mappings or correspondences. The functional notation. Graphical representation. Inverse functions.

(*b*) *Continuous Functions of a Real Variable*

Meaning of continuity at a point and in an interval. Statement of the fundamental properties without proof.

(*c*) *Tangents to Curves*

Gradient or slope. Calculation of the gradient as a limit. Differentiation. The derived function. The significance of the sign of the derived function. Maxima and Minima. Rolle's theorem. The mean value theorem. Differentiation of combinations of functions and of composite functions. Differentiation of rational functions and of simple irrational functions. The relation

$$\frac{dy}{dx} \cdot \frac{dx}{dy} = 1$$

in connection with inverse functions. Geometric meaning on the graph. Calculation of derivative of functions defined implicitly in cases which do not involve partial differentiation. Second derivatives, inflexions and curve tracing.

(*d*) *The Problem of Areas*

Figures bounded by curved lines. The definite integral defined as the limit of a sum. Formal properties of the definite integral. Indefinite integrals and the Fundamental Theorem of the calculus. Calculation of areas and volumes.

(*e*) *The Exponential and Logarithmic Functions*

The theory of these functions should be derived from one or other of the differential equations

$$\frac{dy}{dx} = y, \text{ or } \frac{dy}{dx} = \frac{1}{x}.$$

The definition of e. Derivation of the exponential and logarithmic series. Calculation of e and of $\log_e 2$.

(*f*) *Arcs of Curves; the Trigonometric Functions*

Calculation of length of arc. Length of circular arc. Measure of an angle. The radian. Definition of π. The trigonometric or circular functions defined by reference to the unit circle. Graphs of these functions. Differentiation of the trigonometric functions. The addition formulae and allied formulae.

The inverse circular functions and their "principal" values. Differentiation of $\sin^{-1} x$, $\cos^{-1} x$ and $\tan^{-1} x$.

(*g*) *Integration*

Table of standard integrals. Indefinite integration. Integration of simple rational functions. Integration of simple irrational functions involving the square root of a quadratic expression. Integration by Parts. Integration by change of variable. Integration of simple functions using the above methods.

6. LINEAR TRANSFORMATIONS IN THE PLANE; INTRODUCTION TO MATRIX ALGEBRA

Rotations, reflections, displacements as geometric operators. Composition of operators. Cartesian form of rotations, etc. 2×2 matrices as operators. General form of 2×2 matrices and the rules for their addition and multiplication. Singular and non-singular matrices. The inverse matrix. The transposed matrix. Matrix treatment of rotation of axes. Characteristic polynomial of a symmetric 2×2 matrix: eigenvalues and eigenvectors. Reduction of a quadratic form to a sum of squares; ellipse, hyperbola; lengths and directions of principal axes from eigenvalues and eigenvectors. The focus–directrix property as an alternative definition of conics.

c 48081—2

NOTES TO THE SYLLABUS

The notes on the 3 Unit Course, presented separately, apply to part I of this course except insofar as they restrict the range of a topic which is treated more fully in part II.

The notes on the additional items which form part II are contained below.

1. THEORETICAL ARITHMETIC

(*a*) *The Integers or Whole Numbers*

The fundamental properties will be known from the work of the first four years, but may be summarily reviewed. The essential facts are:

(i) The set of integers is *ordered.*

(ii) Two binary operations, *addition and multiplication*, are defined within the set and these obey "algebraic" laws which have been fully discussed in earlier work.

(iii) There are relations between the properties (i) and (ii) and these may be called the "laws for inequalities". There are two principal relations and these are not independent. They are:

a, b, c being integers, and $a < b$ then $a + c < b + c$;

and

a, b, c being integers, $a < b$ and $c > 0$ then $ac < bc$.

These are the basic facts for work with inequalities.

Divisibility among the integers is defined as follows:

Given a, $b \neq 0$, if there is an integer c such that $a = bc$ we say "b divides a" and indicate this by $b|a$. "b does not divide a" is indicated by $b \nmid a$. We say also "b is a factor of a"; and also c is a factor of a. The two theorems

$b|a$, and $b|a' \rightarrow b|(a + a')$,

and

$b|a$, and $c|b \rightarrow c|a$,

are consequences of the algebraic laws. For much of the discussion it is convenient to restrict attention to the positive integers and their positive factors. If we do this the factorization $a = bc$ implies $1 \leqslant b \leqslant a$ and $1 \leqslant c \leqslant a$. So any positive a can have only a finite number of positive factors. If $a > 1$ it has the two factors 1 and a; and there are numbers a which have no other factors.

Such numbers are prime; and numbers which are not prime are called composite. A composite number a (>1) has a divisor d with $1 < d < a$. It follows that any number >1, if not prime, can be expressed as a product of a finite number of prime factors.

If $\{p_1, \ldots, p_n\}$ is a finite set of primes and we form the number $N = 1 + p_1 p_2 \ldots p_n$ then we see $p_1 \nmid N$; for if $p_1|N$ we infer $p_1|1$, and so $1 = ap_1 > a \geqslant 1$, $1 > 1$. Similarly $p_i \nmid N$, $i = 1, \ldots, n$. Thus N has a prime divisor $p \leqslant N$, which must be different from each of the p_i. Thus we infer that there are infinitely many primes.

If $a > 1$ is composite we may write $a = bc$ with $b \geqslant c > 1$. Hence $a = bc \geqslant c^2$. Thus a has a divisor c whose square does not exceed a. In particular any composite a will have a prime divisor p with $p^2 \leqslant a$. To test whether a given number $a > 1$ is prime it is sufficient to test it for divisibility by all the primes p for which $p^2 \leqslant a$.

Given $b > 0$, there is a unique largest integer q such that $qb \leqslant a$. Then $qb + b = (q + 1)b > a$. If we set $a = qb + r$ then r is uniquely determined and $0 \leqslant r < b$. We call q the quotient and r the remainder when a is divided by b. The relation

$$a = qb + r, \quad 0 \leqslant r < b$$

is called the division transformation.

The highest common factor of two integers a, b is the largest positive number h such that $h|a$, $h|b$. If $h = 1$, we say a, b are relatively prime.

If now $a = qb + r$, $0 \leqslant r < b$ we infer $h|r$.

Given two positive integers $a_0 \geqslant a_1$ we can form the Euclid algorithm

$$a_0 = q_1 a_1 + a_2, \quad 0 \leqslant a_2 < a_1$$

$$a_1 = q_2 a_2 + a_3, \quad 0 \leqslant a_3 < a_2$$

$$\cdot \; \cdot \; \cdot \qquad \cdot \; \cdot \; \cdot \qquad \cdot \; \cdot \; \cdot \qquad \cdot \; \cdot \; \cdot$$

$$a_{r-2} = q_{r-1}\, a_{r-1} + ar, \quad 0 \leqslant a_r < a_{r-1}$$

$$a_{r-1} = q_r a_r.$$

The remainders decrease and the process ends after r steps when the last remainder is zero. Then, by a simple argument, $a_r = h$, the H.C.F. of a_0, a_1.

Now working "up" the algorithm we find

$$h = a_r = A_0 a_0 + A_1 a_1$$

where A_0 and A_1 are integers. Changing the notation we see that if a, b are any positive integers there are integers A, B such that

$$Aa + Bb = h = \text{H.C.F. of } a, b.$$

In this statement the restriction to positive a, b is obviously unnecessary. When a, b are relatively prime, $h = 1$ and we have the relation

$$Aa + Bb = 1.$$

Conversely this relation implies that a, b are relatively prime. If now $a|bc$ we have

$$Aac + Bbc = c$$

but a divides both the terms on the left, so $a|c$. If a divides a product and is prime to one factor then it divides the other. Thus if the prime p divides the product of primes $p_1 \ldots p_n$ then we must have $p = p_i$ for some i. For if not we derive successively

$$p|p_2 \ldots p_n, \; p|p_3 \ldots p_n, \; \ldots, \; p|p_n$$

and this is a contradiction. The fundamental theorem on the uniqueness of the prime factorisation of any number a now follows easily.

(*b*) *The Rational Numbers*

The fundamental properties concerning the order and the arithmetic operations in the set of rational numbers will be known. They are precisely as detailed above for the set of integers with one additional property, namely, given any rational number $a \neq 0$ there is one and only one rational number b such that $ab = 1$. This b is denoted by a^{-1}.

Given any two rationals $a, b \neq 0$ there is just one rational c such that $a = bc$. Thus if we take $c = b^{-1}a$ then $bc = b(b^{-1}a) = 1a = a$. One consequence of this fact is that there is no theory of divisibility among the set of rationals in any way similar to what we have considered for the integers.

Between any two rationals there is another rational and hence infinitely many rationals. Thus if $a < b$ we have

$$2a = a + a < a + b < b + b = 2b.$$

$$\therefore \quad a < \frac{a+b}{2} < b\,.$$

It should be remembered that the rationals are constructed from the integers as ordered pairs of integers on which the operations of addition and multiplication are defined. Mathematicians do this formally, but it should be noted that school children do exactly the same thing, only less formally.

Thus for instance the "fraction" 2/3 is literally an ordered pair of integers separated by a solidus instead of by a comma—and 2/3 is identified with 4/6, etc. Anyone who knows the formal construction will see that it is not much ahead of the school construction, and that the biggest difference is in the mental attitude to the process.

The decimal representation of the rational numbers is also familiar since primary school. This should be reviewed noting especially that the decimals obtained either terminate or recur. In the conversion of rationals to their decimal representation such an infinite decimal as

$$2{\cdot}202002000200002\ldots$$

which neither terminates nor recurs does not arise.

(*c*) *The Real Numbers*

The need for a further extension of the set of numbers beyond the rationals may be illustrated by the simplest arithmetic problems. We may seek, among the rationals, a number x which satisfies any of the equations such as

$$x^2 = 2\,, \quad x^2 = 3\,, \quad x^3 = 4\,.$$

But we will seek in vain! It is easy to prove that a rational number x does not satisfy any one of these equations. In earlier work we have been accustomed to write, as solutions of these equations, $x = \sqrt{2}$, $x = \sqrt{3}$, $x = \sqrt[3]{4}$. If there is any sense in this formalism we have not yet explained it. Our purpose should be to construct, if possible, a number system which will contain all the numbers (i.e., the rationals) that we have already, and others besides which may satisfy these equations. This extended number system should have the essential properties of the system of rationals as we have explained them above.

There are in fact several extensions of the set of rationals which have been studied and which are of great interest. If, for instance, we were only concerned to have a number system in which all polynomial equations have solutions then we would be satisfied to introduce the so-called "algebraic" numbers. If, however, we wish to establish the customary limit processes involved in, say, the calculus we require a much wider extension.

The germ of the appropriate extension has already presented itself in primary school work, in the decimal representation used there. The real numbers may be introduced as the set of all decimals—not merely those which terminate or recur. It can be easily explained how this set is ordered; how addition and multiplication is defined within the set; and how, when these definitions are properly given, (and the proper way is "nearly obvious") the set of real numbers has all the properties which we have picked out above as the essential properties of the set of rationals. This is precisely our justification for calling the set of decimals, with the structure we have imposed on it, a set of numbers; and the word real is used merely as a name. The numbers of this set which do not terminate or recur are the irrational numbers.

But the set of real numbers, now introduced, has further important properties which are not possessed by the set of rationals; naturally it is these new properties which gives this set of reals their importance. A fundamental property of this kind (from which all the other important ones follow) is the monotonic principle of convergence:

An increasing sequence of numbers, bounded above, is convergent.

While this is easy enough to prove, such a formal proof is not required at this stage. What *is* required is that the principle should be *accurately stated*, discussed and explained sufficiently so that pupils feel reasonably convinced of its accuracy, and so that they can use it and *argue from it accurately*. If necessary, the principle could be set up as a postulate, to be accepted for the present, and to be examined more fully at a later stage. There would be no violation of logic in this procedure. It can be shown by examples that the principle does *not* hold within the set of rational numbers.

Now, using this principle, we can prove that each of the equations $x^2 = 2$, etc., quoted above has a solution among the real numbers; the solutions are of course irrational and as a matter of notation they are denoted by $\sqrt{2}$, $\sqrt{3}$, etc. By exactly the same argument we can now show that any equation like

$$x^q = a^p$$

where $a > 0$ is real and p, q are integers has exactly one positive solution. This is denoted by $a^{p/q}$. Here, and for the first time, we have a satisfactory definition of the meaning of fractional exponents. The usual "laws of indices" for such exponents now follow by the usual arguments. For example for positive integers p, q, r, s, if

$$x = a^{p/q}, \quad y = a^{r/s}$$

we have

$$(xy)^{qs} = x^{qs}y^{qs} = a^{ps}a^{rq} = a^{ps+rq}.$$

So $xy = a^{(ps+rq)/qs} = a^{(p/q)+(r/s)}$.

Hence $a^{p/q} \,.\, a^{r/s} = a^{(p/q)+(r/s)}$.

2. ALGEBRA

In a preliminary discussion it may be worthwhile to explain that in the algebra associated with number systems it is the "algebraic" properties of the numbers which is of the greatest importance. We noted that the rational numbers and the real numbers had the same algebraic properties while the set of integers has most of the same properties. It may now seem useful to introduce the notion of a *field* in the algebraic sense and possibly also the notion of an *integral domain*.

(*a*) *Polynomials*

Polynomials may be formed over various sets—and the sets usually have some kind of algebraic structure. For the various useful structures the theories of polynomials have much in common but they differ in some important points of detail. It is not feasible to ask beginners to absorb all the useful theories at the same time so we consider here only sets which have the algebraic structure of a *field*. More definitely the beginner may contemplate the set of rationals or the set of reals.

If $a_0, a_1, \ldots, a_n$ are any numbers in one of these fields the expression

$$a_0 + a_1x + \ldots + a_nx^n$$

is called a "polynomial in x" over the field. In the most abstract approach the letter x has no meaning and the polynomial is a purely formal expression—merely a device for studying ordered finite subsets $\{a_0, \ldots, a_n\}$ of numbers. Here $a_0, \ldots, a_n$ are called the coefficients of the polynomial, x is called the "indeterminate" and if $a_n \neq 0$ the polynomial is said to have degree n. We impose a structure on the set of polynomials by defining addition and multiplication formally in a way which need not be specified here. The usual algebraic laws for addition and multiplication apply. Subtraction is possible but division is not always possible. The polynomials obey the same set of algebraic laws as do the integers—technically polynomials over a field form an integral domain. The zero polynomial, $a_r = 0$ for all r, has no degree but apart from this the degree of a product is equal to the sum of degrees of the factors. It follows that the product of non-zero polynomials is not zero.

Ideas like divisibility and factorisation may be applied to the set of polynomials just as to the set of integers—and the theory is constructed in much the same way. Given $A = A(x)$, $B = B(x)$ $(\neq 0)$ any two polynomials in x, if there is a polynomial $C(x)$ such that $A = BC$ we say "B divides A" and write $B|A$.

We say also B is a factor of A; and of course if $C \neq 0$, C is a factor of A. The theorems

$$B|A \text{ and } B|A' \Rightarrow B|(A + A')$$

and

$$B|A \text{ and } C|B \Rightarrow C|A$$

follow as in arithmetic.

If $A = BC$ and if $\deg B < \deg A$ and $\deg C < \deg A$

then B, C are "proper" factors of A, and A is composite or reducible. If A has no proper factors we call A an irreducible polynomial or a prime polynomial. Whether or not a given polynomial is reducible depends very materially on the field over which the set of polynomials is constructed. For example it is easy to prove that $x^2 - 2$ is reducible when the underlying field is the set of real numbers:

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

but it is irreducible when the underlying field is the set of rationals. The proof is trivial and depends only on the fact that no rational number has its square equal to 2. Also, it is easily seen that $x^2 + 1$ is irreducible over the field of real numbers.

A polynomial, if not prime, can be expressed as a product of primes. The long division process of classical algebra can be used to establish the division transformation. Given two polynomials in x, say A and B, suppose that deg $A \geqslant$ deg $B > 0$. We can find two polynomials q and r such that $A = qB + r$ and either $r = 0$ or deg $r <$ deg B. Such q, r are uniquely determined.

The H.C.F. of two polynomials A and B is a polynomial H of highest degree with coefficient of the highest power equal to 1 such that $H|A$ and $H|B$. Euclid's algorithm for the H.C.F. is constructed now almost exactly as in the case of the set of integers and we make similar deductions from it. Thus if A, B are given polynomials and H is their H.C.F. there are polynomials P, Q such that

$$PA + QB = H$$

If $H = 1$, we say A, B are relatively prime and

$$PA + QB = 1.$$

Conversely, a relation of this kind implies that A, B are relatively prime. Again, also proved as before, if A, B are relatively prime,

$$A|BC \Rightarrow A|C.$$

Finally we may show that the factorisation of a polynomial into irreducible factors is essentially unique, i.e., apart from the introduction of "constant" factors.

(*b*) *Polynomials as functions*

Given a polynomial in the indeterminate x over a field F, say

$$P(x) = a_0 + a_1x + \ldots + a_nx^n$$

if we replace x by b, an element of F, we get

$$a_0 + a_1b + \ldots + a_nb^n.$$

This is an element of F; it is denoted by $P(b)$.

To each element b in F the polynomial determines in this way a definite element of F, i.e., it specifies a function on F into F. Further the definitions of addition and multiplication of polynomials have been such that the polynomial relations

$$P(x) + Q(x) = R(x), \quad P(x)Q(x) = S(x)$$

imply the relations

$$P(b) + Q(b) = R(b), \quad P(b)Q(b) = S(b).$$

In consequence any relation between polynomials derived by the use of these operations yields a corresponding relation between the associated functions. Expressed briefly this means that, in any relation between polynomials over F, we may substitute any element of the set F for the indeterminate x. It is this principle of substitution which is so important; if it were not true any abstract theory of polynomials would be of very much less use than it is.

Consider for instance the division transformation when $B\,(= x - a)$, is of the first degree. Then

$$A = q(x - a) + r$$

and either $r = 0$, or it is a polynomial of degree 0; we may say r is an element of the field. Substituting $x = b$ we get

$$A(b) = q(b) \,.\, (b - a) + r,$$

a relation which holds for every b in F.

In particular, taking $b = a$ we get $A(a) = r$.

So, in this division the remainder r is $A(a)$. This is the so-called "remainder theorem". When $A(a) = 0$, $r = 0$ and we get

$$A(x) = q(x) \,.\, (x - a)$$

so

$$(x - a)|A(x)$$

in the sense of polynomial division. Of course, the converse statement holds also. This is the "factor theorem".

Then also a is called a root of the polynomial $A(x)$. We see that a polynomial over F which is irreducible over F and of degree > 1 has no roots in F.

Again, if a polynomial $A(x)$ of degree n has n roots $a_1, \ldots, a_n$ we find the factorisation

$$A(x) = a(x - a_1) \ldots (x - a_n).$$

Hence for b in F,

$$A(b) = a(b - a_1) \ldots (b - a_n).$$

This means that for $b \neq a_i$, each i, $A(b) \neq 0$.

Hence a polynomial of degree n cannot have more than n roots. Now we may show that two different polynomials $A(x)$ and $B(x)$ cannot specify the same function in the rational, real, or complex fields. For then the polynomial

$$A(x) - B(x)$$

has a degree $n \geqslant 0$. Thus $A(b) - B(b) \neq 0$ except for at most n values of b. If F contains more than n elements then $A(b)$ and $B(b)$ are not the same functions on F.

If $A(x)$ is completely reducible to a product of n linear factors, and if it has highest coefficient 1 we have

$$A(x) = (x - a_1) \ldots (x - a_n).$$

The roots of $A(x)$ are $a_1, \ldots, a_n$; and the coefficients of $A(x)$ are

$$1, -\Sigma a_i, +\Sigma a_i a_j, \ldots, (-1)^n a_1 \ldots a_n.$$

Of course if a polynomial is not completely reducible there can be no sense in speaking of a relation between its roots and its coefficients.

(c) *Rational Functions. Partial Fractions*

If $P(x)$, $Q(x)$ are two polynomials the expression

$$\frac{P(x)}{Q(x)}$$

is called a rational function. Of course properly speaking the word function should not be used in a purely algebraical context since the mathematical meaning of the word function is not involved at all. We shall continue with the old-fashioned use. The sum of two rational functions is defined by the formula

$$\frac{P_1(x)}{Q_1(x)} + \frac{P_2(x)}{Q_2(x)} = \frac{P_1Q_2 + P_2Q_1}{Q_1Q_2},$$

and so is a rational function. The problem of "partial fractions" is to reverse this process. We state here only the main lemma. If f is a polynomial, P, Q relatively prime polynomials, and if $\deg f < \deg PQ$ then there is a unique separation into partial fractions

$$\frac{f}{PQ} = \frac{A}{P} + \frac{B}{Q}$$

where deg A $<$ deg P, deg $B <$ deg Q.

Since P, Q are relatively prime we find A_1, B_1 such that

$$B_1P + A_1Q = 1.$$

Multiply by f and set $A = A_1f$, $B = B_1f$, then

$$BP + AQ = f.$$

Hence

$$\frac{f}{PQ} = \frac{A}{P} + \frac{B}{Q}\,.$$

By long division we may remove the "integral parts" of $\frac{A}{P}$ and $\frac{B}{Q}$ and then these must cancel from considerations of degree.

3. ELEMENTARY DYNAMICS OF A PARTICLE

Throughout the work on this topic the emphasis should be on the use of the calculus to express the relations of mechanics, and not on the solution of complicated problems. It is not intended that a large number of formulae should be committed to memory but rather that the student should be able to derive any result from first principles whenever it is required.

Rectilinear motion of a particle may be described by a functional relation $x = f(t)$ where x is distance measured from an origin and t is time from a given instant. Velocity and acceleration will be defined as differential coefficients $\dot{x}$ and $\ddot{x}$ where the dots indicate differentiation with respect to t.

The kinematical formulae

$$\ddot{x} = a, \quad \dot{x} = u + at, \quad x = ut + \tfrac{1}{2}at^2$$

for uniformly accelerated motion will be derived.

The classical statement of Newton's Laws of motion (excluding the third law) should be given. By choice of units the Newtonian formula $F = ma$ follows. For motion of a particle in a straight line line this becomes

$$m\frac{d^2x}{dt^2} = F$$

where F is the force acting on the particle at time t in the direction in which x increases. If F is known at each instant of the motion this is a differential equation to determine the motion.

Setting $v = \dot{x}$ the equation may be written

$$mv \frac{dv}{dx} = F$$

or

$$\frac{d}{dx} (\tfrac{1}{2} mv^2) = F .$$

Then, integrating from x_1 to x_2 with corresponding velocity values v_1, v_2 we get

$$\tfrac{1}{2}mv_2^2 - \tfrac{1}{2}mv_1^2 = \int_{x_1}^{x_2} F\, dx.$$

On the left we have the increment in the quantity $\tfrac{1}{2}mv^2$, which is called the kinetic energy of the particle of mass m moving with velocity v; on the right we have the work done by the force acting on the particle during its displacement.

Thus

Increase in kinetic energy = work done by the external force acting on the particle.

In cases where the force acting at a position x depends on this position only, then F is a function of x, and

$$V = \int_a^x F\, dx$$

is a function of x also. V is called the potential energy of the particle at x.

Then

$$\int_{x_1}^{x_2} F\, dx = V_1 - V_2$$

and the energy equation may be written

$$\tfrac{1}{2}mv_1^2 + V_1 = \tfrac{1}{2}mv_2^2 + V_2.$$

So the quantity

$$\tfrac{1}{2}mv^2 + V$$

is constant throughout the motion. This is a statement of the conservation of (mechanical) energy—the sum of kinetic and potential energies is constant. Of course it is a mathematical theorem derived by integration of the equation of motion—and applies only in the "conservative" field of force considered.

The important cases in elementary work are (i) the case when F is constant (motion under gravity), (ii) the case when F is proportional to x (motion near an equilibrium position) leading in particular to simple harmonic motion defined by the equation

$$\frac{d^2x}{dt^2} = -n^2x ,$$

with solutions

$$x = a \sin (nt + b),$$

$$x = a \cos (nt + b),$$

or $x = A \cos nt + B \sin nt.$

Discussion of resisted motion should be restricted to the case of a particle moving vertically under gravity and subject to a resistance proportional to a power of the speed.

Motion in two dimensions should deal only with the parabolic motion of a projectile under gravity. Find the height of flight and the range on a horizontal plane for given conditions of projection, and the maximum height and maximum range for a given speed of projection.

It is suggested that the work on dynamics should not be treated as a separate unit, but that, as far as possible the applications should be made whenever appropriate to illustrate the use and value of the calculus. It may be recalled that the subject was developed originally in just this context.

4. SEQUENCES AND SERIES. COMPLEX NUMBERS

(*a*) *Sequences and Series*

A precise statement of the convergence of a sequence to a limit should be given.

The following results should be demonstrated, but the proofs of those items marked with a dagger (†) will not be examined:

† (i) If (as $n \to \infty$) $\lim u_n = U$ and $\lim v_n = V$ then

$$\lim (u_n + v_n) = U + V;$$
$$\lim (u_n - v_n) = U - V;$$
$$\lim (u_n v_n) = UV;$$
$$\lim (u_n/v_n) = U/V \quad \text{(given } V \neq 0\text{)}.$$

† (ii) If $u_n \geqslant v_n$ for all values of n and if $v_n \to \infty$ as $n \to \infty$ then $u_n \to \infty$ as $n \to \infty$.

† (iii) If $0 \leqslant u_n \leqslant v_n$ for all values of n and if $v_n \to 0$ as $n \to \infty$ then $u_n \to 0$ as $n \to \infty$.

(iv) If $a > 1$ then $a^n \to \infty$ as $n \to \infty$.

(v) If $0 < a < 1$ then $a^n \to 0$ as $n \to \infty$.

(vi) If $a > 0$ then $a^{1/n} \to 1$ as $n \to \infty$.

(vii) $n^{1/n} \to 1$ as $n \to \infty$.

(viii) $n^p a^n \to 0$ as $n \to \infty$ for $0 < a < 1$.

(ix) $\dfrac{x^n}{n!} \to 0$ as $n \to \infty$.

(x) An absolutely convergent series is convergent.

(xi) A series whose terms have alternating signs and whose absolute values form a decreasing sequence with limit zero is convergent.

It is not intended that the work on sequences and series should lead to sets of complicated examples. A few straightforward examples to illustrate the concepts and results mentioned above and in the syllabus are all that is required. The series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$ is an example of a conditionally convergent series.

(b) *Complex Numbers*

There remain simple algebraic questions which have no answer within the set of real numbers. There is no real number x which satisfies the equation $x^2 + 1 = 0$. We may wish to extend the number system further so that in the new set of numbers equations of this kind have solutions. Here also the extended number system would be required to retain as far as possible the essential properties of the earlier number systems.

Assuming that such an extended system is possible we denote by i a solution of $x^2 + 1 = 0$. So $i^2 = -1$, and we may then solve equations like $x^2 - 2x + 2 = 0$ to find $x = 1 + i$ or $x = 1 - i$. This leads to consideration of expressions of the form $a + ib$ where a, b are real numbers. If we begin again with these elements and define addition and multiplication formally we find a system which has all the *algebraic* properties of the earlier number system. But the system is not ordered. It is easy to explain that we are effectively operating with ordered pairs of real numbers according to rules which may be formally stated.

Any polynomial of degree exceeding 1 over the complex field is reducible; such a polynomial can be expressed as a product of linear factors. This statement is equivalent to what is often called the fundamental theorem of algebra. (Omit proof.)

5. CALCULUS

(a) *The Function Concept*

When two variable quantities (like pressure and volume of a gas, or length and temperature of a metal bar, or distance travelled and time taken for a moving particle) are related in such a way that the value of one of them determines the value of the other we say that they are functionally related. Usually the value of either variable may be taken arbitrarily and then that of the other follows, or is determined by the first. We speak of the dependent and independent variable. If the value of x determines that of y uniquely, we say y is a function of x and write $y = f(x)$. In mathematics we emphasise especially the formal nature of the functional relation, but it is a mistake to forget its practical origin. Formally we have simply a correspondence between values of x and values of y.

If to each element x of a set X we have associated one and only one element y of a set Y then this association or correspondence is a function defined "on X" (its domain) and having its range in Y. If we like, the associated elements may be written as pairs (x, y): so we come to the modern definition of a function simply as a *set of ordered* pairs in which any particular element of X occurs in only one of the pairs. In this formulation the sets X, Y may be quite arbitrary sets and not necessarily sets of numbers. It should be recognized that this modern definition and the older forms mean the same thing.

In our calculus studies the sets X, Y are always sets of real numbers—we are concerned with real functions of real variables.

(b) *Continuous Functions*

The notion of continuity should be explained simply but informally. Assuming $f(x)$ is defined in an interval containing x_0 we define first what is meant by continuity at x_0. We say "$f(x)$ is continuous at x_0 if the values which $f(x)$ takes for values of x near x_0 are nearly equal to $f(x_0)$". A fuller discussion and a formal definition will be found in the Notes on the 3 Unit Course and 2 Unit Course.

Continuity throughout an interval is defined as continuity at each point of the interval.

Graphs of continuous functions are continuous curves, and the general properties of continuous functions should be explained by reference to their graphs.

The fundamental properties are:

(i) if $f(x)$ is continuous in $a \leqslant x \leqslant b$ there is some point x_0 in this interval such that $f(x) \leqslant f(x_0)$ for all x in the interval. This statement is described by saying that a function which is continuous in a closed interval takes a greatest value in the interval. The difference between closed and open intervals should be emphasised.

(ii) If $f(x)$ is continuous in $a \leqslant x \leqslant b$ and if $f(a)f(b) < 0$ then there is some point x_0 between a and b for which $f(x_0) = 0$.

No formal proofs of these statements need be given.

(*c*) *Tangents to Curves*

We may begin by considering continuous curves, being the graphs of continuous functions $y = f(x)$, defined in an interval $a \leqslant x \leqslant b$. Considered geometrically the first problem of the calculus is that of defining what is meant by the tangent to a curve at a given point. Let P be a point on the curve at which we wish to specify the tangent. Take another point Q on the curve and consider the secant PQ. If, as Q moves towards P along the curve, the secant PQ tends to a limiting position or limiting line this line is the tangent at P to the curve. A curve will have a tangent at the point P only if this definition yields a result.

The definition is expressed in informal and geometric language and it should be regarded as a mere preliminary. The use of the description "limiting line" almost begs the principal question. It is therefore necessary to explain just how the definition is to be understood.

Consider the slope or gradient of the secant PQ. If Q has co-ordinates $(x, f(x))$, and P has co-ordinates $(x_0, f(x_0))$ this slope is

$$g(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

To say that the secant PQ has a limiting position must be understood to mean that the slope $g(x)$ tends to a definite limit s, say, as $x \to x_0$. If this is so then the line through P with the slope s is the tangent to the curve at P. Then also we call s the slope of the curve at P.

The question then whether a given curve has a tangent at a particular point P is the same as the question whether a given function has a limit at a given point x_0—the one question is merely a geometric phrasing of the other. When the above limit s exists we say that the function $f(x)$ is "differentiable at x_0". If $f(x)$ is differentiable at each point of an interval then we say $f(x)$ is differentiable throughout the interval. In this case the differential coefficient s is itself a function defined on the interval; its value at x is denoted by $f'(x)$, and this is called the derived function. In contexts where we write $f(x) = y$, we also write $f'(x) = \dfrac{dy}{dx}$.

With the foundations properly laid the usual discussions of the elementary text books can be quite rigorously interpreted. The usual necessary conditions and sufficient conditions for local maxima and minima may be derived. Rolle's theorem and the mean value theorem should be explained geometrically—but it may be noted here that it is quite easy to give rigorous proofs of these theorems using only the properties of continuous functions already explained An important theoretical application of the mean value theorem is to show that if $f'(x) = 0$ in $a \leqslant x \leqslant b$ then $f(x)$ is constant in that interval. The converse, though important, is quite trivial and would have been noted already.

(*d*) *The Definite Integral*

This is, so to speak, the "other half" of the calculus. The theoretical problem is to define what is meant by the area of a region which is bounded wholly or partly by curves. Here it is sufficient to consider this for a region in the Cartesian plane bounded by a continuous curve $y = f(x) > 0$, the x axis and ordinates at $x = a$ and $x = b$, $(a < b)$; and to suppose that $f(x)$ is an increasing function. Divide the interval (a, b) into sub-intervals and form the sums S, s of areas of outer and inner rectangles. It is now easy to show that if all the sub-intervals have length $<\delta$ then

$$S - s < \delta\,[f(b) - f(a)].$$

Also that, if there is an area A satisfying our intuitive requirements then $s < A < S$.

Thus, by taking a suitably fine subdivision we see that either s or S will be a close approximation to A; and the approximation can be made as close as we wish by choice of the subdivision. Thus we may say that the area A is the limit to which either s or S tends as the subdivision is made finer and finer. This will do for the present, but it may be noted that we have not done quite what we intended. From a more sophisticated point of view we should first establish that the limits of S or s in the sense described exist and then this limit is taken as the *definition* of A, the area. We should then show that this definition satisfies the requirements of geometrico-physical intuition. But this more complete discussion belongs to a later stage.

Examples of the direct calculation of areas as limits in cases of the curves $y = x^2$, $y = x^3$ should be given.

The sum S (or s) may be represented symbolically in the form

$$S = \sum_a^b f(x)\,\Delta x$$

and the limit (which is the area) by $\int_a^b f(x)\,dx$, which is called "the definite integral of $f(x)$ over the interval (a, b)".

The extension of the idea of a definite integral to the usual cases when $f(x)$ is not monotonic over the whole interval or when $f(x)$ changes sign in the interval are formal and trivial. The fundamental properties

(i) $$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx;$$

(ii) if $g_1(x) \leqslant f(x) \leqslant g_2(x)$, $a \leqslant x \leqslant b$, then

$$\int_a^b g_1 dx \leqslant \int_a^b f\,dx \leqslant \int_a^b g_2 dx;$$

are also obvious from the geometric picture. The statement (ii) contains the mean value theorem (of the integral calculus).

We now prove easily (still under the assumption that $f(x)$ is continuous)

$$\frac{d}{dx}\int_a^x f(u)\,du = f(x)$$

which establishes the connection between differentiation and integration (and so "joins" the two halves of the subject); this is one form of the fundamental theorem of the calculus. It leads to the calculus rule for evaluating definite integrals. The power of this calculus may now be impressed on students by showing how easily we can calculate the areas under the curves $y = x^2$ and $y = x^3$, comparing this with the direct calculations already made.

Throughout, the work should be illustrated as far as possible with simple examples and the students should work practice exercises. All the points of theoretical statements are often first fully understood only in connection with such exercises.

For example, the simplification of apparently complicated integrals by the use of relations such as

$$\int_0^a f(x)\,dx = \int_0^a f(a-x)\,dx$$

and of the properties of odd and even functions could be demonstrated.

Simple exercises usually serve this purpose just as well as complicated ones, but significant problems are not always simple and it is necessary to acquire useful technique. When possible there should be some intrinsically interesting and challenging problems. The following development of the theory of the elementary transcendental functions should serve to emphasize the power of the calculus methods.

(*e*) *Exponential and Logarithmic Functions*

According to the principles established the area under the curve $y = x^n$ and the ordinates at 1 and x is

$$H(x) = \int_1^x u^n\,du$$

which clearly holds for all $x < 1$ and for all values of n. For $n \neq -1$ we can evaluate this by the method of calculus. For $n = -1$ the method fails; this is simply because we do not know a function whose derived function is x^{-1}.

But, with $n = -1$

$$H(x) = \int_1^x \frac{du}{u}, \tag{1}$$

and, by the general result

$$\frac{dH}{dx} = \frac{1}{x}. \tag{2}$$

Thus the integral formula itself defines a function whose derived function is x^{-1}. We may therefore decide to study the properties of this function directly from this definition.

The *first* properties appear immediately:

$H(1) = 0$, $H(x)$ is an increasing function of x.

Now, using the function of a function rule,

$$\frac{d}{dx} H(ax) = \frac{1}{x}\ ; \text{ from this } H(ax) - H(x) = \text{const.}$$

Setting $x = 1$ we determine the constant and find $H(ax) = H(x) + H(a)$. Replacing x by b,

$$H(a) + H(b) = H(ab) \tag{3}$$

This is a fundamental property of the function.

A quite similar discussion shows that

$$H(x^k) = k\, H(x) \tag{4}$$

From (4), $H(2^x) = x\, H(2)$. This shows that $H(2^x) \to \infty$ as $x \to \infty$; and equivalently $H(x) \to \infty$ as $x \to \infty$. Again from (3), setting $a = x$, $b = \frac{1}{x}$ we find $H(x) = -H\left(\frac{1}{x}\right)$, and so $H(x) \to -\infty$ as $x \to 0$. The function $H(x)$ increases monotonically from $-\infty$ to $+\infty$ as x increases from 0 to ∞. $H(x) \gtrless 0$ as $x \gtrless 1$, and the gradient of $H(x)$ continually decreases as x increases. The graph of $H(x)$ may now be drawn showing these features.

Since $H(x)$ is continuous and increasing there is a unique value e such that $H(e) = 1$. If we write for a moment $x = e^y$ then by (4)

$$H(x) = H(e^y) = y\, H(e) = y = \log_e x$$

by the usual definition of a logarithm. This logarithm, to base e, is called the natural logarithm of x. From this we have its essential properties.

$$\log_e x = \int_1^x \frac{du}{u}, \quad \frac{d}{dx}(\log_e x) = \frac{1}{x}.$$

Further, since $\sqrt{x} < x$, for $x > 1$, we have

$$\log_e x = \int_1^x \frac{du}{u} < \int_1^x \frac{du}{\sqrt{u}} < 2\sqrt{x},$$

and so

$$\frac{\log x}{x} \to 0 \quad \text{as} \quad x \to \infty.$$

The relations $y = \log_e x$, $x = e^y$ are equivalent and so

$$\frac{d}{dy}(e^y) = \frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}} = x = e^y$$

or, what is the same thing, $\frac{d}{dx}(e^x) = e^x$.

Finally we may obtain the series expansions of e^x and $\log(1 + x)$. A simple derivation of the exponential series can be given by using the following principle. Suppose $v_0 = v_0(x)$ be (say) a positive function defined over some range $(0, X)$; and that $0 < v_0(x) < k$.

If we define a sequence $v_1(x)$, $v_2(x)$, . . . in $(0, X)$ by

$$v_{n+1} = \int_0^x v_n \, dx$$

then we find, by induction, $0 < v_n < \dfrac{kx^n}{n!} < \dfrac{kX^n}{n!}$.

Thus $v_n \to 0$ as $n \to \infty$. Now take $v_0 = e^x$.

Then $v_1 = e^x - 1 = v_0 - 1$, or $v_0 - v_1 = 1$.

Then by integration over $(0, x)$ we get successively

$$v_1 - v_2 = x$$

$$v_2 - v_3 = \frac{x^2}{2!}$$

. . . . .

$$v_n - v_{n+1} = \frac{x^n}{n!}.$$

By addition and transposition of terms

$$e^x - \left[1 + x + \ldots + \frac{x^n}{n!}\right] = v_{n+1} \to 0$$

as $n \to \infty$. This establishes the exponential series

$$e^x = 1 + x + \ldots + \frac{x^n}{n!} + \ldots$$

We have supposed $x > 0$; but the proof is easily extended to the case $x < 0$. By setting $x = 1$ we easily compute $e = 2{\cdot}718\,282 \ldots$

From the G.P. formula

$$\frac{1}{1 - x} = 1 + x + \ldots + x^{n-1} + \frac{x^n}{1 - x}$$

we find by integration and then letting $n \to \infty$

$$-\log(1 - x) = x + \frac{x^2}{2} + \ldots + \frac{x^n}{n} \ldots$$

valid in $-1 \leqslant x < 1$.

(*f*) *Arcs of Curves*

Like areas, the length of an arc of a curved line is a matter of definition. Given a curve $y = f(x)$ consider the arc AB between $x = a$ and $x = b$.

Select points $A = P_0, P_1, \ldots, P_n = B$ in order on the curve and form the sum of the lengths of the chords $P_{i-1} P_i$,

$$\sum_{i=1}^{n} P_{i-1} P_i$$

The limit of this sum as the division is made finer and finer, if it exists, is by definition the length of the arc AB. Under certain conditions, viz., $f(x)$ has a continuous derivative, it is easy to show that the limit does exist and is represented by the definite integral.

$$\int_a^b \left[1 + \left\{f'(x)\right\}^2\right]^{\frac{1}{2}} dx = \int_a^b \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{\frac{1}{2}} dx.$$

For, if P, P' be any adjacent points in the above set with co-ordinate differences Δx, Δy then

$$PQ^2 = (\Delta x)^2 + (\Delta y)^2, \quad \text{and} \quad \frac{\Delta y}{\Delta x} = f'(x)$$

where x is a number between the abscissae of P, Q. (This uses the mean value theorem.) So setting

$$\phi = \sqrt{1 + (f')^2}$$

we find

$$PQ = \phi(x) \,.\, \Delta x.$$

Then for the above sum we have

$$\Sigma P_{i-1} P_i = \Sigma\, \phi(x)\, \Delta x.$$

As the division is made finer and finer the sum on the right has the limit

$$\int_a^b \phi\,(x)\, dx$$

by the definition of the integral. This is the result.

For the unit circle $x^2 + y^2 = 1$, we find

$$x \frac{dx}{dy} + y = 0,$$

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

Hence the length of arc θ between $(1, 0)$ and (x, y) in the first quadrant is

$$\theta = \int_0^y \frac{dy}{x}\,.$$

Now, by definition, θ is the angle subtended by this arc at the origin. Also, by definition, the trigonometric functions $\cos\theta$ and $\sin\theta$ are

$$\cos\theta = x, \ \sin\theta = y.$$

Then

$$\frac{dy}{d\theta} = \frac{1}{d\theta/dy} = \frac{1}{1/x} = x,$$

or

$$\frac{d}{d\theta}(\sin\theta) = \cos\theta.$$

So we have the definitions and the differentiation of the trigonometric functions all in a few lines. From these all the other properties are easily derived.

The elementary theory is completed with the definitions of the inverse circular functions, $\sin^{-1}x$, $\cos^{-1}x$, $\tan^{-1}x$; with the selection of principal values and with their differentiation. Further details need not be given here. Perhaps enough has been written to indicate the advantages of first establishing completely the fundamental principles of the calculus and using these for the development of the theory of the elementary functions from appropriate definitions.

(g) *Integration*

Only simple examples should be worked; it is far more important and of greater interest to teach the essential principles of methods and applications. Examples on the use of change of variable should be confined to simple substitutions such as:

$$u = 1 + x^2, \ \sqrt{x}, \ \cos x, \ \log x;$$

$$x = a\sin\theta, \ a\tan\theta.$$

Integration of high powers of $\sin x$ and $\cos x$ is not intended. The use of reduction formulae should be confined to a few very simple examples and certainly no special formulae in this connection need be memorised. The theory of convergence of integrals is not included in this syllabus.

6. LINEAR TRANSFORMATIONS IN THE PLANE. INTRODUCTION TO MATRIX ALGEBRA

(a) TRANSFORMATIONS OF THE EUCLIDEAN PLANE

We consider a plane, and define a geometrical operation in it which associates with each point P of the plane another point P' of the plane; i.e., which transforms (or maps) the plane into itself. Thus we consider operators (functions, mappings, transformations) whose domain is the whole plane. We may omit parentheses in expressing the functional relation between P and P'; i.e., we may write $P' = \mathscr{F}P$ instead of $P' = \mathscr{F}(P)$.

By definition, the resultant $\mathscr{G}\mathscr{F}$, meaning $\mathscr{F}$ followed by $\mathscr{G}$, is the operation which transforms P into P'', where

$$P' = \mathscr{F}P \text{ and } P'' = \mathscr{G}P' \text{ then } P'' = (\mathscr{G}\mathscr{F})P.$$

The identity transformation $\mathscr{I}$ maps each point into itself:

$$P' = \mathscr{I}P = P.$$

A transformation $\mathscr{F}$ is called non-singular if there is a corresponding transformation $\mathscr{F}^{-1}$ such that

$$\mathscr{F}^{-1}\mathscr{F} = \mathscr{I} = \mathscr{F}\mathscr{F}^{-1}.$$

(i) *Translations* ("*Displacements*")

If A, B are any two points they determine a segment $\ulcorner AB \urcorner$ and a transformation $\mathscr{D}_{AB}$ of the plane in which, for any point P,

$$P' = \mathscr{D}_{AB}P$$

is defined by $PP' \parallel AB$ and $PA \parallel P'B$. The identity transformation may be regarded as displacement $\mathscr{D}_{AA}$. Easily proved, but important, results are:

$$\mathscr{D}_{BC}\mathscr{D}_{AB} = \mathscr{D}_{AC} \quad \text{and} \quad \mathscr{D}_{AB}^{-1} = \mathscr{D}_{BA}.$$

(ii) *Rotations* ("*Turns*")

For a given fixed centre of rotation, C, and a given angle α the point $P' = \mathscr{T}_\alpha P$ associated with P is defined by

$$\underline{/*}PCP' = \alpha, \qquad *CP' = *CP.$$

If $\mathscr{T}_\alpha$, $\mathscr{T}_{-\alpha}$, $\mathscr{T}_\beta$, $\mathscr{T}_{\alpha+\beta}$ represent rotations through angles α, $-\alpha$, β, $\alpha+\beta$ respectively, all about the same centre, then

$$\mathscr{T}_\alpha^{-1} = \mathscr{T}_{-\alpha}, \qquad \mathscr{T}_{\alpha+\beta} = \mathscr{T}_\alpha\mathscr{T}_\beta = \mathscr{T}_\beta\mathscr{T}_\alpha,$$

and

$$\mathscr{T}_{\alpha+2\pi} = \mathscr{T}_\alpha.$$

(iii) *Reflections* ("*Symmetries*")

For a given line h the point P' associated with P given by

$$P' = \mathscr{S}_h P$$

is defined by $PP' \perp h$, h bisects PP'.

Clearly $\mathscr{S}_h\mathscr{S}_h = \mathscr{I}$.

If h, k are lines intersecting in C then

$$\mathscr{S}_k\mathscr{S}_h = \mathscr{T}_{2\underline{/*}hk}$$

where the rotation $\mathscr{T}_{2\underline{/*}hk}$ is about centre C and $\underline{/*}hk$ is the measure of the angle from h to k.

If h, k are parallel then $\mathscr{S}_k\mathscr{S}_h$ is a displacement through twice the separation between h and k.

Further remarks

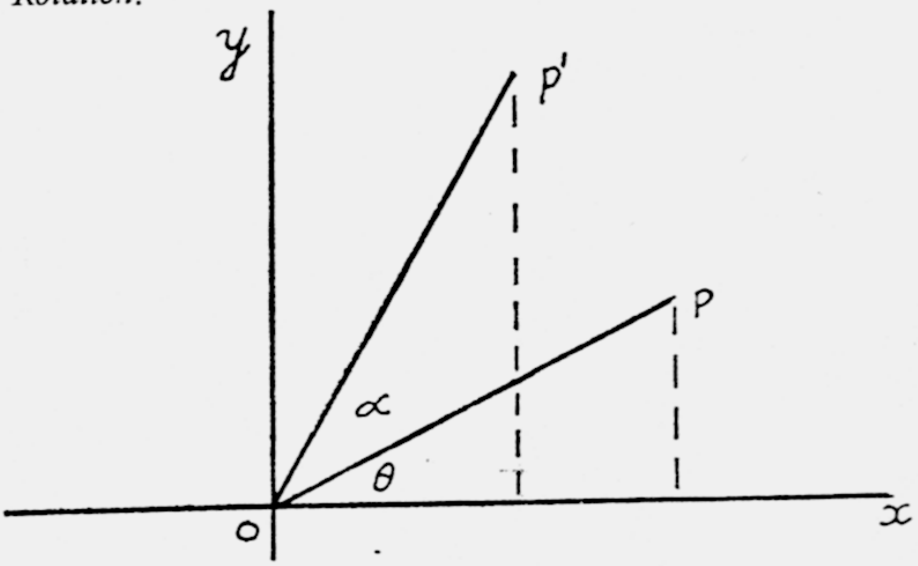
The associative law $\mathscr{F}(\mathscr{G}\mathscr{H}) = (\mathscr{F}\mathscr{G})\mathscr{H}$ holds for the repeated composition of mappings. Another important result is

$$(\mathscr{F}\mathscr{G})^{-1} = \mathscr{G}^{-1}\mathscr{F}^{-1}.$$

(b) CARTESIAN REPRESENTATION OF ROTATIONS, REFLECTIONS

2×2 *Matrices*

(i) *Rotation:*



P is $(r\cos\theta, r\sin\theta)$

P' is $(r\cos(\theta+\alpha), r\sin(\theta+\alpha))$

so that

$x' = r(\cos\theta\cos\alpha - \sin\theta\sin\alpha)$

$y' = r(\cos\theta\sin\alpha + \sin\theta\cos\alpha)$

i.e.,

$$x' = \cos\alpha \,.\, x - \sin\alpha \,.\, y$$
$$y' = \sin\alpha \,.\, x + \cos\alpha \,.\, y \qquad (1)$$

Let us invent an algebraic operator T_α, so that we could write these equations in the form

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2)$$

or

$$\mathbf{r}' = \mathbf{T}_\alpha \mathbf{r}, \qquad (3)$$

where

$$\mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{r}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

are *vectors, column-vectors,* or *column-matrices* and

$$\mathbf{T}_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

is the *matrix* of two *rows* $[\cos\alpha, \ -\sin\alpha]$,

$[\sin\alpha, \ \cos\alpha]$ and two *columns*

$$\begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \quad \begin{bmatrix} -\sin\alpha \\ \cos\alpha \end{bmatrix}$$

The symbolic relation (3) is a condensed form of (2) which in turn corresponds exactly to the two explicit equations (1).

The matrix T_α is an *algebraic operator* transforming the vector **r** into the vector **r**′.

Let us follow the next step in the account of the geometric operators to discover how to combine the algebraic operators, i.e., consider the algebraic equivalent of

$$P'' = \mathscr{T}_\beta P' = \mathscr{T}_\beta \mathscr{T}_\alpha P.$$

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta\sin\alpha - \sin\beta\sin\alpha & \cos\beta(-\sin\alpha) - \sin\beta\cos\alpha \\ \sin\beta\cos\alpha + \cos\beta\sin\alpha & \sin\beta(-\sin\alpha) + \cos\beta\cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

This is exactly $T_\beta T_\alpha \mathbf{r} = T_{\alpha+\beta}\mathbf{r}$ and gives the rules which have to be followed if the "matrices" T_α are to correspond exactly to the geometric operators $\mathscr{T}_\alpha$. This is the line of thought which led Cayley to the invention of *matrices* to represent the operators and *matrix multiplication*, to represent the compounding of two operators.

$$T_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{1}, \text{ the } \textit{unit matrix}.$$

$$T_{\frac{1}{2}\pi} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad T_\pi = -\mathbf{1}.$$

The rule for multiplication shows up most clearly if we use double subscripts. Take in a matrix A, a_{ij} to be the *element* in the *row* numbered i and *column* numbered j, so that A can be written as

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then

$$\begin{matrix} & \text{column 1} & \text{column 2} & \\ AB = & \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} & & \begin{matrix}\text{row 1} \\ \text{row 2}\end{matrix} \end{matrix}$$

or, if $AB = C$, then

$$c_{rs} = a_{r1}b_{1s} + a_{r2}b_{2s}.$$

Clearly in general $\mathbf{AB} \neq \mathbf{BA}$, although

$T_\beta T_\alpha = T_\alpha T_\beta$ (but $\mathbf{S}_b\mathbf{S}_a \neq \mathbf{S}_a\mathbf{S}_b$ for these matrices as used in the next section (ii)).

Matrix multiplication is *associative*.

(ii) *Reflections*

If $y = x \tan\theta$ is the line h, write $\mathscr{S}_\theta$ for $\mathscr{S}_h$, and $\mathbf{S}_\theta$ for the matrix corresponding to $\mathscr{S}_\theta$. Then

$$\mathbf{S}_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{S}_{\frac{1}{4}\pi} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{S}_{\frac{1}{2}\pi} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

$\mathbf{S}_\theta$ can be computed directly from the Cartesian diagram, but it is simpler to use the relation

$$\mathscr{S}_\theta \mathscr{S}_0 = \mathscr{T}_{2\theta},$$

$$\mathscr{S}_\theta = \mathscr{T}_{2\theta} \mathscr{S}_0,$$

so that

$$\mathbf{S}_\theta = \mathbf{T}_{2\theta}\mathbf{S}_0 = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

If we use the perpendicular form of the equation of the line

$$x \cos\phi + y \sin\phi = 0, \text{ where } \theta = \tfrac{1}{2}\pi + \phi,$$

we have

$$\mathbf{S}_{\frac{1}{2}\pi+\phi} = \begin{bmatrix} -\cos 2\phi & -\sin 2\phi \\ -\sin 2\phi & \cos 2\phi \end{bmatrix}$$

(iii) *The Affine Transformation*

This transformation we define algebraically as a generalization of **S** and **T**, and then investigate its geometric properties. The discussion depends on the section formulae

$$x_3 = \frac{k_1 x_1 + k_2 x_2}{k_1 + k_2}, \quad y_3 = \frac{k_1 y_1 + k_2 y_2}{k_1 + k_2}$$

which we can write as

$$\begin{bmatrix} (k_1 + k_2)\, x_3 \\ (k_1 + k_2)\, y_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

or

$$(k_1 + k_2)\,\mathbf{r}_3 = [\mathbf{r}_1 \quad \mathbf{r}_2]\,\mathbf{k}.$$

The transformation is

$$\mathbf{r}' = \mathbf{M}\mathbf{r}$$

where

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$$

When this is applied to the points of a line

$$(k_1 + k_2)\,\mathbf{r} = [\mathbf{r}_1 \quad \mathbf{r}_2]\,\mathbf{k},$$

(k_1, k_2 vary along the line) we find

$$\mathbf{M}\,\{(k_1 + k_2)\,\mathbf{r}\} = \mathbf{M}\,[\mathbf{r}_1 \quad \mathbf{r}_2]\,\mathbf{k},$$

which we can rewrite as

$$(k_1 + k_2)\,\mathbf{Mr} = [\mathbf{Mr}_1 \quad \mathbf{Mr}_2]\,\mathbf{k}$$

i.e., the transformed points are the points of the line $\mathbf{Mr}_1$, $\mathbf{Mr}_2$.

Write $\mathcal{M}$ for the corresponding geometric transformation. Then, if $\mathcal{M}\,P_i = P'_i$ and $P_3 \in P_1P_2$ we have

$$P'_3 \in P_1'\,P_2' \text{ and}$$

$$*P_1P_3/*P_1P_2 = *P_1'P_3'/*P_1'P_2'.$$

Thus, under the transformation $\mathcal{M}$,

points become points

lines become lines

ratios of displacements on a line become equal ratios on the transformed line so that

parallels become parallels

But

distances are altered

and

angles are altered.

(c) SOME PROPERTIES OF MATRICES UNDER MULTIPLICATION

(i) *The inverse matrix*

$\mathbf{M}^{-1}$ is defined by $\mathbf{M}^{-1}\mathbf{M} = \mathbf{1}$.

If

$$\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } \mathbf{M}^{-1} = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$$

we have

$$\begin{bmatrix} x & y \\ z & t \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\left.\begin{aligned} ax + cy &= 1 \\ bx + dy &= 0 \end{aligned}\right\} \quad \left.\begin{aligned} az + ct &= 0 \\ bz + dt &= 1 \end{aligned}\right\}$$

So that

$$\mathbf{M}^{-1} = \begin{bmatrix} \dfrac{d}{ad - bc} & \dfrac{-b}{ad - bc} \\ \\ \dfrac{-c}{ad - bc} & \dfrac{a}{ad - bc} \end{bmatrix}$$

$$\mathbf{MM}^{-1} = \mathbf{M}^{-1}\mathbf{M} = \mathbf{1}.$$

Thus $\mathbf{M}^{-1}$ exists if and only if $ad - bc \neq 0$.

If $\mathbf{A}$ and $\mathbf{B}$ are any two non-singular matrices

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}.$$

(ii) *The determinant of a matrix*

Notation:

$$ad - bc = \det \mathbf{M} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

If $\det \mathbf{M} = 0$, the matrix is *singular* (i.e. it has no inverse).

If the matrix $\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is singular,

the affine transformation

$$\mathbf{M} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} ax_0 + by_0 \\ cx_0 + dy_0 \end{bmatrix} = \begin{bmatrix} ax_0 + by_0 \\ (ax_0 + by_0)\ c/a \end{bmatrix}$$

(provided $a \neq 0$), is such that for all points r the transformed $\mathbf{Mr}$ lies on the line $cx - ay = 0$, i.e., the *transformation is singular*, the whole plane being *mapped onto* the line.

Note that $\det \mathbf{T} = 1$, $\det \mathbf{S} = -1$, so that these matrices are always non-singular.

By direct multiplication we find for any two matrices $\mathbf{M}$, $\mathbf{M}'$

$$\det (\mathbf{MM}') = \det \mathbf{M} \det \mathbf{M}' = \det (\mathbf{M}'\mathbf{M})$$

$$\det (\mathbf{M}^{-1}) = (\det \mathbf{M})^{-1}.$$

(iii) *The zero vector and zero matrix*

$$\mathbf{o} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

For all $\mathbf{M}$, $\mathbf{Mo} = \mathbf{o}$ and $\mathbf{M0} = \mathbf{0M} = \mathbf{0}$.

The most general singular matrix may be written as

$$\begin{bmatrix} rr' & rs' \\ r's & ss' \end{bmatrix}$$

For this matrix we have

$$\begin{bmatrix} rr' & rs' \\ r's & ss' \end{bmatrix} \begin{bmatrix} us' \\ -ur' \end{bmatrix} = \mathbf{o} \text{ for any } u$$

and

$$\begin{bmatrix} rr' & rs' \\ r's & ss' \end{bmatrix} \begin{bmatrix} us & vs' \\ -ur' & -vr' \end{bmatrix} = \mathbf{0} \text{ for any } u, v$$

and

$$\begin{bmatrix} hs & -hr \\ ks & -kr \end{bmatrix} \begin{bmatrix} rr' & rs' \\ r's & ss' \end{bmatrix} = \mathbf{0} \text{ for any } h, k.$$

Thus in matrix algebra there are *divisors of the zero matrix*, i.e., there exist pairs of matrices $\mathbf{M}$ and $\mathbf{N}$, both non-zero such that $\mathbf{MN} = \mathbf{0}$. If we are given $\mathbf{AB} = \mathbf{0}$ we cannot deduce that either $\mathbf{A}$ or $\mathbf{B}$ is the zero matrix, but only that, if neither $\mathbf{A}$ nor $\mathbf{B}$ is the zero matrix, both are singular.

(iv) *The transposed vector and matrix*

If $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ we write $\mathbf{a}^T = [a_1, \ a_2]$ for the transpose of **a**, i.e., the *row-vector* with components identical with those of the column vector **a**. Likewise if

$$\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{we define } \mathbf{M}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

If **A** is any non-singular matrix

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

If **A** and **B** are any two matrices, and **k** any vector

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

$$(\mathbf{Ak})^T = \mathbf{k}^T \mathbf{A}^T.$$

In particular $\mathbf{S}^T = \mathbf{S}^{-1} = \mathbf{S}$, $\mathbf{T}^T = \mathbf{T}^{-1}$.

(v) *Uniqueness of* **S** *and* **T**

Theorem

If under an affine transformation, with fixed origin, distance is invariant, then the transformation is either a rotation or a reflection.

We require

$$x'^2 + y'^2 = x^2 + y^2$$

$$\text{i.e. } \mathbf{r}'^T\mathbf{r}' = \mathbf{r}^T\mathbf{r}$$

$$\text{i.e. } (\mathbf{Mr})^T(\mathbf{Mr}) = \mathbf{r}^T\mathbf{r}$$

$$\text{i.e. } \mathbf{r}^T(\mathbf{M}^T\mathbf{M})\mathbf{r} = \mathbf{r}^T\mathbf{1}\mathbf{r}$$

$$\mathbf{M}^T\mathbf{M} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix}$$

$$\mathbf{r}^T \mathbf{M}^T\mathbf{M}\mathbf{r} = (a^2 + c^2)x^2 + 2\,(ab + cd)xy + (b^2 + d^2)y^2$$

$$\mathbf{r}^T \mathbf{r} = x^2 + y^2$$

Thus, if distance is invariant,

$$\left.\begin{aligned} a^2 + c^2 &= 1 \\ b^2 + d^2 &= 1 \end{aligned}\right\} \quad ab + cd = 0$$

Take

$$a = \cos \alpha, \quad c = \sin \alpha,$$

$$b = \cos \beta, \quad d = \sin \beta,$$

then $\cos(\alpha - \beta) = 0$;

i.e., either $\beta = \alpha - \frac{1}{2}\pi$ or $\beta = \alpha + \frac{1}{2}\pi$

and the two possible matrices are

$$\begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix} = \mathbf{S}_{\frac{1}{2}\alpha}$$

or

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} = \mathbf{T}_{\alpha}$$

A matrix with the property $\mathbf{M}^T\mathbf{M} = \mathbf{1}$, i.e., $\mathbf{M}^T = \mathbf{M}^{-1}$, is called orthogonal.

(*d*) DISPLACEMENT, MATRIX ADDITION, AND MATRIX ALGEBRA

If H is (h, k) and $\mathscr{D}_{OH}P = P'$

where P is (x, y) and P' is (x', y'), we have

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x + h \\ y + k \end{bmatrix}$$

which we write as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} h \\ k \end{bmatrix}$$

or as

$$\mathbf{r}' = \mathbf{r} + \mathbf{h}.$$

This suggests an addition operation for *vectors*, and then immediately an addition operation for *matrices*.

If

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

We define $\mathbf{A} + \mathbf{B}$ by:

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

For any number k, define

$$k\mathbf{A} = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix} = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and hence $\det (k\mathbf{A}) = k^2 \det \mathbf{A}$.

Definition of $\mathbf{A} - \mathbf{B}$, and $\mathbf{A} - \mathbf{A} = \mathbf{0}$ (zero matrix).

Combination of sums and products.

The rules of matrix algebra are the same as those for a field except that:

1. Multiplication is not commutative.
2. There are divisors of zero, i.e., not every element has a multiplicative inverse.

(*e*) CHARACTERISTIC POLYNOMIAL, EIGENVALUES, EIGENVECTORS

The characteristic polynomial of

$$\mathbf{A} = \begin{bmatrix} a & h \\ h & b \end{bmatrix}$$

is

$$\det [x\mathbf{1} - \mathbf{A}] = x^2 - (a + b)x + \det \mathbf{A}.$$

The roots of this characteristic polynomial are the eigenvalues of $\mathbf{A}$; for each eigenvalue λ the non-zero vectors $\mathbf{u}$ satisfying $(\lambda\mathbf{1} - \mathbf{A})\mathbf{u} = \mathbf{o}$ are called *eigenvectors* belonging to λ.

(*f*) CHANGE OF CO-ORDINATES

(i) *Parallel shift of axes*

The equation $\mathbf{r}' = \mathbf{r} + \mathbf{h}$ may be regarded *either* as a transformation of the plane in which $\mathbf{r}$ is transformed into $\mathbf{r}'$ *or* as a change of co-ordinates with new origin at $\mathbf{r} = -\mathbf{h}$ and axes parallely displaced.

(ii) *Rotation of axes*

The equation $\mathbf{r}' = \mathbf{T}_\alpha \mathbf{r}$ may be interpreted *either* as a transformation of the plane *or* as a change of axes by rotation through an angle of $-\alpha$.

(*g*) REDUCTION OF QUADRATIC FORMS. CONICS

It should be shown that the quadratic form

$$ax^2 + 2hxy + by^2$$

can, by a suitable rotation of the co-ordinate axes, be written as

$$\alpha x'^2 + \beta y'^2$$

where α, β are the eigenvalues of the matrix

$$\mathbf{A} = \begin{bmatrix} a & h \\ h & b \end{bmatrix}.$$

The rotation of axes may be represented by the transformation

$$\mathbf{r} = \mathbf{P}\mathbf{r}'$$

where

$$\mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{r}' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

and

$$(a - b)\sin 2\theta = 2\,h\cos 2\theta.$$

The above is essentially equivalent to the reduction of the matrix **A** to diagonal form in accordance with the equation

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}.$$

The locus

$$ax^2 + 2hxy + by^2 = 1$$

assumes, in the (x', y') frame, the simpler form

$$\alpha x'^2 + \beta y'^2 = 1$$

which is more readily identified. Depending on the values of α, β it may be a central conic (ellipse or hyperbola), or a pair of parallel lines, or it may represent the empty set. It should also be noted that the principal axes of the conic will lie along eigenvectors of **A**.

D. West, Government Printer, New South Wales—1974