CHAPTER VI

COMPLEX NUMBERS AND VECTORS

70. Operators and operations. If an entity u is changed into an entity v by some law, the change may be regarded as an *operation* performed upon u, the *operand*, to convert it into v; and if f be introduced as the symbol of the operation, the result may be written as $v = fu$. For brevity the symbol f is often called an *operator*. Various sorts of operand, operator, and result are familiar. Thus if u is a positive number n, the application of the operator $\sqrt{}$ gives the square root; if u represents a range of values of a variable x, the expression $f(x)$ or fx denotes a function of x; if u be a function of x, the operation of differentiation may be symbolized by D and the result Du is the derivative; the symbol of definite integration $\int_a^b (*)\,d*$ converts a function $u(x)$ into a number; and so on in great variety.

The reason for making a short study of operators is that a considerable number of the concepts and rules of arithmetic and algebra may be so defined for operators themselves as to lead to a *calculus of operations* which is of frequent use in mathematics; the single application to the integration of certain differential equations (§ 95) is in itself highly valuable. The fundamental concept is that of a *product: If u is operated upon by f to give $fu = v$ and if v is operated upon by g to give $gv = w$, so that*

$$fu = v, \qquad gv = gfu = w, \qquad gfu = w, \tag{1}$$

then the operation indicated as gf which converts u directly into w is called the product of f by g. If the functional symbols sin and log be regarded as operators, the symbol log sin could be regarded as the product. The transformations of turning the xy-plane over on the x-axis, so that $x' = x$, $y' = -y$, and over the y-axis, so that $x' = -x$, $y' = y$, may be regarded as operations; the combination of these operations gives the transformation $x' = -x$, $y' = -y$, which is equivalent to rotating the plane through 180° about the origin.

The products of arithmetic and algebra satisfy the *commutative law* $gf = fg$, that is, the products of g by f and of f by g are equal. This is not true of operators in general, as may be seen from the fact that log sin x and sin log x are different. Whenever the order of the factors is immaterial, as in the case of the transformations just considered, the operators are said to be *commutative*. Another law of arithmetic and algebra is that when there are three or more factors in a product, the factors may be grouped at pleasure without altering the result, that is,

$$h(gf) = (hg)f = hgf. \tag{2}$$

This is known as the *associative law* and operators which obey it are called *associative*. Only associative operators are considered in the work here given.

For the repetition of an operator several times

$$ff = f^2, \quad fff = f^3, \quad f^m f^n = f^{m+n}, \tag{3}$$

the usual notation of powers is used. *The law of indices clearly holds;* for f^{m+n} means that f is applied $m+n$ times successively, whereas $f^m f^n$ means that it is applied n times and then m times more. Not applying the operator f at all would naturally be denoted by f^0, so that $f^0 u = u$ and the operator f^0 would be equivalent to multiplication by 1; the notation $f^0 = 1$ is adopted.

If for a given operation f there can be found an operation g such that the product $fg = f^0 = 1$ is equivalent to no operation, then g is called the *inverse* of f and notations such as

$$fg = 1, \quad g = f^{-1} = \frac{1}{f}, \quad ff^{-1} = f\frac{1}{f} = 1 \tag{4}$$

are regularly borrowed from arithmetic and algebra. Thus the inverse of the square is the square root, the inverse of sin is $\sin^{-1}$, the inverse of the logarithm is the exponential, the inverse of D is $\int$. Some operations have no inverse; multiplication by 0 is a case, and so is the square when applied to a negative number if only real numbers are considered. Other operations have more than one inverse; integration, the inverse of D, involves an arbitrary additive constant, and the inverse sine is a multiple valued function. It is therefore not always true that $f^{-1}f = 1$, but it is customary to mean by f^{-1} that particular inverse of f for which $f^{-1}f = ff^{-1} = 1$. Higher negative powers are defined by the equation $f^{-n} = (f^{-1})^n$, and it readily follows that $f^n f^{-n} = 1$, as may be seen by the example

$$f^3 f^{-3} = ff(f \cdot f^{-1})f^{-1}f^{-1} = f(f \cdot f^{-1})f^{-1} = ff^{-1} = 1.$$

The law of indices $f^m f^n = f^{m+n}$ *also holds for negative indices,* except in so far as $f^{-1}f$ may not be equal to 1 and may be required in the reduction of $f^m f^n$ to f^{m+n}.

If u, v, and $u + v$ are operands for the operator f and if

$$f(u + v) = fu + fv, \tag{5}$$

so that the operator applied to the sum gives the same result as the sum of the results of operating on each operand, then the operator f is called *linear* or *distributive*. If f denotes a function such that $f(x + y) = f(x) + f(y)$, it has been seen (Ex. 9, p. 45) that f must be equivalent to multiplication by a constant and $fx = Cx$. For a less specialized interpretation this is not so; for

$$D(u + v) = Du + Dv \quad \text{and} \quad \int (u + v) = \int u + \int v$$

are two of the fundamental formulas of calculus and show operators which are distributive and not equivalent to multiplication by a constant. Nevertheless it does follow by the same reasoning as used before (Ex. 9, p. 45), that $fnu = nfu$ if f is distributive and if n is a rational number.

Some operators have also the property of addition. Suppose that u is an operand and f, g are operators such that fu and gu are things that may be added together as $fu + gu$, then the *sum* of the operators, $f + g$, is defined by the equation $(f + g)u = fu + gu$. If furthermore the operators f, g, h are distributive, then

$$h(f + g) = hf + hg \quad \text{and} \quad (f + g)h = fh + gh, \tag{6}$$

and the multiplication of the operators becomes itself distributive. To prove this fact, it is merely necessary to consider that

$$h[(f + g)u] = h(fu + gu) = hfu + hgu$$

and

$$(f + g)(hu) = fhu + ghu.$$

Operators which are associative, commutative, distributive, and which admit addition may be treated algebraically, in so far as polynomials are concerned, by the ordinary algorisms of algebra; for it is by means of the associative, commutative, and distributive laws, and the law of indices that ordinary algebraic polynomials are rearranged, multiplied out, and factored. Now the operations of multiplication by constants and of differentiation or partial differentiation as applied to a function of one or more variables x, y, z, $\cdots$ do satisfy these laws. For instance

$$c(Du) = D(cu), \quad D_xD_yu = D_yD_xu, \quad (D_x + D_y)D_zu = D_xD_zu + D_yD_zu. \tag{7}$$

Hence, for example, if y be a function of x, the expression

$$D^ny + a_1D^{n-1}y + \cdots + a_{n-1}Dy + a_ny,$$

where the coefficients a are constants, may be written as

$$(D^n + a_1D^{n-1} + \cdots + a_{n-1}D + a_n)y \tag{8}$$

and may then be factored into the form

$$[(D-\alpha_1)(D-\alpha_2)\cdots(D-\alpha_{n-1})(D-\alpha_n)]y, \tag{8'}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the roots of the algebraic polynomial

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = 0.$$

EXERCISES

1. Show that $(fgh)^{-1} = h^{-1}g^{-1}f^{-1}$, that is, that the reciprocal of a product of operations is the product of the reciprocals in inverse order.

2. By definition the operator gfg^{-1} is called the transform of f by g. Show that (α) the transform of a product is the product of the transforms of the factors taken in the same order, and (β) the transform of the inverse is the inverse of the transform.

3. If $s \neq 1$ but $s^2 = 1$, the operator s is by definition said to be *involutory*. Show that (α) an involutory operator is equal to its own inverse; and conversely (β) if an operator and its inverse are equal, the operator is involutory; and (γ) if the product of two involutory operators is commutative, the product is itself involutory; and conversely (δ) if the product of two involutory operators is involutory, the operators are commutative.

4. If f and g are both distributive, so are the products fg and gf.

5. If f is distributive and n rational, show $fnu = nfu$.

6. Expand the following operators first by ordinary formal multiplication and second by applying the operators successively as indicated, and show the results are identical by translating both into familiar forms.

$$(\alpha)\ (D-1)(D-2)y, \quad Ans.\ \frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y,$$

$$(\beta)\ (D-1)D(D+1)y, \qquad (\gamma)\ D(D-2)(D+1)(D+3)y.$$

7. Show that $(D-a)\left[e^{ax}\int e^{-ax}Xdx\right] = X$, where X is a function of x, and hence infer that $e^{ax}\int e^{-ax}(*)\,dx$ is the inverse of the operator $(D-a)(*)$.

8. Show that $D(e^{ax}y) = e^{ax}(D+a)y$ and hence generalize to show that if $P(D)$ denote any polynomial in D with constant coefficients, then

$$P(D)\cdot e^{ax}y = e^{ax}P(D+a)y.$$

Apply this to the following and check the results.

$$(\alpha)\ (D^2 - 3D + 2)e^{2x}y = e^{2x}(D^2 + D)y = e^{2x}\left(\frac{d^2y}{dx^2} + \frac{dy}{dx}\right),$$

$$(\beta)\ (D^2 - 3D - 2)e^{x}y, \qquad (\gamma)\ (D^3 - 3D + 2)e^{x}y.$$

9. If y is a function of x and $x = e^t$ show that

$$D_xy = e^{-t}D_ty,\ D_x^2y = e^{-2t}D_t(D_t - 1)y,\ \cdots,\ D_x^py = e^{-pt}D_t(D_t-1)\cdots(D_t - p + 1)y.$$

10. Is the expression $(hD_x + kD_y)^n$, which occurs in Taylor's Formula (§ 54), the nth power of the operator $hD_x + kD_y$ or is it merely a conventional symbol? The same question relative to $(xD_x + yD_y)^k$ occurring in Euler's Formula (§ 53)?

71. Complex numbers. In the formal solution of the equation $ax^2 + bx + c = 0$, where $b^2 < 4ac$, numbers of the form $m + n\sqrt{-1}$, where m and n are real, arise. Such numbers are called *complex* or *imaginary*; the part m is called the *real part* and $n\sqrt{-1}$ the *pure imaginary part* of the number. It is customary to write $\sqrt{-1} = i$ and to treat i as a literal quantity subject to the relation $i^2 = -1$. The definitions for the *equality*, *addition*, and *multiplication* of complex numbers are

$$\begin{aligned} a + bi &= c + di \quad \text{if and only if} \quad a = c,\ b = d, \\ [a + bi] + [c + di] &= (a + c) + (b + d)\, i, \\ [a + bi]\,[c + di] &= (ac - bd) + (ad + bc)\, i. \end{aligned} \tag{9}$$

It readily follows that *the commutative, associative, and distributive laws hold in the domain of complex numbers*, namely,

$$\begin{aligned} \alpha + \beta &= \beta + \alpha, & (\alpha + \beta) + \gamma &= \alpha + (\beta + \gamma), \\ \alpha\beta &= \beta\alpha, & (\alpha\beta)\,\gamma &= \alpha\,(\beta\gamma), \\ \alpha\,(\beta + \gamma) &= \alpha\beta + \alpha\gamma, & (\alpha + \beta)\,\gamma &= \alpha\gamma + \beta\gamma, \end{aligned} \tag{10}$$

where Greek letters have been used to denote complex numbers.

Division is accomplished by the method of rationalization.

$$\frac{a + bi}{c + di} = \frac{a + bi}{c + di}\,\frac{c - di}{c - di} = \frac{(ac + bd) + (bc - ad)\, i}{c^2 + d^2}. \tag{11}$$

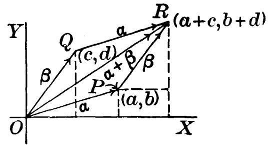
This is always possible except when $c^2 + d^2 = 0$, that is, when both c and d are 0. A complex number is defined as 0 when and only when its real and pure imaginary parts are both zero. With this definition 0 has the ordinary properties that $\alpha + 0 = \alpha$ and $\alpha 0 = 0$ and that $\alpha/0$ is impossible. Furthermore *if a product $\alpha\beta$ vanishes, either α or β vanishes.* For suppose

$$[a + bi]\,[c + di] = (ac - bd) + (ad + bc)\, i = 0.$$

Then
$$ac - bd = 0 \quad \text{and} \quad ad + bc = 0, \tag{12}$$

from which it follows that either $a = b = 0$ or $c = d = 0$. From the fact that a product cannot vanish unless one of its factors vanishes follow the ordinary laws of cancellation. In brief, *all the elementary laws of real algebra hold also for the algebra of complex numbers.*

By assuming a set of Cartesian coördinates in the xy-plane and associating the number $a + bi$ to the point (a, b), a *graphical representation* is obtained which is the counterpart of the number scale for real numbers. The point (a, b) alone or the directed line from the origin to the point (a, b) may be considered as representing the number $a + bi$. If OP and OQ are two directed lines representing the two numbers $a + bi$ and $c + di$, a reference to the figure shows that the line which represents the sum of the numbers is OR, the diagonal of the parallelogram of which OP and OQ are sides. Thus *the geometric law for adding complex numbers is the same as the law for compounding forces and is known as the parallelogram law.* A segment AB of a line possesses magnitude, the length AB, and direction, the direction of the line AB from A to B. *A quantity which has magnitude and direction is called a vector; and the parallelogram law is called the law of vector addition. Complex numbers* may therefore be regarded as *vectors.*



From the figure it also appears that OQ and PR have the same magnitude and direction, so that as vectors they are equal although they start from different points. As $OP + PR$ will be regarded as equal to $OP + OQ$, the definition of addition may be given as the triangle law instead of as the parallelogram law; namely, from the terminal end P of the first vector lay off the second vector PR and close the triangle by joining the initial end O of the first vector to the terminal end R of the second. The *absolute value* of a complex number $a + bi$ is the magnitude of its vector OP and is equal to $\sqrt{a^2 + b^2}$, the square root of the sum of the squares of its real part and of the coefficient of its pure imaginary part. The absolute value is denoted by $|a + bi|$ as in the case of reals. If α and β are two complex numbers, the rule $|\alpha| + |\beta| \geqq |\alpha + \beta|$ is a consequence of the fact that one side of a triangle is less than the sum of the other two. If the absolute value is given and the initial end of the vector is fixed, the terminal end is thereby constrained to lie upon a circle concentric with the initial end.

72. When the complex numbers are laid off from the origin, polar coördinates may be used in place of Cartesian. Then

$$r = \sqrt{a^2 + b^2}, \qquad \phi = \tan^{-1} b/a^*, \qquad a = r\cos\phi, \qquad b = r\sin\phi$$
$$\text{and} \qquad a + ib = r(\cos\phi + i\sin\phi). \tag{13}$$

The absolute value r is often called the *modulus* or *magnitude* of the complex number; the angle ϕ is called the *angle* or *argument* of the number and suffers a certain indetermination in that $2n\pi$, where n is a positive or negative integer, may be added to ϕ without affecting the number. This polar representation is particularly useful in discussing products and quotients. For if

$$\alpha = r_1(\cos\phi_1 + i\sin\phi_1), \qquad \beta = r_2(\cos\phi_2 + i\sin\phi_2),$$
$$\text{then} \qquad \alpha\beta = r_1 r_2[\cos(\phi_1 + \phi_2) + i\sin(\phi_1 + \phi_2)], \tag{14}$$

* As both $\cos\phi$ and $\sin\phi$ are known, the quadrant of this angle is determined.

as may be seen by multiplication according to the rule. Hence the *magnitude of a product is the product of the magnitudes of the factors, and the angle of a product is the sum of the angles of the factors;* the general rule being proved by induction.

The interpretation of *multiplication by a complex number as an operation* is illuminating. Let β be the multiplicand and α the multiplier. As the product $\alpha\beta$ has a magnitude equal to the product of the magnitudes and an angle equal to the sum of the angles, the factor α used as a multiplier may be interpreted as effecting the rotation of β through the angle of α and the stretching of β in the ratio $|\alpha|:1$. From the geometric viewpoint, therefore, *multiplication by a complex number is an operation of rotation and stretching in the plane.* In the case of $\alpha = \cos\phi + i\sin\phi$ with $r = 1$, the operation is only of rotation and hence the factor $\cos\phi + i\sin\phi$ is often called a cyclic factor or versor. In particular the number $i = \sqrt{-1}$ will effect a rotation through 90° when used as a multiplier and is known as a quadrantal versor. The series of powers i, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$ give rotations through 90°, 180°, 270°, 360°. This fact is often given as the reason for laying off pure imaginary numbers bi along an axis at right angles to the axis of reals.

As a particular product, the nth power of a complex number is

$$\alpha^n = (a + ib)^n = [r(\cos\phi + i\sin\phi)]^n = r^n(\cos n\phi + i\sin n\phi); \quad (15)$$

and

$$(\cos\phi + i\sin\phi)^n = \cos n\phi + i\sin n\phi, \quad (15')$$

which is a special case, is known as *De Moivre's Theorem* and is of use in evaluating the functions of $n\phi$; for the binomial theorem may be applied and the real and imaginary parts of the expansion may be equated to $\cos n\phi$ and $\sin n\phi$. Hence

$$\begin{aligned}\cos n\phi &= \cos^n\phi - \frac{n(n-1)}{2!}\cos^{n-2}\phi\sin^2\phi \\ &\quad + \frac{n(n-1)(n-2)(n-3)}{4!}\cos^{n-4}\phi\sin^4\phi - \cdots \qquad (16)\\ \sin n\phi &= n\cos^{n-1}\phi\sin\phi - \frac{n(n-1)(n-2)}{3!}\cos^{n-3}\phi\sin^3\phi + \cdots.\end{aligned}$$

As the nth root $\sqrt[n]{\alpha}$ of α must be a number which when raised to the nth power gives α, the nth root may be written as

$$\sqrt[n]{\alpha} = \sqrt[n]{r}(\cos\phi/n + i\sin\phi/n). \quad (17)$$

The angle ϕ, however, may have any of the set of values

$$\phi, \quad \phi + 2\pi, \quad \phi + 4\pi, \quad \cdots, \quad \phi + 2(n-1)\pi,$$

and the nth parts of these give the n different angles

$$\frac{\phi}{n}, \quad \frac{\phi}{n}+\frac{2\pi}{n}, \quad \frac{\phi}{n}+\frac{4\pi}{n}, \quad \cdots, \quad \frac{\phi}{n}+\frac{2(n-1)\pi}{n}. \tag{18}$$

Hence there may be found just n different nth roots of any given complex number (including, of course, the reals).

The *roots of unity* deserve mention. The equation $x^n = 1$ has in the real domain one or two roots according as n is odd or even. But if 1 be regarded as a complex number of which the pure imaginary part is zero, it may be represented by a point at a unit distance from the origin upon the axis of reals; the magnitude of 1 is 1 and the angle of 1 is $0, 2\pi, \cdots, 2(n-1)\pi$. The nth roots of 1 will therefore have the magnitude 1 and one of the angles $0, 2\pi/n, \cdots, 2(n-1)\pi/n$. The n nth roots are therefore

$$1, \quad \alpha = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}, \quad \alpha^2 = \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n}, \quad \cdots,$$
$$\alpha^{n-1} = \cos\frac{2(n-1)\pi}{n} + i\sin\frac{2(n-1)\pi}{n},$$

and may be evaluated with a table of natural functions. Now $x^n - 1 = 0$ is factorable as $(x-1)(x^{n-1}+x^{n-2}+\cdots+x+1)=0$, and it therefore follows that the nth roots other than 1 must all satisfy the equation formed by setting the second factor equal to 0. As α in particular satisfies this equation and the other roots are $\alpha^2, \cdots, \alpha^{n-1}$, it follows that the sum of the n nth roots of unity is zero.

EXERCISES

1. Prove the distributive law of multiplication for complex numbers.

2. By definition the pair of imaginaries $a + bi$ and $a - bi$ are called *conjugate imaginaries*. Prove that (α) the sum and the product of two conjugate imaginaries are real; and conversely (β) if the sum and the product of two imaginaries are both real, the imaginaries are conjugate.

3. Show that if $P(x, y)$ is a symmetric polynomial in x and y with real coefficients so that $P(x, y) = P(y, x)$, then if conjugate imaginaries be substituted for x and y, the value of the polynomial will be real.

4. Show that if $a + bi$ is a root of an algebraic equation $P(x) = 0$ with real coefficients, then $a - bi$ is also a root of the equation.

5. Carry out the indicated operations algebraically and make a graphical representation for every number concerned and for the answer:

(α) $(1+i)^3$, (β) $(1+\sqrt{3}\,i)(1-i)$, (γ) $(3+\sqrt{-2})(4+\sqrt{-5})$,

(δ) $\dfrac{1+i}{1-i}$, (ϵ) $\dfrac{1+i\sqrt{3}}{1-i\sqrt{3}}$, ($\zeta$) $\dfrac{5}{\sqrt{2}-i\sqrt{3}}$,

(η) $\dfrac{(1-i)^2}{(1+i)^3}$, (θ) $\dfrac{1}{(1+i)^2}+\dfrac{1}{(1-i)^2}$, ($\iota$) $\left(\dfrac{-1+\sqrt{-3}}{2}\right)^3$.

6. Plot and find the modulus and angle in the following cases:

(α) -2, (β) $-2\sqrt{-1}$, (γ) $3+4i$, (δ) $\frac{1}{2}-\frac{1}{2}\sqrt{-3}$.

7. Show that *the modulus of a quotient of two numbers is the quotient of the moduli and that the angle is the angle of the numerator less that of the denominator.*

8. Carry out the indicated operations trigonometrically and plot:

(α) The examples of Ex. 5, (β) $\sqrt{1+i}\sqrt{1-i}$, (γ) $\sqrt{-2+2\sqrt{3}i}$,
(δ) $(\sqrt{1+i}+\sqrt{1-i})^2$, ($\epsilon$) $\sqrt{\sqrt{2}+\sqrt{-2}}$, ($\zeta$) $\sqrt[3]{2+2\sqrt{3}i}$,
(η) $\sqrt[4]{16(\cos 200^\circ + i\sin 200^\circ)}$, ($\theta$) $\sqrt[5]{-1}$, (ι) $\sqrt[6]{8i}$.

9. Find the equations of analytic geometry which represent the transformation equivalent to multiplication by $\alpha = -1+\sqrt{-3}$.

10. Show that $|z-\alpha| = r$, where z is a variable and α a fixed complex number, is the equation of the circle $(x-a)^2+(y-b)^2 = r^2$.

11. Find $\cos 5x$ and $\cos 8x$ in terms of $\cos x$, and $\sin 6x$ and $\sin 7x$ in terms of $\sin x$.

12. Obtain to four decimal places the five roots $\sqrt[5]{1}$.

13. If $z = x+iy$ and $z' = x'+iy'$, show that $z' = (\cos\phi - i\sin\phi)z - \alpha$ is the formula for shifting the axes through the vector distance $\alpha = a+ib$ to the new origin (a, b) and turning them through the angle ϕ. Deduce the ordinary equations of transformation.

14. Show that $|z-\alpha| = k|z-\beta|$, where k is real, is the equation of a circle; specify the position of the circle carefully. Use the theorem: The locus of points whose distances to two fixed points are in a constant ratio is a circle the diameter of which is divided internally and externally in the same ratio by the fixed points.

15. The transformation $z' = \dfrac{az+b}{cz+d}$, where a, b, c, d are complex and $ad-bc \neq 0$, is called the *general linear transformation* of z into z'. Show that

$$|z'-\alpha'| = k|z'-\beta'| \quad \text{becomes} \quad |z-\alpha| = k\left|\frac{c\alpha'+d}{c\beta'+d}\right|\cdot|z-\beta|.$$

Hence infer that the transformation carries circles into circles, and points which divide a diameter internally and externally in the same ratio into points which divide some diameter of the new circle similarly, but generally with a different ratio.

73. Functions of a complex variable. Let $z = x+iy$ be a complex variable representable geometrically as a variable point in the xy-plane, which may be called the *complex plane.* As z determines the two real numbers x and y, any function $F(x, y)$ which is the sum of two single valued real functions in the form

$$F(x, y) = X(x, y) + iY(x, y) = R(\cos\Phi + i\sin\Phi) \qquad (19)$$

will be completely determined in value if z is given. Such a function is called a *complex function* (and not a function of the complex variable, for reasons that will appear later). The magnitude and angle of the function are determined by

$$R = \sqrt{X^2+Y^2}, \qquad \cos\Phi = \frac{X}{R},\ \sin\Phi = \frac{Y}{R}. \qquad (20)$$

The function F is continuous by definition when and only when both X and Y are continuous functions of (x, y); R is then continuous in (x, y) and F can vanish only when $R = 0$; the angle Φ regarded as a function of (x, y) is also continuous and determinate (except for the additive $2n\pi$) unless $R = 0$, in which case X and Y also vanish and the expression for Φ involves an indeterminate form in two variables and is generally neither determinate nor continuous (§ 44).

If the derivative of F with respect to z were sought for the value $z = a + ib$, the procedure would be entirely analogous to that in the case of a real function of a real variable. The increment $\Delta z = \Delta x + i\Delta y$ would be assumed for z and ΔF would be computed and the quotient $\Delta F/\Delta z$ would be formed. Thus by the Theorem of the Mean (§ 46),

$$\frac{\Delta F}{\Delta z} = \frac{\Delta X + i\Delta Y}{\Delta x + i\Delta y} = \frac{(X'_x + iY'_x)\Delta x + (X'_y + iY'_y)\Delta y}{\Delta x + i\Delta y} + \zeta, \qquad (21)$$

where the derivatives are formed for (a, b) and where ζ is an infinitesimal complex number. When Δz approaches 0, both Δx and Δy must approach 0 without any implied relation between them. In general the limit of $\Delta F/\Delta z$ is a double limit (§ 44) and may therefore depend on the way in which Δx and Δy approach their limit 0.

Now if first $\Delta y \doteq 0$ and then subsequently $\Delta x \doteq 0$, the value of the limit of $\Delta F/\Delta z$ is $X'_x + iY'_x$ taken at the point (a, b); whereas if first $\Delta x \doteq 0$ and then $\Delta y \doteq 0$, the value is $-iX'_y + Y'_y$. Hence if the limit of $\Delta F/\Delta z$ is to be independent of the way in which Δz approaches 0, it is surely necessary that

$$\frac{\partial X}{\partial x} + i\frac{\partial Y}{\partial x} = -i\frac{\partial X}{\partial y} + \frac{\partial Y}{\partial y},$$

or

$$\frac{\partial X}{\partial x} = \frac{\partial Y}{\partial y} \quad \text{and} \quad \frac{\partial X}{\partial y} = -\frac{\partial Y}{\partial x}. \qquad (22)$$

And conversely if these relations are satisfied, then

$$\frac{\Delta F}{\Delta z} = \left(\frac{\partial X}{\partial x} + i\frac{\partial Y}{\partial x}\right) + \zeta = \left(\frac{\partial Y}{\partial y} - i\frac{\partial X}{\partial y}\right) + \zeta;$$

and the limit is $X'_x + iY'_x = Y'_y - iX'_y$ taken at the point (a, b), and is independent of the way in which Δz approaches zero. The desirability of having at least the ordinary functions differentiable suggests the definition: *A complex function $F(x, y) = X(x, y) + iY(x, y)$ is considered as a function of the complex variable $z = x + iy$ when and only when X and Y are in general differentiable and satisfy the relations* (22). In this case *the derivative is*

$$F'(z) = \frac{dF}{dz} = \frac{\partial X}{\partial x} + i\frac{\partial Y}{\partial x} = \frac{\partial Y}{\partial y} - i\frac{\partial X}{\partial y}. \tag{23}$$

These conditions may also be expressed in polar coördinates (Ex. 2).

A few words about the function $\Phi(x, y)$. This is a multiple valued function of the variables (x, y), and the difference between two neighboring branches is the constant 2π. The application of the discussion of § 45 to this case shows at once that, in any simply connected region of the complex plane which contains no point (a, b) such that $R(a, b) = 0$, the different branches of $\Phi(x, y)$ may be entirely separated so that the value of Φ must return to its initial value when any closed curve is described by the point (x, y). If, however, the region is multiply connected or contains points for which $R = 0$ (which makes the region multiply connected because these points must be cut out), it may happen that there will be circuits for which Φ, although changing continuously, will not return to its initial value. Indeed if it can be shown that Φ does not return to its initial value when changing continuously as (x, y) describes the boundary of a region simply connected except for the excised points, it may be inferred that there must be points in the region for which $R = 0$.

An application of these results may be made to give a very simple demonstration of *the fundamental theorem of algebra that every equation of the nth degree has at least one root.* Consider the function

$$F(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n = X(x, y) + iY(x, y),$$

where X and Y are found by writing z as $x + iy$ and expanding and rearranging. The functions X and Y will be polynomials in (x, y) and will therefore be everywhere finite and continuous in (x, y). Consider the angle Φ of F. Then

$$\Phi = \text{ang. of } F = \text{ang. of } z^n\left(1 + \frac{a_1}{z} + \cdots + \frac{a_{n-1}}{z^{n-1}} + \frac{a_n}{z^n}\right) = \text{ang. of } z^n + \text{ang. of } (1 + \cdots).$$

Next draw about the origin a circle of radius r so large that

$$\left|\frac{a_1}{z}\right| + \cdots + \left|\frac{a_{n-1}}{z^{n-1}}\right| + \left|\frac{a_n}{z^n}\right| = \frac{|a_1|}{r} + \cdots + \frac{|a_{n-1}|}{r^{n-1}} + \frac{|a_n|}{r^n} < \epsilon.$$

Then for all points z upon the circumference the angle of F is

$$\Phi = \text{ang. of } F = n(\text{ang. of } z) + \text{ang. of } (1 + \eta), \qquad |\eta| < \epsilon.$$

Now let the point (x, y) describe the circumference. The angle of z will change by 2π for the complete circuit. Hence Φ must change by $2n\pi$ and does not return to its initial value. Hence there is within the circle at least one point (a, b) for which $R(a, b) = 0$ and consequently for which $X(a, b) = 0$ and $Y(a, b) = 0$ and $F(a, b) = 0$. Thus if $\alpha = a + ib$, then $F(\alpha) = 0$ and the equation $F(z) = 0$ is seen to have at least the one root α. It follows that $z - \alpha$ is a factor of $F(z)$; and hence by induction it may be seen that $F(z) = 0$ has just n roots.

74. The discussion of the algebra of complex numbers showed how the sum, difference, product, quotient, real powers, and real roots of such numbers could be found, and hence made it possible to compute the value of any given algebraic expression or function of z for a given value of z. It remains to show that any algebraic expression in z is really a function of z in the sense that it has a derivative with respect to z, and to find the derivative. Now the differentiation of an algebraic function of the variable x was made to depend upon the formulas of differentiation, (6) and (7) of § 2. A glance at the methods of derivation of these formulas shows that they were proved by ordinary algebraic manipulations such as have been seen to be equally possible with imaginaries as with reals. It therefore may be concluded that *an algebraic expression in z has a derivative with respect to z and that derivative may be found just as if z were a real variable.*

The case of the elementary functions e^z, $\log z$, $\sin z$, $\cos z$, $\cdots$ other than algebraic is different; for these functions have not been defined for complex variables. Now in seeking to define these functions when z is complex, an effort should be made to define in such a way that: 1° when z is real, the new and the old definitions become identical; and 2° the rules of operation with the function shall be as nearly as possible the same for the complex domain as for the real. Thus it would be desirable that $De^z = e^z$ and $e^{z+w} = e^z e^w$, when z and w are complex. With these ideas in mind one may proceed to define the elementary functions for complex arguments. Let

$$e^z = R(x, y)[\cos \Phi(x, y) + i \sin \Phi(x, y)]. \tag{24}$$

The derivative of this function is, by the first rule of (23),

$$\begin{aligned} De^z &= \frac{\partial}{\partial x}(R\cos\Phi) + i\frac{\partial}{\partial x}(R\sin\Phi) \\ &= (R'_x \cos\Phi - R\sin\Phi \cdot \Phi'_x) + i(R'_x \sin\Phi + R\cos\Phi\cdot\Phi'_x), \end{aligned}$$

and if this is to be identical with e^z above, the equations

$$\begin{matrix} R'_x\cos\Phi - R\Phi'_x\sin\Phi = R\cos\Phi \\ R'_x\sin\Phi + R\Phi'_x\cos\Phi = R\sin\Phi \end{matrix} \quad \text{or} \quad \begin{matrix} R'_x = R \\ \Phi'_x = 0 \end{matrix}$$

must hold, where the second pair is obtained by solving the first. If the second form of the derivative in (23) had been used, the results would have been $R'_y = 0$, $\Phi'_y = 1$. It therefore appears that if the derivative of e^z, however computed, is to be e^z, then

$$R'_x = R, \quad R'_y = 0, \quad \Phi'_x = 0, \quad \Phi'_y = 1$$

are four conditions imposed upon R and Φ. These conditions will be satisfied if $R = e^x$ and $\Phi = y$.* Hence define

$$e^z = e^{x+iy} = e^x(\cos y + i\sin y). \tag{25}$$

* The use of the more general solutions $R = Ge^x$, $\Phi = y + C$ would lead to expressions which would not reduce to e^x when $y = 0$ and $z = x$ or would not satisfy $e^{z+w} = e^z e^w$.

With this definition De^z is surely e^z, and it is readily shown that the exponential law $e^{z+w} = e^z e^w$ holds.

For the special values $\frac{1}{2}\pi i$, πi, $2\pi i$ of z the value of e^z is

$$e^{\frac{1}{2}\pi i} = i, \quad e^{\pi i} = -1, \quad e^{2\pi i} = 1.$$

Hence it appears that if $2n\pi i$ be added to z, e^z is unchanged;

$$e^{z+2n\pi i} = e^z, \quad \text{period} \quad 2\pi i. \tag{26}$$

Thus *in the complex domain e^z has the period $2\pi i$*, just as $\cos x$ and $\sin x$ have the real period 2π. This relation is inherent; for

$$e^{yi} = \cos y + i \sin y, \quad e^{-yi} = \cos y - i \sin y,$$

and
$$\cos y = \frac{e^{yi} + e^{-yi}}{2}, \quad \sin y = \frac{e^{yi} - e^{-yi}}{2i}. \tag{27}$$

The trigonometric functions of a real variable y may be expressed in terms of the exponentials of yi and $-yi$. As the exponential has been defined for all complex values of z, it is natural to use (27) to define the trigonometric functions for complex values as

$$\cos z = \frac{e^{zi} + e^{-zi}}{2}, \quad \sin z = \frac{e^{zi} - e^{-zi}}{2i}. \tag{27'}$$

With these definitions the ordinary formulas for $\cos(z + w)$, $D \sin z$, $\cdots$ may be obtained and be seen to hold for complex arguments, just as the corresponding formulas were derived for the hyperbolic functions (§ 5).

As in the case of reals, the logarithm $\log z$ will be defined for complex numbers as the inverse of the exponential. Thus

$$\text{if} \quad e^z = w, \quad \text{then} \quad \log w = z + 2n\pi i, \tag{28}$$

where the periodicity of the function e^z shows that *the logarithm is not uniquely determined but admits the addition of $2n\pi i$ to any one of its values*, just as $\tan^{-1} x$ admits the addition of $n\pi$. If w is written as a complex number $u + iv$ with modulus $r = \sqrt{u^2 + v^2}$ and with the angle ϕ, it follows that

$$w = u + iv = r(\cos\phi + i \sin\phi) = re^{\phi i} = e^{\log r + \phi i}; \tag{29}$$

and
$$\log w = \log r + \phi i = \log\sqrt{u^2 + v^2} + i \tan^{-1}(v/u)$$

is the expression for the logarithm of w in terms of the modulus and angle of w; the $2n\pi i$ may be added if desired.

To this point the expression of a power a^b, where the exponent b is imaginary, has had no definition. The definition may now be given in terms of exponentials and logarithms. Let

$$a^b = e^{b \log a} \quad \text{or} \quad \log a^b = b \log a.$$

In this way the problem of computing a^b is reduced to one already solved. From the very definition it is seen that the logarithm of a power is the product of the exponent by the logarithm of the base, as in the case of reals. To indicate the path that has been followed in defining functions, a sort of family tree may be made.

real numbers, x — real powers and roots of reals, x^n

real angles, x — real trigonometric functions, $\cos x$, $\sin x$, $\tan^{-1} x$, $\cdots$

exponentials, logarithms of reals, e^x, $\log x$

real powers and roots of imaginaries, z^n

exponentials of imaginaries, e^z

logarithms of imaginaries, $\log z$ — imaginary powers, z^α

trigonometric functions of imaginaries

EXERCISES

1. Show that the following complex functions satisfy the conditions (22) and are therefore functions of the complex variable z. Find $F'(z)$:

(α) $x^2 - y^2 + 2\,ixy$, (β) $x^3 - 3(xy^2 + x^2 - y^2) + i(3x^2y - y^3 - 6xy)$,

(γ) $\dfrac{x}{x^2+y^2} - i\dfrac{y}{x^2+y^2}$, ($\delta$) $\log\sqrt{x^2+y^2} + i\tan^{-1}\dfrac{y}{x}$,

(ϵ) $e^x\cos y + ie^x\sin y$, (ζ) $\sin x\sinh y + i\cos x\cosh y$.

2. Show that in polar coördinates the conditions for the existence of $F'(z)$ are

$$\frac{\partial X}{\partial r} = \frac{1}{r}\frac{\partial Y}{\partial \phi}, \quad \frac{\partial Y}{\partial r} = -\frac{1}{r}\frac{\partial X}{\partial \phi} \quad \text{with} \quad F'(z) = \left(\frac{\partial X}{\partial r} + i\frac{\partial Y}{\partial r}\right)(\cos\phi - i\sin\phi).$$

3. Use the conditions of Ex. 2 to show from $D\log z = z^{-1}$ that $\log z = \log r + \phi i$.

4. From the definitions given above prove the formulas

$$(\alpha)\ \sin(x+iy) = \sin x\cosh y + i\cos x\sinh y,$$
$$(\beta)\ \cos(x+iy) = \cos x\cosh y - i\sin x\sinh y,$$
$$(\gamma)\ \tan(x+iy) = \frac{\sin 2x + i\sinh 2y}{\cos 2x + \cosh 2y}.$$

5. Find to three decimals the complex numbers which express the values of:

(α) $e^{\frac{1}{4}\pi i}$, (β) e^i, (γ) $e^{\frac{1}{2}+\frac{1}{2}\sqrt{-3}}$, ($\delta$) e^{-1-i},

(ϵ) $\sin\frac{1}{4}\pi i$, (ζ) $\cos i$, (η) $\sin(\frac{1}{2}+\frac{1}{2}\sqrt{-3})$, ($\theta$) $\tan(-1-i)$,

(ι) $\log(-1)$, (κ) $\log i$, (λ) $\log(\frac{1}{2}+\frac{1}{2}\sqrt{-3})$, ($\mu$) $\log(-1-i)$.

6. Owing to the fact that $\log a$ is multiple valued, a^b is multiple valued in such a manner that any one value may be multiplied by $e^{2n\pi bi}$. Find one value of each of the following and several values of one of them:

(α) 2^i, (β) i^i, (γ) $\sqrt[i]{i}$, (δ) $\sqrt[i]{2}$, (ϵ) $(\frac{1}{2}+\frac{1}{2}\sqrt{-3})^{\frac{3}{\pi}i+1}$,

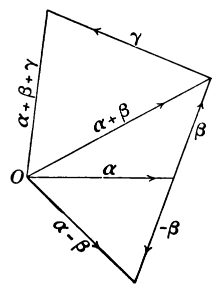
7. Show that $Da^z = a^z \log a$ when a and z are complex.

8. Show that $(a^b)^c = a^{bc}$; and fill in such other steps as may be suggested by the work in the text, which for the most part has merely been sketched in a broad way.

9. Show that if $f(z)$ and $g(z)$ are two functions of a complex variable, then $f(z) \pm g(z)$, $\alpha f(z)$ with α a complex constant, $f(z)g(z)$, $f(z)/g(z)$ are also functions of z.

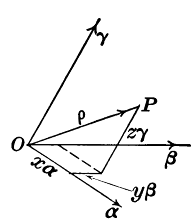
10. Obtain logarithmic expressions for the inverse trigonometric functions. Find $\sin^{-1} i$.

75. Vector sums and products. As stated in § 71, a vector is a quantity which has magnitude and direction. If the magnitudes of two vectors are equal and the directions of the two vectors are the same, the vectors are said to be equal irrespective of the position which they occupy in space. The vector $-\alpha$ is by definition a vector which has the same magnitude as α but the opposite direction. The vector $m\alpha$ is a vector which has the same direction as α (or the opposite) and is m (or $-m$) times as long. The law of vector or geometric addition is the parallelogram or triangle law (§ 71) and is still applicable when the vectors do not lie in a plane but have any directions in space; for any two vectors brought end to end determine a plane in which the construction may be carried out. Vectors will be designated by Greek small letters or by letters in heavy type. The relations of equality or similarity between triangles establish the rules



$$\alpha + \beta = \beta + \alpha, \quad \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma, \quad m(\alpha + \beta) = m\alpha + m\beta \qquad (30)$$

as true for vectors as well as for numbers whether real or complex. A vector is said to be zero when its magnitude is zero, and it is written 0. From the definition of addition it follows that $\alpha + 0 = \alpha$. In fact *as far as addition, subtraction, and multiplication by numbers are concerned, vectors obey the same formal laws as numbers.*



A vector ρ may be resolved into components parallel to any three given vectors α, β, γ which are not parallel to any one plane. For let a parallelepiped be constructed with its edges parallel to the three given vectors and with its diagonal equal to the vector whose components are desired. The edges of the parallelepiped are then certain multiples $x\alpha$, $y\beta$, $z\gamma$ of α, β, γ; and these are the desired components of ρ. The vector ρ may be written as

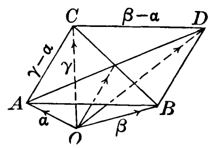
$$\rho = x\alpha + y\beta + z\gamma.^* \tag{31}$$

It is clear that two equal vectors would necessarily have the same components along three given directions and that the components of a zero vector would all be zero. Just as the equality of two complex numbers involved the two equalities of the respective real and imaginary parts, so the equality of two vectors as

$$\rho = x\alpha + y\beta + z\gamma = x'\alpha + y'\beta + z'\gamma = \rho' \tag{31'}$$

involves the three equations $x = x'$, $y = y'$, $z = z'$.

As a problem in the use of vectors let there be given the three vectors α, β, γ from an assumed origin O to three vertices of a parallelogram; required the vector to the other vertex, the vector expressions for the sides and diagonals of the parallelogram, and the proof of the fact that the diagonals bisect each other. Consider the figure. The side AB is, by the triangle law, that vector which when added to $OA = \alpha$ gives $OB = \beta$, and hence it must be that $AB = \beta - \alpha$. In like manner $AC = \gamma - \alpha$. Now OD is the sum of OC and CD, and $CD = AB$; hence $OD = \gamma + \beta - \alpha$. The diagonal AD is the difference of the vectors OD and OA, and is therefore $\gamma + \beta - 2\alpha$. The diagonal BC is $\gamma - \beta$. Now the vector from O to the middle point of BC may be found by adding to OB one half of BC. Hence this vector is $\beta + \frac{1}{2}(\gamma - \beta)$ or $\frac{1}{2}(\beta + \gamma)$. In like manner the vector to the middle point of AD is seen to be $\alpha + \frac{1}{2}(\gamma + \beta - 2\alpha)$ or $\frac{1}{2}(\gamma + \beta)$, which is identical with the former. The two middle points therefore coincide and the diagonals bisect each other.



Let α and β be any two vectors, $|\alpha|$ and $|\beta|$ their respective lengths, and $\angle(\alpha, \beta)$ the angle between them. For convenience the vectors may be considered to be laid off from the same origin. The product of the lengths of the vectors by the cosine of the angle between the vectors is called the *scalar product*,

$$\text{scalar product} = \alpha\cdot\beta = |\alpha||\beta|\cos\angle(\alpha, \beta), \tag{32}$$

of the two vectors and is denoted by placing a dot between the letters. This combination, called the scalar product, is a number, not a vector. As $|\beta|\cos\angle(\alpha, \beta)$ is the projection of β upon the direction of α, the scalar product may be stated to be equal to the product of the length of either vector by the length of the projection of the other upon it. In particular if either vector were of unit length, the scalar product would be the projection of the other upon it, with proper regard for

* The numbers x, y, z are the oblique coördinates of the terminal end of ρ (if the initial end be at the origin) referred to a set of axes which are parallel to α, β, γ and upon which the unit lengths are taken as the lengths of α, β, γ respectively.

the sign; and if both vectors are unit vectors, the product is the cosine of the angle between them.

The scalar product, from its definition, is *commutative* so that $\alpha\cdot\beta=\beta\cdot\alpha$. Moreover $(m\alpha)\cdot\beta = \alpha\cdot(m\beta) = m(\alpha\cdot\beta)$, thus allowing a numerical factor m to be combined with either factor of the product. Furthermore the *distributive law*

$$\alpha\cdot(\beta+\gamma) = \alpha\cdot\beta + \alpha\cdot\gamma \quad \text{or} \quad (\alpha+\beta)\cdot\gamma = \alpha\cdot\gamma + \beta\cdot\gamma \tag{33}$$

is satisfied as in the case of numbers. For if α be written as the product $a\alpha_1$ of its length a by a vector α_1 of unit length in the direction of α, the first equation becomes

$$a\alpha_1\cdot(\beta+\gamma) = a\alpha_1\cdot\beta + a\alpha_1\cdot\gamma \quad \text{or} \quad \alpha_1\cdot(\beta+\gamma) = \alpha_1\cdot\beta + \alpha_1\cdot\gamma.$$

And now $\alpha_1\cdot(\beta+\gamma)$ is the projection of the sum $\beta+\gamma$ upon the direction of α, and $\alpha_1\cdot\beta + \alpha_1\cdot\gamma$ is the sum of the projections of β and γ upon this direction; by the law of projections these are equal and hence the distributive law is proved.

The associative law does not hold for scalar products; for $(\alpha\cdot\beta)\gamma$ means that the vector γ is multiplied by the number $\alpha\cdot\beta$, whereas $\alpha(\beta\cdot\gamma)$ means that α is multiplied by $(\beta\cdot\gamma)$, a very different matter. The laws of cancellation cannot hold; for if

$$\alpha\cdot\beta = 0, \quad \text{then} \quad |\alpha||\beta|\cos\angle(\alpha, \beta) = 0, \tag{34}$$

and the vanishing of the scalar product $\alpha\cdot\beta$ implies either that one of the factors is 0 or that the two vectors are perpendicular. In fact $\alpha\cdot\beta = 0$ is called the *condition of perpendicularity*. It should be noted, however, that if a vector ρ satisfies

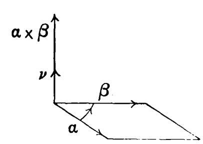
$$\rho\cdot\alpha = 0, \qquad \rho\cdot\beta = 0, \qquad \rho\cdot\gamma = 0, \tag{35}$$

three conditions of perpendicularity with three vectors α, β, γ not parallel to the same plane, the inference is that $\rho = 0$.

76. Another product of two vectors is the *vector product*,

$$\text{vector product} = \alpha\times\beta = \nu|\alpha||\beta|\sin\angle(\alpha, \beta), \tag{36}$$

where ν represents a vector of unit length normal to the plane of α and β upon that side on which rotation from α to β through an angle of less than 180° appears positive or counterclockwise. Thus the vector product is itself a vector of which the direction is perpendicular to each factor, and of which the magnitude is the product of the magnitudes into the sine of the included angle. The magnitude is therefore equal to the area of the parallelogram of which the vectors α and β are the sides.



The vector product will be represented by a cross inserted between the letters.

As rotation from β to α is the opposite of that from α to β, it follows from the definition of the vector product that

$$\beta\times\alpha = -\alpha\times\beta, \quad \text{not} \quad \alpha\times\beta = \beta\times\alpha, \tag{37}$$

and the product is *not commutative*, the order of the factors must be carefully observed. Furthermore the equation

$$\alpha\times\beta = \nu|\alpha||\beta|\sin\angle(\alpha, \beta) = 0 \tag{38}$$

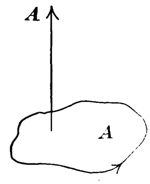
implies either that one of the factors vanishes or that the vectors α and β are parallel. Indeed the condition $\alpha\times\beta = 0$ is called the *condition of parallelism*. The laws of cancellation do not hold. The associative law also does not hold; for $(\alpha\times\beta)\times\gamma$ is a vector perpendicular to $\alpha\times\beta$ and γ, and since $\alpha\times\beta$ is perpendicular to the plane of α and β, the vector $(\alpha\times\beta)\times\gamma$ perpendicular to it must lie in the plane of α and β; whereas the vector $\alpha\times(\beta\times\gamma)$, by similar reasoning, must lie in the plane of β and γ; and hence the two vectors cannot be equal except in the very special case where each was parallel to β which is common to the two planes.

But the operation $(m\alpha)\times\beta = \alpha\times(m\beta) = m(\alpha\times\beta)$, which consists in allowing the transference of a numerical factor to any position in the product, does hold; and so does the *distributive law*

$$\alpha\times(\beta+\gamma) = \alpha\times\beta + \alpha\times\gamma \quad \text{and} \quad (\alpha+\beta)\times\gamma = \alpha\times\gamma + \beta\times\gamma, \tag{39}$$

the proof of which will be given below. In expanding according to the distributive law care must be exercised to keep the order of the factors in each vector product the same on both sides of the equation, owing to the failure of the commutative law; an interchange of the order of the factors changes the sign. It might seem as if any algebraic operations where so many of the laws of elementary algebra fail as in the case of vector products would be too restricted to be very useful; that this is not so is due to the astonishingly great number of problems in which the analysis can be carried on with only the laws of addition and the distributive law of multiplication combined with the possibility of transferring a numerical factor from one position to another in a product; in addition to these laws, the scalar product $\alpha\cdot\beta$ is commutative and the vector product $\alpha\times\beta$ is commutative except for change of sign.

In addition to segments of lines, *plane areas may be regarded as vector quantities;* for a plane area has magnitude (the amount of the area) and direction (the direction of the normal to its plane). To specify on which side of the plane the normal lies, some convention must be made. If the area is part of a surface inclosing a portion of space, the normal is taken as the exterior normal. If the area lies in an isolated plane, its positive side is determined only in connection with some assigned direction of description of its bounding curve; the rule is: If a person is assumed to walk along the boundary of an area in an assigned direction and upon that side of the plane which causes the inclosed area to lie upon his left, he is said to be upon the positive side (for the assigned direction of description of the boundary), and the vector which represents the area is the normal to that side. It has been mentioned that the vector product represented an area.



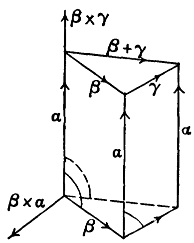
That the projection of a plane area upon a given plane gives an area which is the original area multiplied by the cosine of the angle between the two planes is a fundamental fact of projection, following from the simple fact that lines parallel to the intersection of the two planes are unchanged in length whereas lines perpendicular to the intersection are multiplied by the cosine of the angle between the planes. As the angle between the normals is the same as that between the planes, *the projection of an area upon a plane and the projection of the vector representing the area upon the normal to the plane are equivalent.* The projection of a *closed* area upon a plane is zero; for the area in the projection is covered twice (or an even number of times) with opposite signs and the total algebraic sum is therefore 0.

To prove the law $\alpha\times(\beta+\gamma)=\alpha\times\beta+\alpha\times\gamma$ and illustrate the use of the vector interpretation of areas, construct a triangular prism with the triangle on β, γ, and $\beta+\gamma$ as base and α as lateral edge. The total vector expression for the surface of this prism is

$$\beta\times\alpha+\gamma\times\alpha+\alpha\times(\beta+\gamma)+\tfrac{1}{2}(\beta\times\gamma)-\tfrac{1}{2}\beta\times\gamma=0,$$

and vanishes because the surface is closed. A cancellation of the equal and opposite terms (the two bases) and a simple transposition combined with the rule $\beta\times\alpha=-\alpha\times\beta$ gives the result

$$\alpha\times(\beta+\gamma)=-\beta\times\alpha-\gamma\times\alpha=\alpha\times\beta+\alpha\times\gamma.$$



A system of *vectors of reference* which is particularly useful consists of three vectors **i**, **j**, **k** of unit length directed along the axes X, Y, Z drawn so that rotation from X to Y appears positive from the side of the xy-plane upon which Z lies. The components of any vector **r** drawn from the origin to the point (x, y, z) are

$$x\mathbf{i},\quad y\mathbf{j},\quad z\mathbf{k},\quad \text{and}\quad \mathbf{r}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}.$$

The products of **i**, **j**, **k** into each other are, from the definitions,

$$\begin{aligned} \mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1, \\ \mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{i} = \mathbf{i}\cdot\mathbf{k} = 0, \\ \mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} = 0, \\ \mathbf{i}\times\mathbf{j} = -\mathbf{j}\times\mathbf{i} = \mathbf{k}, \quad \mathbf{j}\times\mathbf{k} = -\mathbf{k}\times\mathbf{j} = \mathbf{i}, \quad \mathbf{k}\times\mathbf{i} = -\mathbf{i}\times\mathbf{k} = \mathbf{j}. \end{aligned} \tag{40}$$

By means of these products and the distributive laws for scalar and vector products, any given products may be expanded. Thus if

$$\alpha = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad \text{and} \quad \beta = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k},$$

then

$$\alpha\cdot\beta = a_1b_1 + a_2b_2 + a_3b_3, \tag{41}$$

$$\alpha\times\beta = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k},$$

by direct multiplication. In this way a passage may be made from vector formulas to Cartesian formulas whenever desired.

EXERCISES

1. Prove geometrically that $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ and $m(\alpha + \beta) = m\alpha + m\beta$.

2. If α and β are the vectors from an assumed origin to A and B and if C divides AB in the ratio $m : n$, show that the vector to C is $\gamma = (n\alpha + m\beta)/(m + n)$.

3. In the parallelogram $ABCD$ show that the line BE connecting the vertex to the middle point of the opposite side CD is trisected by the diagonal AD and trisects it.

4. Show that the medians of a triangle meet in a point and are trisected.

5. If m_1 and m_2 are two masses situated at P_1 and P_2, the *center of gravity* or *center of mass* of m_1 and m_2 is defined as that point G on the line P_1P_2 which divides P_1P_2 inversely as the masses. Moreover if G_1 is the center of mass of a number of masses of which the total mass is M_1 and if G_2 is the center of mass of a number of other masses whose total mass is M_2, the same rule applied to M_1 and M_2 and G_1 and G_2 gives the center of gravity G of the total number of masses. Show that

$$\bar{\mathbf{r}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \quad \text{and} \quad \bar{\mathbf{r}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n}{m_1 + m_2 + \cdots + m_n} = \frac{\Sigma m\mathbf{r}}{\Sigma m},$$

where $\bar{\mathbf{r}}$ denotes the vector to the center of gravity. Resolve into components to show

$$\bar{x} = \frac{\Sigma mx}{\Sigma m}, \quad \bar{y} = \frac{\Sigma my}{\Sigma m}, \quad \bar{z} = \frac{\Sigma mz}{\Sigma m}.$$

6. If α and β are two fixed vectors and ρ a variable vector, all being laid off from the same origin, show that $(\rho - \beta)\cdot\alpha = 0$ is the equation of a plane through the end of β perpendicular to α.

7. Let α, β, γ be the vectors to the vertices A, B, C of a triangle. Write the three equations of the planes through the vertices perpendicular to the opposite sides. Show that the third of these can be derived as a combination of the other two; and hence infer that the three planes have a line in common and that the perpendiculars from the vertices of a triangle meet in a point.

8. Solve the problem analogous to Ex. 7 for the perpendicular bisectors of the sides.

9. Note that the length of a vector is $\sqrt{\alpha\cdot\alpha}$. If α, β, and $\gamma = \beta - \alpha$ are the three sides of a triangle, expand $\gamma\cdot\gamma = (\beta - \alpha)\cdot(\beta - \alpha)$ to obtain the law of cosines.

10. Show that the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of the sides. What does the difference of the squares of the diagonals equal ?

11. Show that $\dfrac{\alpha\cdot\beta}{\alpha\cdot\alpha}\,\alpha$ and $\dfrac{(\alpha\times\beta)\times\alpha}{\alpha\cdot\alpha}$ are the components of β parallel and perpendicular to α by showing 1° that these vectors have the right direction, and 2° that they have the right magnitude.

12. If α, β, γ are the three edges of a parallelepiped which start from the same vertex, show that $(\alpha\times\beta)\cdot\gamma$ is the volume of the parallelepiped, the volume being considered positive if γ lies on the same side of the plane of α and β with the vector $\alpha\times\beta$.

13. Show by Ex. 12 that $(\alpha\times\beta)\cdot\gamma = \alpha\cdot(\beta\times\gamma)$ and $(\alpha\times\beta)\cdot\gamma = (\beta\times\gamma)\cdot\alpha$; and hence infer that in a product of three vectors with cross and dot, the position of the cross and dot may be interchanged and the order of the factors may be permuted cyclically without altering the value. Show that the vanishing of $(\alpha\times\beta)\cdot\gamma$ or any of its equivalent expressions denotes that α, β, γ are parallel to the same plane; the condition $\alpha\times\beta\cdot\gamma = 0$ is called the condition of complanarity.

14. Assuming $\alpha = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\beta = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, $\gamma = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$, expand $\alpha\cdot\gamma$, $\alpha\cdot\beta$, and $\alpha\times(\beta\times\gamma)$ in terms of the coefficients to show

$$\alpha\times(\beta\times\gamma) = (\alpha\cdot\gamma)\,\beta - (\alpha\cdot\beta)\,\gamma; \quad \text{and hence} \quad (\alpha\times\beta)\times\gamma = (\alpha\cdot\gamma)\,\beta - (\gamma\cdot\beta)\,\alpha.$$

15. The formulas of Ex. 14 for expanding a product with two crosses and the rule of Ex. 13 that a dot and a cross may be interchanged may be applied to expand

$$(\alpha\times\beta)\times(\gamma\times\delta) = (\alpha\cdot\gamma\times\delta)\,\beta - (\beta\cdot\gamma\times\delta)\,\alpha = (\alpha\times\beta\cdot\delta)\,\gamma - (\alpha\times\beta\cdot\gamma)\,\delta$$

and

$$(\alpha\times\beta)\cdot(\gamma\times\delta) = (\alpha\cdot\gamma)(\beta\cdot\delta) - (\beta\cdot\gamma)(\alpha\cdot\delta).$$

16. If α and β are two unit vectors in the xy-plane inclined at angles θ and ϕ to the x-axis, show that

$$\alpha = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta, \quad \beta = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi;$$

and from the fact that $\alpha\cdot\beta = \cos(\phi - \theta)$ and $\alpha\times\beta = \mathbf{k}\sin(\phi - \theta)$ obtain by multiplication the trigonometric formulas for $\sin(\phi - \theta)$ and $\cos(\phi - \theta)$.

17. If l, m, n are direction cosines, the vector $l\mathbf{i} + m\mathbf{j} + n\mathbf{k}$ is a vector of unit length in the direction for which l, m, n are direction cosines. Show that the condition for perpendicularity of two directions (l, m, n) and (l', m', n') is $ll' + mm' + nn' = 0$.

18. With the same notations as in Ex. 14 show that

$$\alpha\cdot\alpha = a_1^2 + a_2^2 + a_3^2 \quad \text{and} \quad \alpha\times\beta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \quad \text{and} \quad \alpha\times\beta\cdot\gamma = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

19. Compute the scalar and vector products of these pairs of vectors:

$$(\alpha)\ \begin{cases} 6\mathbf{i}+0.3\mathbf{j}-5\mathbf{k} \\ 0.1\mathbf{i}-4.2\mathbf{j}+2.5\mathbf{k}, \end{cases} \qquad (\beta)\ \begin{cases} \mathbf{i}+2\mathbf{j}+3\mathbf{k} \\ -3\mathbf{i}-2\mathbf{j}+\mathbf{k}, \end{cases} \qquad (\gamma)\ \begin{cases} \mathbf{i}+\mathbf{k} \\ \mathbf{j}+\mathbf{i}. \end{cases}$$

20. Find the areas of the parallelograms defined by the pairs of vectors in Ex. 19. Find also the sine and cosine of the angles between the vectors.

21. Prove $\alpha\times[\beta\times(\gamma\times\delta)] = (\alpha\cdot\gamma\times\delta)\beta - \alpha\cdot\beta\,\gamma\times\delta = \beta\cdot\delta\,\alpha\times\gamma - \beta\cdot\gamma\,\alpha\times\delta$.

22. What is the area of the triangle (1, 1, 1), (0, 2, 3), (0, 0, $-$1)?

77. Vector differentiation. As the fundamental rules of differentiation depend on the laws of subtraction, multiplication by a number, the distributive law, and the rules permitting rearrangement, it follows that the rules must be applicable to expressions containing vectors without any changes except those implied by the fact that $\alpha\times\beta \neq \beta\times\alpha$. As an illustration consider the application of the definition of differentiation to the vector product $\mathbf{u}\times\mathbf{v}$ of two vectors which are supposed to be functions of a numerical variable, say x. Then

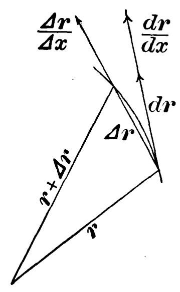
$$\Delta(\mathbf{u}\times\mathbf{v}) = (\mathbf{u}+\Delta\mathbf{u})\times(\mathbf{v}+\Delta\mathbf{v}) - \mathbf{u}\times\mathbf{v}$$
$$= \mathbf{u}\times\Delta\mathbf{v} + \Delta\mathbf{u}\times\mathbf{v} + \Delta\mathbf{u}\times\Delta\mathbf{v},$$

$$\frac{\Delta(\mathbf{u}\times\mathbf{v})}{\Delta x} = \mathbf{u}\times\frac{\Delta\mathbf{v}}{\Delta x} + \frac{\Delta\mathbf{u}}{\Delta x}\times\mathbf{v} + \frac{\Delta\mathbf{u}\times\Delta\mathbf{v}}{\Delta x},$$

$$\frac{d(\mathbf{u}\times\mathbf{v})}{dx} = \lim_{\Delta x \doteq 0}\frac{\Delta(\mathbf{u}\times\mathbf{v})}{\Delta x} = \mathbf{u}\times\frac{d\mathbf{v}}{dx} + \frac{d\mathbf{u}}{dx}\times\mathbf{v}.$$

Here the ordinary rule for a product is seen to hold, except that *the order of the factors must not be interchanged.*

The interpretation of the derivative is important. Let the variable vector $\mathbf{r}$ be regarded as a function of some variable, say x, and suppose $\mathbf{r}$ is laid off from an assumed origin so that, as x varies, the terminal point of $\mathbf{r}$ describes a curve. The increment $\Delta\mathbf{r}$ of $\mathbf{r}$ corresponding to Δx is a vector quantity and in fact is the chord of the curve as indicated. *The derivative*



$$\frac{d\mathbf{r}}{dx} = \lim\frac{\Delta\mathbf{r}}{\Delta x}, \quad \frac{d\mathbf{r}}{ds} = \lim\frac{\Delta\mathbf{r}}{\Delta s} = \mathbf{t} \qquad (42)$$

is therefore a vector tangent to the curve; in particular if the variable x were the arc s, the derivative would have the magnitude unity and would be a unit vector tangent to the curve.

The derivative or differential of a vector of constant length is perpendicular to the vector. This follows from the fact that the vector then describes a circle concentric with the origin. It may also be seen analytically from the equation

$$d(\mathbf{r}\cdot\mathbf{r}) = d\mathbf{r}\cdot\mathbf{r} + \mathbf{r}\cdot d\mathbf{r} = 2\,\mathbf{r}\cdot d\mathbf{r} = d\text{ const.} = 0. \tag{43}$$

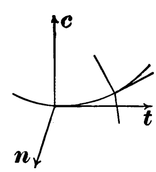
If the vector of constant length is of length unity, the increment $\Delta\mathbf{r}$ is the chord in a unit circle and, apart from infinitesimals of higher order, it is equal in magnitude to the angle subtended at the center. Consider then the derivative of the unit tangent $\mathbf{t}$ to a curve with respect to the arc s. The magnitude of $d\mathbf{t}$ is the angle the tangent turns through and the direction of $d\mathbf{t}$ is normal to $\mathbf{t}$ and hence to the curve. The vector quantity,

$$\text{curvature} \quad \mathbf{C} = \frac{d\mathbf{t}}{ds} = \frac{d^2\mathbf{r}}{ds^2}, \tag{44}$$

therefore has the magnitude of the curvature (by the definition in § 42) and the direction of the interior normal to the curve.

This work holds equally for plane or space curves. In the case of a space curve the plane which contains the tangent $\mathbf{t}$ and the curvature $\mathbf{C}$ is called the osculating plane (§ 41). By definition (§ 42) the *torsion of a space curve* is the rate of turning of the osculating plane with the arc, that is, $d\psi/ds$. To find the torsion by vector methods let $\mathbf{c}$ be a unit vector $\mathbf{C}/\sqrt{\mathbf{C}\cdot\mathbf{C}}$ along $\mathbf{C}$. Then as $\mathbf{t}$ and $\mathbf{c}$ are perpendicular, $\mathbf{n} = \mathbf{t}\times\mathbf{c}$ is a unit vector perpendicular to the osculating plane and $d\mathbf{n}$ will equal $d\psi$ in magnitude. Hence as a vector quantity the torsion is

$$\mathbf{T} = \frac{d\mathbf{n}}{ds} = \frac{d(\mathbf{t}\times\mathbf{c})}{ds} = \frac{d\mathbf{t}}{ds}\times\mathbf{c} + \mathbf{t}\times\frac{d\mathbf{c}}{ds} = \mathbf{t}\times\frac{d\mathbf{c}}{ds}, \tag{45}$$

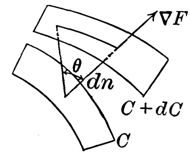


where (since $d\mathbf{t}/ds = \mathbf{C}$, and $\mathbf{c}$ is parallel to $\mathbf{C}$) the first term drops out. Next note that $d\mathbf{n}$ is perpendicular to $\mathbf{n}$ because it is the differential of a unit vector, and is perpendicular to $\mathbf{t}$ because $d\mathbf{n} = d(\mathbf{t}\times\mathbf{c}) = \mathbf{t}\times d\mathbf{c}$ and $\mathbf{t}\cdot(\mathbf{t}\times d\mathbf{c}) = 0$ since $\mathbf{t}$, $\mathbf{t}$, $d\mathbf{c}$ are necessarily complanar (Ex. 12, p. 169). Hence $\mathbf{T}$ is parallel to $\mathbf{c}$. It is convenient to consider the torsion as positive when the osculating plane seems to turn in the positive direction when viewed from the side of the normal plane upon which $\mathbf{t}$ lies. An inspection of the figure shows that in this case $d\mathbf{n}$ has the direction $-\mathbf{c}$ and not $+\mathbf{c}$. As $\mathbf{c}$ is a unit vector, the numerical value of the torsion is therefore $-\mathbf{c}\cdot\mathbf{T}$. Then

$$\begin{aligned} T = -\mathbf{c}\cdot\mathbf{T} &= -\,\mathbf{c}\cdot\mathbf{t}\times\frac{d\mathbf{c}}{ds} = -\,\mathbf{c}\cdot\mathbf{t}\times\frac{d}{ds}\frac{\mathbf{C}}{\sqrt{\mathbf{C}\cdot\mathbf{C}}} \\ &= -\,\mathbf{c}\cdot\mathbf{t}\times\left[\frac{d^3\mathbf{r}}{ds^3}\frac{1}{\sqrt{\mathbf{C}\cdot\mathbf{C}}} + \mathbf{C}\frac{d}{ds}\frac{1}{\sqrt{\mathbf{C}\cdot\mathbf{C}}}\right] = -\mathbf{c}\cdot\mathbf{t}\times\frac{d^3\mathbf{r}}{ds^3}\frac{1}{\sqrt{\mathbf{C}\cdot\mathbf{C}}} \\ &= \mathbf{t}\cdot\frac{\mathbf{C}}{\mathbf{C}\cdot\mathbf{C}}\times\frac{d^3\mathbf{r}}{ds^3} = \frac{\mathbf{r}'\cdot\mathbf{r}''\times\mathbf{r}'''}{\mathbf{r}''\cdot\mathbf{r}''}, \end{aligned} \tag{45'}$$

where differentiation with respect to s is denoted by accents.

78. Another sort of relation between vectors and differentiation comes to light in connection with the normal and directional derivatives (§ 48). If $F(x, y, z)$ is a function which has a definite value at each point of space and if the two neighboring surfaces $F = C$ and $F = C + dC$ are considered, the normal derivative of F is the rate of change of F along the normal to the surfaces and is written dF/dn. The rate of change of F along the normal to the surface $F = C$ is more rapid than along any other direction; for the change in F between the two surfaces is $dF = dC$ and is constant, whereas the distance dn between the two surfaces is least (apart from infinitesimals of higher order) along the normal. In fact if dr denote the distance along any other direction, the relations shown by the figure are



$$dr = \sec\theta dn \quad \text{and} \quad \frac{dF}{dr} = \frac{dF}{dn}\cos\theta. \tag{46}$$

If now $\mathbf{n}$ denote a vector of unit length normal to the surface, *the product* $\mathbf{n}dF/dn$ *will be a vector quantity which has both the magnitude and the direction of most rapid increase of* F. Let

$$\mathbf{n}\frac{dF}{dn} = \nabla F = \text{grad } F \tag{47}$$

be the symbolic expressions for this vector, where ∇F is read as "del F" and grad F is read as "the gradient of F." If $d\mathbf{r}$ be the vector of which dr is the length, the scalar product $\mathbf{n}\cdot d\mathbf{r}$ is precisely $\cos\theta dr$, and hence it follows that

$$d\mathbf{r}\cdot\nabla F = dF \quad \text{and} \quad \mathbf{r}_1\cdot\nabla F = \frac{dF}{dr}, \tag{48}$$

where $\mathbf{r}_1$ is a unit vector in the direction $d\mathbf{r}$. The second of the equations shows that *the directional derivative in any direction is the component or projection of the gradient in that direction.*

From this fact the expression of the gradient may be found in terms of its components along the axes. For the derivatives of F along the axes are $\partial F/\partial x$, $\partial F/\partial y$, $\partial F/\partial z$, and as these are the components of ∇F along the directions $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, the result is

$$\nabla F = \text{grad } F = \mathbf{i}\frac{\partial F}{\partial x} + \mathbf{j}\frac{\partial F}{\partial y} + \mathbf{k}\frac{\partial F}{\partial z}.$$

Hence

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{49}$$

may be regarded as a symbolic vector-differentiating operator which when applied to F gives the gradient of F. The product

$$d\mathbf{r}\cdot\nabla F = \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y} + dz\frac{\partial}{\partial z}\right)F = dF \tag{50}$$

is immediately seen to give the ordinary expression for dF. From this form of grad F it does not appear that the gradient of a function is independent of the choice of axes, but from the manner of derivation of ∇F first given it does appear that grad F is a definite vector quantity independent of the choice of axes.

In the case of any given function F the gradient may be found by the application of the formula (49); but in many instances it may also be found by means of the important relation $d\mathbf{r}\cdot\nabla F = dF$ of (48). For instance to prove the formula $\nabla(FG) = F\nabla G + G\nabla F$, the relation may be applied as follows:

$$d\mathbf{r}\cdot\nabla(FG) = d(FG) = FdG + GdF$$
$$= Fd\mathbf{r}\cdot\nabla G + Gd\mathbf{r}\cdot\nabla F = d\mathbf{r}\cdot(F\nabla G + G\nabla F).$$

Now as these equations hold for any direction $d\mathbf{r}$, the $d\mathbf{r}$ may be canceled by (35), p. 165, and the desired result is obtained.

The use of vector notations for treating assigned practical problems involving computation is not great, but for handling the general theory of such parts of physics as are essentially concerned with direct quantities, mechanics, hydromechanics, electromagnetic theories, etc., the actual use of the vector algorisms considerably shortens the formulas and has the added advantage of operating directly upon the magnitudes involved. At this point some of the elements of mechanics will be developed.

79. According to Newton's Second Law, when a force acts upon a particle of mass m, *the rate of change of momentum is equal to the force acting, and takes place in the direction of the force.* It therefore appears that the rate of change of momentum and momentum itself are to be regarded as vector or directed magnitudes in the application of the Second Law. Now if the vector $\mathbf{r}$, laid off from a fixed origin to the point at which the moving mass m is situated at any instant of time t, be differentiated with respect to the time t, the derivative $d\mathbf{r}/dt$ is a vector, tangent to the curve in which the particle is moving and of magnitude equal to ds/dt or v, the velocity of motion. As vectors*, then, the velocity $\mathbf{v}$ and the momentum and the force may be written as

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}, \qquad m\mathbf{v}, \qquad \mathbf{F} = \frac{d}{dt}(m\mathbf{v}). \tag{51}$$

Hence $$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = m\frac{d^2\mathbf{r}}{dt^2} = m\mathbf{f} \quad \text{if} \quad \mathbf{f} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2}.$$

From the equations it appears that the force $\mathbf{F}$ is the product of the mass m by a vector $\mathbf{f}$ which is the rate of change of the velocity regarded

* In applications, it is usual to denote vectors by heavy type and to denote the magnitudes of those vectors by corresponding italic letters.

as a vector. The vector **f** is called the *acceleration;* it must not be confused with the rate of change dv/dt or d^2s/dt^2 of the speed or magnitude of the velocity. The components f_x, f_y, f_z of the acceleration along the axes are the projections of **f** along the directions **i**, **j**, **k** and may be written as **f·i**, **f·j**, **f·k**. Then by the laws of differentiation it follows that

$$f_x = \mathbf{f}\cdot\mathbf{i} = \frac{d\mathbf{v}}{dt}\cdot\mathbf{i} = \frac{d\,(\mathbf{v}\cdot\mathbf{i})}{dt} = \frac{dv_x}{dt},$$

or

$$f_x = \mathbf{f}\cdot\mathbf{i} = \frac{d^2\mathbf{r}}{dt^2}\cdot\mathbf{i} = \frac{d^2\,(\mathbf{r}\cdot\mathbf{i})}{dt^2} = \frac{d^2x}{dt^2}.$$

Hence

$$f_x = \frac{d^2x}{dt^2}, \qquad f_y = \frac{d^2y}{dt^2}, \qquad f_z = \frac{d^2z}{dt^2},$$

and it is seen that the components of the acceleration are the accelerations of the components. If X, Y, Z are the components of the force, the equations of motion in rectangular coördinates are

$$m\frac{d^2x}{dt^2} = X, \qquad m\frac{d^2y}{dt^2} = Y, \qquad m\frac{d^2z}{dt^2} = Z. \tag{52}$$

Instead of resolving the acceleration, force, and displacement along the axes, it may be convenient to resolve them along the tangent and normal to the curve. The velocity **v** may be written as $v\mathbf{t}$, where v is the magnitude of the velocity and **t** is a unit vector tangent to the curve. Then

$$\mathbf{f} = \frac{d\mathbf{v}}{dt} = \frac{d(v\mathbf{t})}{dt} = \frac{dv}{dt}\mathbf{t} + v\frac{d\mathbf{t}}{dt}.$$

But

$$\frac{d\mathbf{t}}{dt} = \frac{d\mathbf{t}}{ds}\frac{ds}{dt} = \mathbf{C}v = \frac{v}{R}\mathbf{n}, \tag{53}$$

where R is the radius of curvature and **n** is a unit normal. Hence

$$\mathbf{f} = \frac{d^2s}{dt^2}\mathbf{t} + \frac{v^2}{R}\mathbf{n}, \qquad f_t = \frac{d^2s}{dt^2}, \qquad f_n = \frac{v^2}{R}. \tag{53'}$$

It therefore is seen that the component of the acceleration along the tangent is d^2s/dt^2, or the rate of change of the velocity regarded as a number, and the component normal to the curve is v^2/R. If T and N are the components of the force along the tangent and normal to the curve of motion, the equations are

$$T = mf_t = m\frac{d^2s}{dt^2}, \qquad N = mf_n = m\frac{v^2}{R}.$$

It is noteworthy that the force must lie in the osculating plane.

If **r** and $\mathbf{r} + \Delta\mathbf{r}$ are two positions of the radius vector, the area of the sector included by them is (except for infinitesimals of higher order)

$\Delta\mathbf{A} = \frac{1}{2}\mathbf{r}\times(\mathbf{r} + \Delta\mathbf{r}) = \frac{1}{2}\mathbf{r}\times\Delta\mathbf{r}$, and is a vector quantity of which the direction is normal to the plane of $\mathbf{r}$ and $\mathbf{r} + \Delta\mathbf{r}$, that is, to the plane through the origin tangent to the curve. The rate of description of area, or the *areal velocity*, is therefore

$$\frac{d\mathbf{A}}{dt} = \lim \tfrac{1}{2}\mathbf{r}\times\frac{\Delta\mathbf{r}}{\Delta t} = \tfrac{1}{2}\mathbf{r}\times\frac{d\mathbf{r}}{dt} = \tfrac{1}{2}\mathbf{r}\times\mathbf{v}. \tag{54}$$

The projections of the areal velocities on the coördinate planes, which are the same as the areal velocities of the projection of the motion on those planes, are (Ex. 11 below)

$$\frac{1}{2}\left(y\frac{dz}{dt} - z\frac{dy}{dt}\right), \quad \frac{1}{2}\left(z\frac{dx}{dt} - x\frac{dz}{dt}\right), \quad \frac{1}{2}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right). \tag{54'}$$

If the force $\mathbf{F}$ acting on the mass m passes through the origin, then $\mathbf{r}$ and $\mathbf{F}$ lie along the same direction and $\mathbf{r}\times\mathbf{F} = 0$. The equation of motion may then be integrated at sight.

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F}, \qquad m\mathbf{r}\times\frac{d\mathbf{v}}{dt} = \mathbf{r}\times\mathbf{F} = 0,$$

$$\mathbf{r}\times\frac{d\mathbf{v}}{dt} = \frac{d}{dt}(\mathbf{r}\times\mathbf{v}) = 0, \qquad \mathbf{r}\times\mathbf{v} = \text{const.}$$

It is seen that in this case the rate of description of area is a constant vector, which means that the rate is not only constant in magnitude but is constant in direction, that is, the path of the particle m must lie in a plane through the origin. When the force passes through a fixed point, as in this case, the force is said to be *central*. Therefore when a particle moves under the action of a central force, the motion takes place in a plane passing through the center and the rate of description of areas, or the areal velocity, is constant.

80. If there are several particles, say n, in motion, each has its own equation of motion. These equations may be combined by addition and subsequent reduction.

$$m_1\frac{d^2\mathbf{r}_1}{dt^2} = \mathbf{F}_1,\ m_2\frac{d^2\mathbf{r}_2}{dt^2} = \mathbf{F}_2,\ \cdots,\ m_n\frac{d^2\mathbf{r}_n}{dt^2} = \mathbf{F}_n,$$

and $$m_1\frac{d^2\mathbf{r}_1}{dt^2} + m_2\frac{d^2\mathbf{r}_2}{dt^2} + \cdots + m_n\frac{d^2\mathbf{r}_n}{dt^2} = \mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n.$$

But $$m_1\frac{d^2\mathbf{r}_1}{dt^2} + m_2\frac{d^2\mathbf{r}_2}{dt^2} + \cdots + m_n\frac{d^2\mathbf{r}_n}{dt^2} = \frac{d^2}{dt^2}(m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n).$$

Let $$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n = (m_1 + m_2 + \cdots + m_n)\,\bar{\mathbf{r}} = M\,\bar{\mathbf{r}}$$

or $$\bar{\mathbf{r}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n}{m_1 + m_2 + \cdots + m_n} = \frac{\Sigma m\mathbf{r}}{\Sigma m} = \frac{\Sigma m\mathbf{r}}{M}.$$

Then $$M\frac{d^2\bar{\mathbf{r}}}{dt^2} = \mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n = \sum\mathbf{F}. \tag{55}$$

Now the vector $\mathbf{r}$ which has been here introduced is the vector of the center of mass or center of gravity of the particles (Ex. 5, p. 168). The result (55) states, on comparison with (51), that the center of gravity of the n masses moves as if all the mass M were concentrated at it and all the forces applied there.

The force $\mathbf{F}_i$ acting on the ith mass may be wholly or partly due to attractions, repulsions, pressures, or other actions exerted on that mass by one or more of the other masses of the system of n particles. In fact let $\mathbf{F}_i$ be written as

$$\mathbf{F}_i = \mathbf{F}_{i0} + \mathbf{F}_{i1} + \mathbf{F}_{i2} + \cdots + \mathbf{F}_{in},$$

where $\mathbf{F}_{ij}$ is the force exerted on m_i by m_j and $\mathbf{F}_{i0}$ is the force due to some agency external to the n masses which form the system. Now by Newton's Third Law, when one particle acts upon a second, the second reacts upon the first with a force which is equal in magnitude and opposite in direction. Hence to $\mathbf{F}_{ij}$ above there will correspond a force $\mathbf{F}_{ji} = - \mathbf{F}_{ij}$ exerted by m_i on m_j. In the sum $\Sigma\mathbf{F}_i$ all these equal and opposite actions and reactions will drop out and $\Sigma\mathbf{F}_i$ may be replaced by $\Sigma\mathbf{F}_{i0}$, the sum of the external forces. Hence the important theorem that: *The motion of the center of mass of a set of particles is as if all the mass were concentrated there and all the external forces were applied there* (the internal forces, that is, the forces of mutual action and reaction between the particles being entirely neglected).

The *moment of a force* about a given point is defined as the product of the force by the perpendicular distance of the force from the point. If $\mathbf{r}$ is the vector from the point as origin to any point in the line of the force, the moment is therefore $\mathbf{r}\times\mathbf{F}$ when considered as a vector quantity, and is perpendicular to the plane of the line of the force and the origin. The equations of n moving masses may now be combined in a different way and reduced. Multiply the equations by $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n$ and add. Then

$$m_1\mathbf{r}_1\times\frac{d\mathbf{v}_1}{dt} + m_2\mathbf{r}_2\times\frac{d\mathbf{v}_2}{dt} + \cdots + m_n\mathbf{r}_n\times\frac{d\mathbf{v}_n}{dt} = \mathbf{r}_1\times\mathbf{F}_1 + \mathbf{r}_2\times\mathbf{F}_2 + \cdots + \mathbf{r}_n\times\mathbf{F}_n$$

or $$m_1\frac{d}{dt}\mathbf{r}_1\times\mathbf{v}_1 + m_2\frac{d}{dt}\mathbf{r}_2\times\mathbf{v}_2 + \cdots + m_n\frac{d}{dt}\mathbf{r}_n\times\mathbf{v}_n = \mathbf{r}_1\times\mathbf{F}_1 + \mathbf{r}_2\times\mathbf{F}_2 + \cdots + \mathbf{r}_n\times\mathbf{F}_n$$

or $$\frac{d}{dt}(m_1\mathbf{r}_1\times\mathbf{v}_1 + m_2\mathbf{r}_2\times\mathbf{v}_2 + \cdots + m_n\mathbf{r}_n\times\mathbf{v}_n) = \Sigma\mathbf{r}\times\mathbf{F}. \tag{56}$$

This equation shows that if the areal velocities of the different masses are multiplied by those masses, and all added together, the derivative of the sum obtained is equal to the moment of all the forces about the origin, the moments of the different forces being added as vector quantities.

This result may be simplified and put in a different form. Consider again the resolution of $\mathbf{F}_i$ into the sum $\mathbf{F}_{i0} + \mathbf{F}_{i1} + \cdots + \mathbf{F}_{in}$, and in particular consider the action $\mathbf{F}_{ij}$ and the reaction $\mathbf{F}_{ji} = - \mathbf{F}_{ij}$ between two particles. Let it be assumed that the action and reaction are not only equal and opposite, but lie along the line connecting the two particles. Then the perpendicular distances from the origin to the action and reaction are equal and the moments of the action and reaction are equal and opposite, and may be dropped from the sum $\Sigma\mathbf{r}_i\times\mathbf{F}_i$, which then reduces to $\Sigma\mathbf{r}_i\times\mathbf{F}_{i0}$. On the other hand a term like $m_i\mathbf{r}_i\times\mathbf{v}_i$ may be written as $\mathbf{r}_i\times(m_i\mathbf{v}_i)$. This product is formed from the momentum in exactly the same way that the moment is formed from the force, and it is called the moment of momentum. Hence the equation (56) becomes

$$\frac{d}{dt}\text{(total moment of momentum)} = \text{moment of external forces.}$$

Hence the result that, as vector quantities: *The rate of change of the moment of momentum of a system of particles is equal to the moment of the external forces* (the forces between the masses being entirely neglected under the assumption that action and reaction lie along the line connecting the masses).

EXERCISES

1. Apply the definition of differentiation to prove

(α) $d(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot d\mathbf{v} + \mathbf{v}\cdot d\mathbf{u}$, ($\beta$) $d[\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})] = d\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(d\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{v}\times d\mathbf{w})$.

2. Differentiate under the assumption that vectors denoted by early letters of the alphabet are constant and those designated by the later letters are variable:

(α) $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})$, ($\beta$) $\mathbf{a}\cos t + \mathbf{b}\sin t$, ($\gamma$) $(\mathbf{u}\cdot\mathbf{u})\,\mathbf{u}$,

(δ) $\mathbf{u}\times\dfrac{d\mathbf{u}}{dx}$, ($\epsilon$) $\mathbf{u}\cdot\left(\dfrac{d\mathbf{u}}{dx}\times\dfrac{d^2\mathbf{u}}{dx^2}\right)$, ($\zeta$) $\mathbf{c}(\mathbf{a}\cdot\mathbf{u})$.

3. Apply the rules for change of variable to show that $\dfrac{d^2\mathbf{r}}{ds^2} = \dfrac{\mathbf{r}''s' - \mathbf{r}'s''}{s'^3}$, where accents denote differentiation with respect to x. In case $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ show that $1/\sqrt{\mathbf{C}\cdot\mathbf{C}}$ takes the usual form for the radius of curvature of a plane curve.

4. The equation of the helix is $r = \mathbf{i}a\cos\phi + \mathbf{j}a\sin\phi + \mathbf{k}b\phi$ with $s = \sqrt{a^2 + b^2}\,\phi$; show that the radius of curvature is $(a^2 + b^2)/a$.

5. Find the torsion of the helix. It is $b/(a^2 + b^2)$.

6. Change the variable from s to some other variable t in the formula for torsion.

7. In the following cases find the gradient either by applying the formula which contains the partial derivatives, or by using the relation $d\mathbf{r}\cdot\nabla F = dF$, or both:

(α) $\mathbf{r}\cdot\mathbf{r} = x^2 + y^2 + z^2$, ($\beta$) $\log r$, (γ) $r = \sqrt{\mathbf{r}\cdot\mathbf{r}}$,

(δ) $\log(x^2 + y^2) = \log[\mathbf{r}\cdot\mathbf{r} - (\mathbf{k}\cdot\mathbf{r})^2]$, ($\epsilon$) $(\mathbf{r}\times\mathbf{a})\cdot(\mathbf{r}\times\mathbf{b})$.

8. Prove these laws of operation with the symbol ∇:

(α) $\nabla(F + G) = \nabla F + \nabla G$, ($\beta$) $G^2\nabla(F/G) = G\nabla F - F\nabla G$.

9. If r, ϕ are polar coördinates in a plane and $\mathbf{r}_1$ is a unit vector along the radius vector, show that $d\mathbf{r}_1/dt = \mathbf{n}\,d\phi/dt$ where $\mathbf{n}$ is a unit vector perpendicular to the radius. Thus differentiate $\mathbf{r} = r\mathbf{r}_1$ twice and separate the result into components along the radius vector and perpendicular to it so that

$$f_r = \frac{d^2r}{dt^2} - r\left(\frac{d\phi}{dt}\right)^2, \qquad f_\phi = r\frac{d^2\phi}{dt^2} + 2\frac{d\phi}{dt}\frac{dr}{dt} = \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right).$$

10. Prove conversely to the text that if the vector rate of description of area is constant, the force must be central, that is, $\mathbf{r}\times\mathbf{F} = 0$.

11. Note that $\mathbf{r}\times\mathbf{v}\cdot\mathbf{i}$, $\mathbf{r}\times\mathbf{v}\cdot\mathbf{j}$, $\mathbf{r}\times\mathbf{v}\cdot\mathbf{k}$ are the projections of the areal velocities upon the planes $x = 0$, $y = 0$, $z = 0$. Hence derive (54′) of the text.

12. Show that the Cartesian expressions for the magnitude of the velocity and of the acceleration and for the rate of change of the speed dv/dt are

$$v = \sqrt{x'^2 + y'^2 + z'^2}, \quad f = \sqrt{x''^2 + y''^2 + z''^2}, \quad v' = \frac{x'x'' + y'y'' + z'z''}{\sqrt{x'^2 + y'^2 + z'^2}},$$

where accents denote differentiation with respect to the time.

13. Suppose that a body which is rigid is rotating about an axis with the angular velocity $\omega = d\phi/dt$. Represent the angular velocity by a vector $\mathbf{a}$ drawn along the axis and of magnitude equal to ω. Show that the velocity of any point in space is $\mathbf{v} = \mathbf{a}\times\mathbf{r}$, where $\mathbf{r}$ is the vector drawn to that point from any point of the axis as origin. Show that the acceleration of the point determined by $\mathbf{r}$ is in a plane through the point and perpendicular to the axis, and that the components are

$$\mathbf{a}\times(\mathbf{a}\times\mathbf{r}) = (\mathbf{a}\cdot\mathbf{r})\,\mathbf{a} - \omega^2\mathbf{r} \text{ toward the axis,} \quad (d\mathbf{a}/dt)\times\mathbf{r} \text{ perpendicular to the axis,}$$

under the assumption that the axis of rotation is invariable.

14. Let $\bar{\mathbf{r}}$ denote the center of gravity of a system of particles and $\mathbf{r}_i'$ denote the vector drawn from the center of gravity to the ith particle so that $\mathbf{r}_i = \bar{\mathbf{r}} + \mathbf{r}_i'$ and $\mathbf{v}_i = \bar{\mathbf{v}} + \mathbf{v}_i'$. The kinetic energy of the ith particle is by definition

$$\tfrac{1}{2} m_i v_i^2 = \tfrac{1}{2} m_i \mathbf{v}_i\cdot\mathbf{v}_i = \tfrac{1}{2} m_i (\bar{\mathbf{v}} + \mathbf{v}_i')\cdot(\bar{\mathbf{v}} + \mathbf{v}_i').$$

Sum up for all particles and simplify by using the fact $\Sigma m_i \mathbf{r}_i' = 0$, which is due to the assumption that the origin for the vectors $\mathbf{r}_i'$ is at the center of gravity. Hence prove the important theorem: *The total kinetic energy of a system is equal to the kinetic energy which the total mass would have if moving with the center of gravity plus the energy computed from the motion relative to the center of gravity as origin*, that is,

$$T = \tfrac{1}{2}\Sigma m_i v_i^2 = \tfrac{1}{2} M\bar{v}^2 + \tfrac{1}{2}\Sigma m_i v_i'^2.$$

15. Consider a rigid body moving in a plane, which may be taken as the xy-plane. Let any point $\mathbf{r}_0$ of the body be marked and other points be denoted relative to it by $\mathbf{r}'$. The motion of any point $\mathbf{r}'$ is compounded from the motion of $\mathbf{r}_0$ and from the angular velocity $\mathbf{a} = \mathbf{k}\omega$ of the body about the point $\mathbf{r}_0$. In fact the velocity $\mathbf{v}$ of any point is $\mathbf{v} = \mathbf{v}_0 + \mathbf{a}\times\mathbf{r}$. Show that the velocity of the point denoted by $\mathbf{r}' = \mathbf{k}\times\mathbf{v}_0/\omega$ is zero. This point is known as the instantaneous center of rotation (§ 39). Show that the coördinates of the instantaneous center referred to axes at the origin of the vectors $\mathbf{r}$ are

$$x = \mathbf{r}\cdot\mathbf{i} = x_0 - \frac{1}{\omega}\frac{dy_0}{dt}, \qquad y = \mathbf{r}\cdot\mathbf{j} = y_0 + \frac{1}{\omega}\frac{dx_0}{dt}.$$

16. If several forces $\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_n$ act on a body, the sum $\mathbf{R} = \Sigma\mathbf{F}_i$ is called the *resultant* and the sum $\Sigma\mathbf{r}_i\times\mathbf{F}_i$, where $\mathbf{r}_i$ is drawn from an origin O to a point in the line of the force $\mathbf{F}_i$, is called the *resultant moment* about O. Show that the resultant moments $\mathbf{M}_O$ and $\mathbf{M}_{O'}$ about two points are connected by the relation $\mathbf{M}_{O'} = \mathbf{M}_O + \mathbf{M}_{O'}(\mathbf{R}_O)$, where $\mathbf{M}_{O'}(\mathbf{R}_O)$ means the moment about O' of the resultant $\mathbf{R}$ considered as applied at O. Infer that moments about all points of any line parallel to the resultant are equal. Show that in any plane perpendicular to $\mathbf{R}$ there is a point O' given by $\mathbf{r} = \mathbf{R}\times\mathbf{M}_O/\mathbf{R}\cdot\mathbf{R}$, where O is any point of the plane, such that $\mathbf{M}_{O'}$ is parallel to $\mathbf{R}$.