THE PENTAGON

Volume XXVI Spring, 1967 Number 2

CONTENTS

National Officers

Kappa Mu Epsilon, national honorary mathematics society, was founded in 1931. The object of the fraternity is fourfold: to further the interests of mathematics in those schools which place their primary emphasis on the undergraduate program; to help the undergraduate realize the important role that mathematics has played in the development of western civilization; to develop an appreciation of the power and beauty possessed by mathematics, due, mainly, to its demands for logical and rigorous modes of thought; and to provide a society for the recognition of outstanding achievements in the study of mathematics at the undergraduate level. The official journal, THE PENTAGON, is designed to assist in achieving these objectives as well as to aid in establishing fraternal ties between the chapters.

For A Product of Matrices What Result Do You Want?*

GARY L. EERKES
Western Washington State College
and
F. MAX STEIN
Colorado State University

1. **Introduction.** Matrices have long played an important role in mathematics, and over the years a number of matrix operations have been developed. The student in a course in Matrix Theory or Linear Algebra, however, is seldom given the opportunity to inquire as to whether these operations are the only possible ones. For example, the definition for the product of matrices is normally given as being predetermined without any discussion of other possibilities or of the desired results. To show how other definitions may arise and why one definition is more desirable than another, we shall construct various possible definitions for the product and give a bit of the motivation involved.

2. **Preliminary Definitions.** Throughout this paper we will be dealing with the set S of square matrices of order two whose elements are *scalars* from F, the complex number field unless restricted otherwise. Over this set we shall define several different products. Each definition will be viewed in the light of the algebraic structure it gives us. That is, we shall examine the set for each product to see whether the set forms a group, field, linear space, etc., under the defined operation. For future reference we shall list the requirements that must be met for various algebraic structures.

A nonempty set of elements G is said to form a *group* if there is defined a binary operation $\oplus$ between elements of G such that:

(1) $a, b \in G$ implies that $a \oplus b \in G$, (closure property).

(2) $a, b, c \in G$ implies that $a \oplus (b \oplus c) = (a \oplus b) \oplus c$, (associative law).

(3) There exists an element $e \in G$ such that $a \oplus e = e \oplus a = a$ for all $a \in G$, (identity element).

*Prepared in a National Science Foundation Undergraduate Science Education Program in Mathematics at Colorado State University under the direction of Professor Stein.

(4) For every $a \varepsilon G$ there exists an element $a^{-1} \varepsilon G$ such that $a \oplus a^{-1} = a^{-1} \oplus a = e$, (inverses).

If in addition to properties (1) through (4) it is true that

(5) for every $a, b \varepsilon G$, $a \oplus b = b \oplus a$, (commutative law), then the group G is said to form an Abelian (or commutative) group.

For a *field* we require two binary operations, called addition (+) and multiplication (• or juxtaposition), between elements of the set F such that the elements form an Abelian group under (+) and all elements except the additive identity form an Abelian group under (•), and further that

(6) for all $a, b, c \varepsilon F$, $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$, (the two distributive laws).

For a linear space (or vector space) an operation is defined which uses elements from two sets, as can be seen by the following definition. A set V is a *linear space* over a field F if V is an Abelian group under an operation (+) such that for every $\gamma, \lambda \varepsilon F$, and every $v, w \varepsilon V$ a product represented by the juxtaposition of symbols is such that the following properties hold:

(7) $\gamma v = v\gamma \varepsilon V$, (closure property).

(8) $\gamma(v + w) = \gamma v + \gamma w$, (one distributive property).

(9) $(\gamma + \lambda)v = \gamma v + \lambda v$, (another distributive property).

(10) $\gamma(\lambda v) = (\gamma\lambda)v$, (associative property).

(11) $1v = v$, (identity property).

Here 1 represents the multiplicative identity of F.

The requirements that a set A be a *linear algebra* over a field F are that A be a linear space over F that satisfies the following additional properties:

(12) $ab = c$, (a closure property).

(13) $a(bc) = (ab)c$, (an associative property).

(14) $a(\gamma b + \lambda c) = \gamma(ab) + \lambda(ac)$, (a distributive property).

(15) $(\gamma b + \lambda c)a = \gamma(ba) + \lambda(ca)$, (another distributive property).

Note that if (14) and (15) hold then (6) holds as well.

In the following we shall use the standard definition for matrices of element-wise addition and the conventional definition for the product of a scalar and a matrix. That is, the operation $\oplus$ in (1) through (4) is (+), the usual matrix addition.

3. **Definitions of Multiplication.** We now consider various definitions for the product of matrices of order two, the elements of S. Each definition will be viewed in light of the algebraic structure of S under the defined product.

DEFINITION I.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1b_1 + a_2b_3 & a_1b_2 + a_2b_4 \\ a_3b_1 + a_4b_3 & a_3b_2 + a_4b_4 \end{bmatrix}$$

This definition is obviously the standard definition for the multiplication of matrices. It is well-known that our set of matrices forms an algebra over the scalars for this definition. However, our set S falls short of a group structure for this definition of multiplication since property (4), requiring an inverse for each element, is not met unless our matrices are greatly restricted.

DEFINITION II.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_2b_2 \\ a_3b_3 & a_4b_4 \end{bmatrix}$$

For this definition it readily follows that properties (1) through (15) are satisfied, and the set under this definition of multiplication satisfies all of the properties required for an algebra.

In attempting to arrive at a group under this operation, however, we are faced with complications. The identity, both right and left, is

$$I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

but no matrix with one or more of its scalar elements equal to zero has an inverse. If we desire a group we must restrict our matrices to arrays for which none of the scalar elements is zero. We shall refer to this restricted set of matrices as S^*. With this restriction we have that

$$\begin{bmatrix} 1/a_1 & 1/a_2 \\ 1/a_3 & 1/a_4 \end{bmatrix}$$

is the inverse of

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad a_1a_2a_3a_4 \neq 0.$$

Therefore, we find that for this definition all elements of S^* satisfy the properties required for a group. It is evident that this group is Abelian, and it is a simple matter to show that (6), the two distributive laws, holds as well. However, we cannot conclude that our system is a field since S^* is not closed under addition. We find that for a field structure we must further restrict our matrices to arrays for which the scalar elements are equal.

DEFINITION III.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} =$$

$$\begin{bmatrix} a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4 & a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3 \\ a_1b_3 + a_3b_1 - a_2b_4 + a_4b_2 & a_1b_4 + a_4b_1 + a_2b_3 - a_3b_2 \end{bmatrix}$$

This definition seems more complicated than it actually is, as will be seen in the next section. It is evident that under this definition properties (1) through (12) hold, and it can be shown that (13) holds also. Properties (14) and (15) follow quite readily, so S forms an algebra over the complex numbers.

Concerning a group structure under this product we have properties (12) and (13) already. Next we observe that the identity, right and left, is

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

If we attempt to find an inverse, we find that

$$\begin{bmatrix} \dfrac{a_1}{a_1^2 + a_2^2 + a_3^2 + a_4^2} & \dfrac{-a_2}{a_1^2 + a_2^2 + a_3^2 + a_4^2} \\ \dfrac{-a_3}{a_1^2 + a_2^2 + a_3^2 + a_4^2} & \dfrac{-a_4}{a_1^2 + a_2^2 + a_3^2 + a_4^2} \end{bmatrix}$$

is the inverse of

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \text{ if } a_1^2 + a_2^2 + a_3^2 + a_4^2 \neq 0.$$

Since $a_1^2 + a_2^2 + a_3^2 + a_4^2 = 0$ has solutions in the complex field, other than the obvious zero solutions, we must restrict our set S if it is to be a group under this definition. A natural restriction is to consider our matrices only over the real number field. In this case every non-zero matrix (the zero matrix being the additive identity) has an inverse. Note that this is the exact condition needed for a field structure. We shall denote this set of matrices over the real number field as S_R.

We have shown that our set of non-zero matrices in S_R forms a group under multiplication, but thus far we have not considered the commutativity of this product. Note, for example, that

$$\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ -4 & -1 \end{bmatrix},$$

whereas

$$\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 4 & 1 \end{bmatrix}.$$

Thus, this product is not commutative, and our group is non-Abelian. With this development it is evident that we will be unable to obtain a field structure unless we place many more restrictions on its possible members. Apparently we are asking for too much in attempting to arrive at a field in this case.

DEFINITION IV.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} a_2 & a_3 & a_4 \\ b_2 & b_3 & b_4 \\ c_2 & c_3 & c_4 \end{vmatrix} & - \begin{vmatrix} a_1 & a_3 & a_4 \\ b_1 & b_3 & b_4 \\ c_1 & c_3 & c_4 \end{vmatrix} \\ \begin{vmatrix} a_1 & a_2 & a_4 \\ b_1 & b_2 & b_4 \\ c_1 & c_2 & c_4 \end{vmatrix} & - \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \end{bmatrix}$$

Here we have a ternary product rather than a binary operation. (One should note that no associative property is involved in this definition.) Since our algebraic structures require a binary operation we do not consider this case further at this time, except to note that the closure property holds.

DEFINITION V.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4$$

It is obvious that we can also dispose of this case immediately if we desire an algebraic structure since (12), the closure property, is not satisfied.

4. **A Second Look at the Products.** We shall now examine each of the above definitions from another point of view. In Definition I we have our most familiar product. We observed in Section 3 that algebraically this product showed little promise, especially when compared to Definition III. Nevertheless, of all the definitions this is by far the most useful.

Definition II is probably the most natural of all the definitions. However, we observed that when this definition is examined in relation to a group structure that our set of matrices is restricted to such an extent that the product seems to be of little value. But rather than drop the product or restrict the set, we see that valuable insights can be gained if we view our matrices geometrically under this product.

For example, if we consider only those matrices in S_R for which the third and fourth scalar elements are zero we have a means of representing points in two-space. That is, if we consider all matrices of the type

$$\begin{bmatrix} a_1 & a_2 \\ 0 & 0 \end{bmatrix}, \quad a_1, a_2 \; \varepsilon \text{ Reals,}$$

we can pair each such matrix with the ordered pair (a_1, a_2) which we consider as coordinates of a point in Euclidean two-space E_2. However, since these matrices do not have multiplicative inverses because of the zero elements we shall consider instead the isomorphic set of one-by-two matrices over the real field under the same element-wise multiplication. Taking the inverse of a point in E_2 reduces to the problem of inverting the point about two *inversion-lines*, $x = 1$ or $x = -1$ and $y = 1$ or $y = -1$. With respect to those matrices in S_R in this case which do not have inverses we see that geometrically they reduce to two *zero-lines* (see [1]), $x = 0$ and $y = 0$. The inverse of $[3 \; -\frac{1}{2}]$, say, is then $[\frac{1}{3} \; -2]$, see Fig. 1.

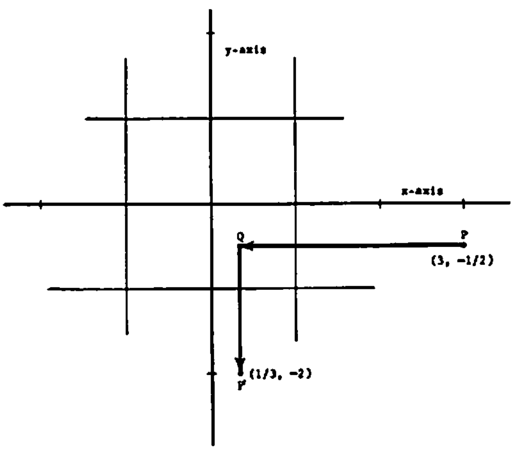


Fig. 1

In short, our system $(S_R, +, \cdot)$ is a *near-field* (see [1]), since properties (1), (2), (3), (5) and (6), are satisfied for all of S_R under multiplication, and property (4) is satisfied for all elements in S_R except a known set of points, the two zero-lines.

Similarly, if we consider the set of all real one-by-three matrices we can pair each such matrix $[a_1 a_2 a_3]$ with the ordered triple (a_1, a_2, a_3) and its corresponding point in three-space E_3. In this case, again considering element-wise multiplication, the problem of finding the inverse of a matrix reduces to taking the inversion of the corresponding point in three-space about three *inversion-planes*, $x = \pm 1$, $y = \pm 1$, and $z = \pm 1$, the signs again corresponding to the signs of the scalar elements a_1, a_2, and a_3 respectively. Those matrices which do not have inverses are represented by three planes $x = 0$, $y = 0$, and $z = 0$, which we denote as *zero-planes*. The inverse of [2 5/2 9/4], for example, is [1/2 2/5 4/9], see Fig. 2.

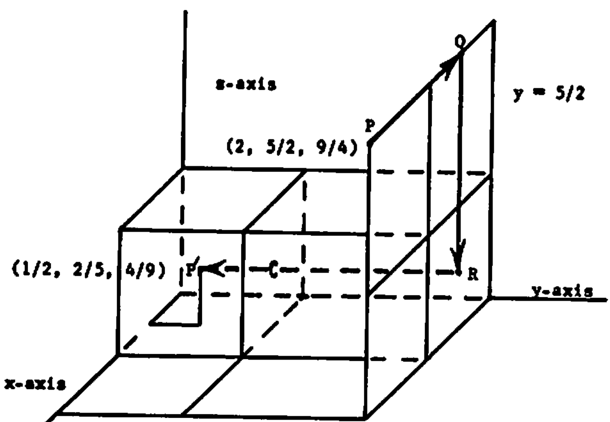


Fig. 2

Finally, to carry the analogy to our set S_R of square matrices we must expand our geometrical space to four dimensions. In this case, we can readily follow the same line of reasoning as in three-space with the zero-planes being replaced by the four *zero-hyperplanes* (see [2]), $x = 0$, $y = 0$, $z = 0$, $w = 0$, and the inversion-planes being replaced by the *inversion-hyperplanes* $x = \pm 1$, $y = \pm 1$, $z = \pm 1$, and $w = \pm 1$.

Along the same line of thought, if we now consider those matrices which form a field under this definition (recalling that that they are exactly those matrices in S_R for which all the scalar elements are equal), we observe that they are represented by the line in four-space, $x = y = z = w$, a field isomorphic to the real numbers themselves.

In Definition III we have a definition that is interesting algebraically in that it satisfies every property except the commutative property without any restrictions on the matrices other than that they be real. But other than that, this definition appears to have little motivation initially. Upon closer observation we see that our set S_R of real matrices under this product is isomorphic to Hamilton's quaternions under multiplication. Algebraically this definition is quite productive but it is not as interesting geometrically as Definition II; certainly it is not as useful as Definition I.

Definition IV is unique in that it is a ternary operation and consequently does not appear to lead anywhere algebraically. However, if we take the ternary product and employ a particular matrix in the product an interesting development arises. For example, consider

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_2b_3 - a_3b_2 & a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 & 0 \end{bmatrix}.$$

This result suggests the cross product of the two vectors (a_1, a_2, a_3), (b_1, b_2, b_3) in three-space. Likewise,

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & a_3c_4 - a_4c_3 \\ a_4c_2 - a_2c_4 & a_2c_3 - a_3c_2 \end{bmatrix}$$

suggests the cross product of (a_2, a_3, a_4) and (c_2, c_3, c_4). In short, we find that when we replace any one of the three matrices in our ternary product by any one of the following:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

and consider the other two matrices as vectors in four-space, that this product projects these two vectors onto a hyperplane and gives the cross product of them in this three-space (actually plus or minus the cross product, depending on the matrix replaced and the particular matrix replacing it). That is, we have a product that is highly suggestive of a relationship to the cross product of vectors in three-space or to a generalized cross product. Real matrices under this product are isomorphic to the vectors in four-space under the ternary product as defined by Williams and Stein [3].

In this light we see that if we interpret our matrices under this product as vectors in four-space we have that the product $A\ B\ C$ can be described as an operation that yields a vector (matrix) *orthogonal* to the three vectors (matrices) A, B, and C in the product. That is, if we use A, say, then $A(A\ B\ C)$ is the zero matrix if we use Definition II, or the result is the scalar O if we use Definition V. The result in either case is analogous to the cross product in three-space where $A \times B$ results in a vector perpendicular to both A and B. The magnitude of the resultant vector in four-space is, according to Williams and Stein [3], equal to the volume of a portion of a hypersurface with sides OA, OB, and OC. Again, we see that this is a generalization of the fact that the cross product $A \times B$ in three-space produces a vector whose magnitude is equal to the area of a parallelogram with sides OA and OB.

Continuing along this line, if we now consider Definition V with Definition IV in the following manner

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \left(\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix} \right)$$

we have

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{vmatrix},$$

which is the quadruple scalar product in four-space (see [3]), analogous to the triple scalar product in three-space. One would correctly suspect that this product yields a scalar whose absolute value is the volume of a hypersolid corresponding to the parallelepiped in three-space. Note that although Definition V does not lead to any of the algebraic structure that apparently would be desirable, since we do not even have the closure property, the definition can be interpreted as the inner product of vectors in four-space, a very valuable concept.

5. **Conclusion.** We have taken a fresh look at various products of matrices. In the process we have found, among other things, that the definition for the product is not always a predetermined one. The algebraist may well view matrices in an entirely different light from the analysist or the geometrist and as a consequence might employ a different product. So before you choose your definition you should ask, for a product of matrices what result do you want?

REFERENCES

1. J. Dorman, J. L. Tollefson, and F. Max Stein, "Fields and Near-Fields of Ordered Pairs of Reals," *The Mathematics Teacher,* 59 (1966), pp. 335-341.
2. H. P. Manning, *Geometry of Four Dimensions,* Macmillan, New York, 1914.
3. M. Z. Williams and F. Max Stein, "A Triple Product of Vectors in Four-Space," *Mathematics Magazine,* 37 (1964), pp. 230-235.

Mathematics is the science which draws necessary conclusions.

—Benjamin Pierce

Least Squares

DERALD WALLING
Faculty, Texas Technological College

Mood [2, page 309] points out that "there is a general problem of curve fitting which is entirely unrelated to normal regression theory but which may be solved by formulas identical with those . . . obtained for estimating regression coefficients."

In various courses, the subject of curve fitting or the subject of regression arises. We want to take this opportunity to outline simply the difference between the two problems.

In order to illustrate the difference, we want to consider the problem where we are given a set of points (x_i, y_i), $i = 1, 2, \cdots, N$; and where we suppose we want to fit

$$(1) \qquad y = a + bx$$

to these points.

Suppose we consider the set of N points as a sample set. Assume that for any given x, y is a normally distributed variate with mean $a + bx$ and variance σ^2 independent of x. Then, the density of y is

$$(2) \qquad f(y;a,b,\sigma^2,x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[y - (a + bx)]^2}{2\sigma^2}}.$$

Thus, we have the one-parameter family of normal distributions for which a, b, and σ^2 are fixed.

The problem is to determine the unknown parameters a, b and σ^2. The method used to find a, b, and σ^2 is called the method of maximum likelihood. The likelihood of y is

$$(3) \qquad L = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N e^{-\frac{\sum_{i=1}^{N}[y_i - a - bx_i]^2}{2\sigma^2}}.$$

and thus

$$\text{(4)} \qquad \ln L = -\frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{N} [y_i - a - bx_i]^2.$$

On putting the derivatives of ln L with respect to σ^2, a, and b equal to zero, we obtain the equations

$$\text{(5)} \qquad N\sigma^2 = \sum_{i=1}^{N} [y_i - a - bx_i]^2,$$

$$\text{(6)} \qquad \sum_{i=1}^{N} [y_i - a - bx_i], = 0,$$

$$\text{(7)} \qquad \sum_{i=1}^{N} x_i[y_i - a - bx_i] = 0.$$

The last two above equations, equations (6) and (7), are called the normal equations. We can solve the normal equations to find a and b:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_1^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{N} y_1 \\ \sum_{i=1}^{N} x_i y_i \end{bmatrix}$$

Once we have a and b, we can compute σ^2 where

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} [y_i - a - bx_i]^2.$$

Suppose we consider the same basic problem; that is, the fitting of $y = a + bx$ to the set of N points (x_1,y_1), (x_2,y_2), $\cdots$, (x_N,y_N). Suppose, however, that no assumption is made concerning the distribution of y. Suppose we decide to fit $y = a + bx$ to these N points in the sense of least squares. To find a and b, we simply minimize

$$\text{(8)} \qquad I(a,b) = \sum_{i=1}^{N} [y_i - a - bx_i]^2$$

which leads to

$$\begin{bmatrix} a \\ \\ b \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} x_i \\ \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \\ \sum_{i=1}^{N} x_i y_i \end{bmatrix}$$

Now, in this second case, the only reason that we used least squares is that it leads us to a system of two linear equations in two unknowns. Least cubes, least distance, least absolute values, etc., would not have led us to such a nice system of equations for a and b.

Once we have a and b, we can easily compute an estimate of the error by finding

$$I(a,b) = \sum_{i=1}^{N} [y_i - a - bx_i]^2.$$

We observe that in both methods we are led to the same set of two linear equations in two unknowns. Thus, we could (and it is often done) refer to both problems as least squares and approach the first problem by the method of least squares. However, we must remember that there is a difference in the problems depending on the initial assumptions.

Hildebrand [1, page 264] notes this difference in the problems by pointing out that the first problem is handled under the assumption that the "true" function is such that the residuals at each of the N points can be reduced to zero, but that the impossibility of achieving this end in the case at hand is due to the presence of independent random errors in the several observed values. He points out also that under this assumption, certain additional statistical information about the calculated coefficients can also be found.

REFERENCES

1. Hildebrand, F. B., *Introduction to Numerical Analysis*, New York, New York, McGraw-Hill Book Company, Inc., 1956.
2. Mood, A. M., *Introduction to the Theory of Statistics*. New York, New York, McGraw-Hill Book Company, Inc., 1950.

Linear Involutions

SHANNON SMYRL
Student, Texas Technological College

This paper attempts to study some of the properties of a linear involution on a vector space over a field.

We first study the matrix and some geometric properties of a linear involution on a two-dimensional real space. Then we generalize the results. We would also like to point out how geometric observations suggest theorems in a vector space.

1. **Definitions and Notations:** Linear transformations are denoted by capital letters $A, B, \cdots$. The identity transformation is denoted by I. Vectors are denoted by $\alpha, \beta, \gamma, \cdots$. Scalars are denoted by small letters $x, y, z, \cdots$. We use row and column matrices for vectors [1]. If **M** and **N** are two complementary subspaces of a vector space **V** over a field F, then $\mathbf{M} \oplus \mathbf{N}$ is the direct sum of these subspaces [2]. The zero vector will be denoted by $\vec{0}$.

2. **Involutions on a Euclidean Plane:** Let A be a linear transformation on a real two-dimensional Euclidean space such that $A^2 = I$. Let the matrix of A with respect to a rectangular co-ordinate system be

$$[A] = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then

$$\begin{pmatrix} a^2 + bc & ab + bd \\ ac + cd & cb + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This is equivalent to the set of equations

$$(1) \qquad \begin{aligned} a^2 + bc &= 1 \\ ab + bd &= 0 \\ ac + cd &= 0 \\ cb + d^2 &= 1\,. \end{aligned}$$

To solve this set we consider different cases.

I. If $b \neq 0$, $c \neq 0$, then

$$A = \begin{pmatrix} a & b \\ \frac{1 - a^2}{b} & -a \end{pmatrix}.$$

II. If $b = 0$, then we have the set of equations

$$\begin{aligned} a^2 &= 1 \\ ac + cd &= 0 \\ d^2 &= 1, \end{aligned} \tag{2}$$

which implies $a = \pm 1$, $c(a + d) = 0$, $d = \pm 1$.

In this case the matrix of an involution may be any of

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ c & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} -1 & 0 \\ c & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

where $c \neq 0$.

The characteristic equation of a linear involution on a two dimensional real space, in all cases, is $m^2 - 1 = 0$, except for I and $-I$.

Thus the proper values are $m = 1$, and $m = -1$.

Let us consider only

$$A = \begin{pmatrix} a & b \\ \frac{1 - a^2}{b} & -a \end{pmatrix}, \; b \neq 0, \; c \neq 0.$$

Then the characteristic vectors of A corresponding to 1 and -1 are respectively $\alpha = (a - 1, b)$ and $\beta = (a + 1, b)$. We observe that these vectors are respectively on the lines $bx - (a - 1)y = 0$ and $bx - (a + 1)y = 0$. We respectively call these lines L_1 and L_2 (Fig. 1). Note that $\{\alpha, \beta\}$ is linearly independent since

$$\begin{vmatrix} a - 1 & b \\ a + 1 & b \end{vmatrix} = -2b \neq 0.$$

For any vector γ in the plane we have

$$\gamma = x\alpha + y\beta.$$

Thus $A\gamma = xA\alpha + yA\beta = x\alpha - y\beta$. (Fig. 2).

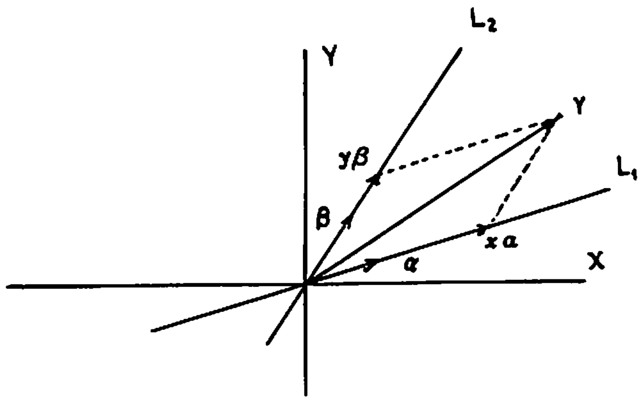


Fig. 1

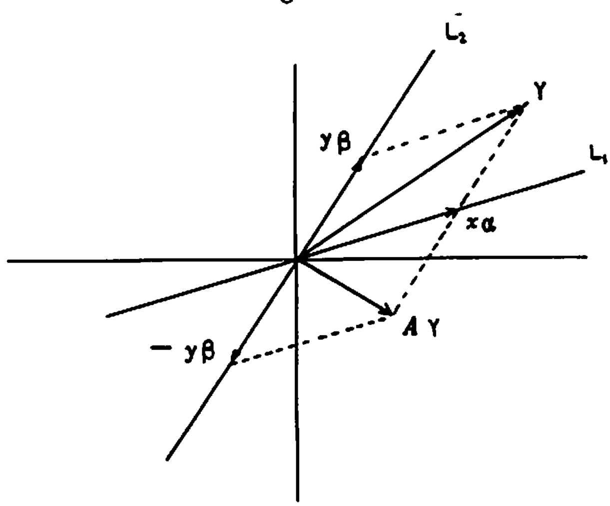


Fig. 2

We observe that $x\alpha = P\alpha$, where P is the projection on the line L_1 along L_2 and $y\beta = Q\beta$, where Q is the projection on L_2 along L_1. Thus

$$A\gamma = P\gamma - Q\gamma = (P - Q)\gamma.$$

This shows that a linear involution is the difference of two projections.

3. **Generalization to a Vector Space:** In what follows we exclude fields of characteristic two. Let **V** be a vector space over a field F. Consider A to be a linear involution on **V**. That is A is a linear transformation on **V** such that $A^2 = I$.

It is clear that I and $-I$ are linear involutions. We shall study cases where $A \neq I$ and $A \neq -I$. Let m be a proper value of A and α the corresponding proper vector, i.e. $A\alpha = m\alpha$, $\alpha \neq \vec{0}$. Then $\alpha = A^2\alpha = mA\alpha = m^2\alpha$ which implies $m = \pm 1$.

THEOREM 1. Let $\mathbf{K} = \{ \xi \mid A\xi = \xi \}$ and $\mathbf{L} = \{ \xi \mid A\xi = -\xi \}$ be subspaces of **V** corresponding to proper value 1 and -1 respectively. Then $\mathbf{V} = \mathbf{K} \oplus \mathbf{L}$.

Proof. Let $\xi \in \mathbf{K} \cap \mathbf{L}$. Then $\xi \in \mathbf{K}$ and $\xi \in \mathbf{L}$. Now $\xi \in \mathbf{K}$ implies $A\xi = \xi$ and $\xi \in \mathbf{L}$ implies $A\xi = -\xi$. Hence $\xi = \vec{0}$. Thus $\mathbf{K} \cap \mathbf{L} = \{\vec{0}\}$.

Now suppose $\mathbf{V} = \mathbf{K} \oplus \mathbf{L} \oplus \mathbf{M}$, where $\mathbf{M} \neq \{0\}$. We shall show that $\xi \in \mathbf{M}$ implies $A\xi \in \mathbf{M}$. For if $\xi \in \mathbf{M}$ and $A\xi \in \mathbf{K} \oplus \mathbf{L} \oplus \mathbf{M}$, then $A\xi = \eta + \zeta + \delta$ where $\eta \in \mathbf{K}$, $\zeta \in \mathbf{L}$, and $\delta \in \mathbf{M}$.

$$\xi = A^2\xi = \eta - \zeta + A\delta \in \mathbf{M}$$

which implies $\eta = \vec{0}$, $\zeta = \vec{0}$, and $A\xi = A^2\delta \in \mathbf{M}$. Therefore, **M** is invariant under A.

Let B be equal to A restricted to **M**. Then B has a proper vector β. That is $A\beta = B\beta = \pm\beta$, $\beta \neq \vec{0}$, $\beta \in \mathbf{M}$. But $A\beta = \beta$ implies $\beta \in \mathbf{K}$ and $A\beta = -\beta$ implies $\beta \in \mathbf{L}$. This contradicts the fact that $\mathbf{K} \cap \mathbf{M} = \{\vec{0}\}$ and $\mathbf{L} \cap \mathbf{M} = \{\vec{0}\}$. Therefore $\mathbf{M} = \{\vec{0}\}$ which proves the theorem.

THEOREM 2. A linear transformation A on **V** is an involution if and only if $A = P - Q$, where P is the projection on a subspace **K** along a complementary subspace **L** and **Q** is the projection on **L** along **K**.

Proof. By theorem 1 we know $\mathbf{V} = \mathbf{K} \oplus \mathbf{L}$ where $\mathbf{K} = \{ \xi \mid A\xi = \xi \}$ and $\mathbf{L} = \{ \xi \mid A\xi = -\xi \}$. Any $\xi \in \mathbf{V}$ can be written as $\xi = \eta + \zeta$ where $\eta \in \mathbf{K}$ and $\zeta \in \mathbf{L}$.

$$A\xi = A(\eta + \zeta) = A\eta + A\zeta = \eta - \zeta.$$

We observe that η is the projection of ξ on the subspace **K** along **L** and that ζ is the projection of ξ on **L** along **K**. Let P be the projection on **K** along **L** and let Q be the projection on **L** along **K**. Then

$$A\xi = P\xi - Q\xi = (P - Q)\xi.$$

This implies $A = P - Q$.

Conversely let $A = P - Q$, where P is the projection on a subspace **K** along the subspace **L** and Q is the projection on **L** along **K**. Then $PQ = QP = 0$, and

$$A^2 = P^2 + Q^2 = P + Q = I.$$

4. **A Geometric Observation.** Let **K** be a fixed line through the origin and let **L** be a variable line through the origin distinct from **K**. Define $A = P - Q$ where P is the projection on **K** along **L** and Q is the projection on **L** along **K**. Let ξ be a vector in the plane (Fig. 3).

Then $A\xi = P\xi - Q\xi$. We observe that $Q\xi = \xi - P\xi$ for

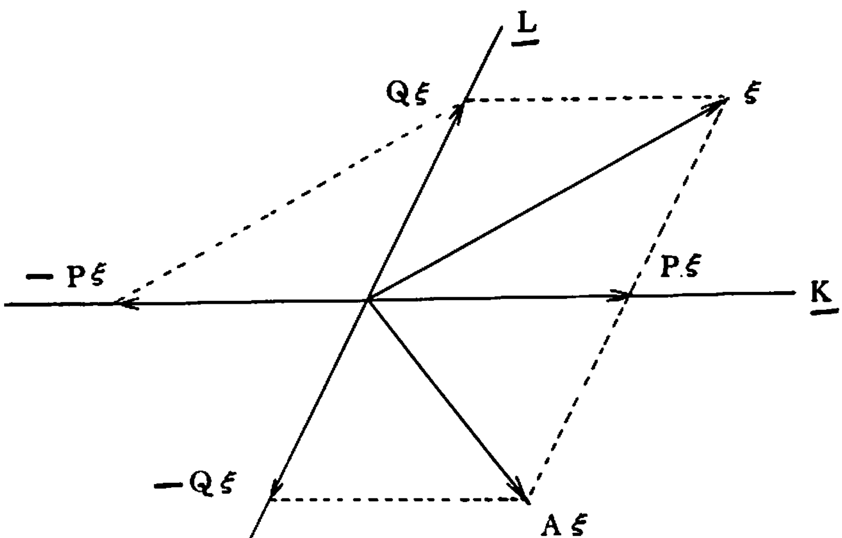


Fig. 3

every **L**. Define $\overline{X} = \{\eta \mid \eta = \xi + \zeta, \zeta \in \mathbf{K}\}$. Then $Q\xi \in \overline{X}$ and for every $\eta \in \overline{X}$, $A\eta \in -\overline{X}$. We define the transformation A such that $B\overline{X} = \{\zeta \mid \zeta = A\xi, \xi \in \overline{X}\}$. Then B is linear and for every $\xi \in \overline{X}$, $B^2\xi = P^2\xi + Q^2\xi = \xi$ which implies B is an involution.

5. Generalization to a Vector Space:

THEOREM 1. Let $\mathbf{K}$ be a subspace of a vector space $\mathbf{V}$ over a field F. Let $\{\overline{X}\}$ be the quotient space of $\mathbf{V}$ corresponding to $\mathbf{K}$ [2]. That is $\overline{X} = \{ \eta \mid \eta = \xi + \zeta, \xi \in \mathbf{V}, \zeta \in \mathbf{K} \}$. Then B is a linear involution $\{\overline{X}\}$ if and only if $B = \pm I$.

Proof. Let A be a linear transformation on $\mathbf{V}$ such that

$$\text{if } \xi \in \mathbf{K}, \text{ then } A\xi \in \mathbf{K};$$

$$\text{if } \xi \in \overline{X}, \text{ then } A\xi \in \pm\overline{X}.$$

This induces a linear transformation B on $\{\overline{X}\}$ by

$$B\overline{X} = \{ \zeta \mid \zeta = A\xi, \xi \in \overline{X}\}.$$

As was done in 4 we can easily show B is a linear involution on $\{\overline{X}\}$. We omit the proof since it is very similar to the one in 4. Conversely let B be a linear involution on $\{\overline{X}\}$, i.e.,

$$B^2\overline{X} = \overline{X}.$$

Let $\overline{K} = \{ \xi \mid \xi = \vec{0} + \zeta, \zeta \in \mathbf{K} \}$. Clearly $\overline{K}$ is the zero vector of $\{\overline{X}\}$. Since B is non-singular

$$B\overline{X} = \overline{K} \tag{1}$$

if and only if $\overline{X} = \overline{K}$. Now let $B\overline{X} = \overline{Y}$. Since $\overline{X}$ and $\overline{Y}$ are in $\{X\}$ either $\overline{X} = \overline{Y}$ which implies $B = I$ or $\overline{X} \cap \overline{Y} = \overline{K}$. Now suppose $\overline{X} \neq \overline{Y}$. Consider $\overline{X} + \overline{Y}$. We note that

$$B(\overline{X} + \overline{Y}) = \overline{Y} + \overline{X} = \overline{X} + \overline{Y}.$$

Then by (1) we have $\overline{X} + \overline{Y} = \overline{K}$. This implies $\overline{Y} = -\overline{X}$ and $B = -I$.

REFERENCES

1. Ali R. Amir-Moez, *Elements of Linear Spaces*. Edwards Brothers, Inc., Ann Arbor, Michigan (1961).
2. Paul R. Halmos, *Finite-Dimensional Vector Spaces*, D Van Nostrand Company, Inc., Princeton, New Jersey (1958).

A Generalization of the Multiplication of A Matrix By A Scalar*

BARRY R. WALLERSTEDT
Coe College
and
F. MAX STEIN
Colorado State University

1. Introduction. In the theory of matrices the usual definition of multiplication requires the matrices to be conformal; i.e. the number of columns in the first matrix must equal the number of rows in the second. Other definitions of multiplication do not require this condition to be met; one such definition is for the Kronecker product [2] and another is for the multiplication of a matrix by a scalar.

In the latter case the multiplication of an m by n matrix $\mathbf{A} = [a_{rs}]$ with elements, called **scalars**, from the real or complex number field $\mathcal{F}$ by any element d from $\mathcal{F}$ results in another m by n matrix $\mathbf{B} = [b_{rs}]$ with $b_{rs} = da_{rs}$. We note that d operates only on elements from $\mathcal{F}$. If we consider d as a 1 by 1 matrix we are led to a multiplicative operation between non-conformal matrices.

In this paper we shall define an operation (X), called X-multiplication, between certain non-conformal matrices in a manner analogous to that for scalar multiplication. Basically in X-multiplication we replace all scalar elements by square matrices of order k. After determining the structure created by X-multiplication, a few theorems similar to those for ordinary matrix multiplication are proved. In the final sections of the paper we specialize the discussion to the particular cases of scalar multiplication and ordinary multiplication of square matrices.

2. Notation. Square matrices are represented by bold faced lower case or capital Latin letters. In general the former denote matrices of smaller order than the latter. As a rule the lower case letters represent k by k matrices and the capital letters represent

*Prepared in a National Science Foundation Undergraduate Science Education Program in Mathematics at Colorado State University under the direction of Professor Stein.

n by n matrices, the elements being scalars in each case. Thus for X-multiplication we have matrices of order k, such as $\mathbf{y} = [y_{rs}]$, and matrices of order n that are partitioned, such as

$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{bmatrix} = [\alpha_{rs}],$$

where the elements of $\mathbf{A}$ are considered as matrices of order k. Note that the subscripts on α_{rs} indicate the position, not the order of the sub-matrix. Hence, X-multiplication implies that $n = mk$ for m and k positive integers. We shall consider the general case first, assuming that $k \neq 1$ and $m \neq 1$.

Diagonal matrices can be expressed as diag $(b_{11}, \cdots, b_{kk})_k$ and diag $(\alpha_{11}, \cdots, \alpha_{mm})_{mk}$, where the elements on the diagonal are enclosed in parentheses, and the final subscript denotes the order of the matrix.

Finally we let $S_k = \{\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \cdots, \mathbf{i}, \mathbf{o}\}$ be the set of k by k matrices and $S_n = \{\mathbf{A}, \mathbf{B}, \mathbf{X}, \cdots, \mathbf{I}, \mathbf{O}\}$ be the set of n by n matrices, where $n = mk$. The symbols $\mathbf{i}$, $\mathbf{o}$, $\mathbf{I}$, and $\mathbf{O}$ represent the multiplicative and additive identities in S_k and S_n respectively.

3. The structure of S_k or S_n over the scalars. We assume the usual definitions of equality, addition $(+)$, and ordinary multiplication $(\cdot)$ of square matrices. From these definitions either set of square matrices forms an **algebra** [3] over the set of scalars $\mathcal{F}$. For the purpose of reference we shall list briefly the requirements that the set S_k, say, forms an algebra over $\mathcal{F}$. Under the operation $(+)$ we have

(1) $\mathbf{x} + \mathbf{y} = \mathbf{z}$, closure;

(2) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$, associativity;

(3) $\mathbf{x} + \mathbf{o} = \mathbf{x}$, the existence of an additive identity;

(4) $\mathbf{x} + (-\mathbf{x}) = \mathbf{o}$, each element has an inverse;

(5) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$, commutativity.

These five properties render S_k a **commutative group** under $(+)$. Combining elements of S_k and $\mathcal{F}$ by juxtaposition we have

(6) $a(b\mathbf{x}) = (ab)\mathbf{x}$, a kind of associativity;
(7) $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$, one distributive property;
(8) $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$, another distributive property;
(9) $1\mathbf{x} = \mathbf{x}$, the existence of a multiplicative identity.

From properties (1) - (9) we have that S_k forms a **vector space** over $\mathcal{F}$. The ordinary multiplication (•) of elements of S_k gives us

(10) $\mathbf{x} \cdot \mathbf{y} = \mathbf{z}$, a closure property;
(11) $\mathbf{x} \cdot (\mathbf{y} \cdot \mathbf{z}) = (\mathbf{x} \cdot \mathbf{y}) \cdot \mathbf{z}$, an associative property;
(12) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$, a distributive property;
(13) $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$, another distributive property.

By properties (1) - (5) and (10) - (13) S_k forms a **ring**.

Finally the property,

(14) $(\mathbf{x} \cdot \mathbf{y})a = \mathbf{x} \cdot (\mathbf{y}a) = (\mathbf{x}a) \cdot \mathbf{y}$,

along with (1) - (13) yields S_k an **algebra** over $\mathcal{F}$.

4. The definition of X-multiplication. If $\mathbf{x} \in S_k$ and $\mathbf{A} \in S_n$, we define the operation (X), called X-multiplication, between elements of these two sets as

$$\mathbf{x} \times \mathbf{A} = \begin{bmatrix} \mathbf{x} \cdot \alpha_{11} & \cdots & \mathbf{x} \cdot \alpha_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{x} \cdot \alpha_{m1} & \cdots & \mathbf{x} \cdot \alpha_{mm} \end{bmatrix} \text{ and}$$

$$\mathbf{A} \times \mathbf{x} = \begin{bmatrix} \alpha_{11} \cdot \mathbf{x} & \cdots & \alpha_{1m} \cdot \mathbf{x} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_{m1} \cdot \mathbf{x} & \cdots & \alpha_{mm} \cdot \mathbf{x} \end{bmatrix}.$$

Thus X-multiplication from the left (right) gives a matrix in S_n which has matrices from S_k as elements; i.e. every k by k submatrix of $\mathbf{A}$ is multiplied on the left (right) by the matrix $\mathbf{x}$.

5. The structure of S_n over S_k. Our investigation of the structure of S_n over S_k involves the proof of statements analogous to properties (1) - (14) for the algebra of square matrices over the field $\mathcal{F}$.

It is obvious that properties (1) - (5) and (10) - (13) hold for addition and ordinary multiplication in S_n since these involve merely a substitution of matrices of order n for those of order k. We may also easily verify that properties analogous to (6) - (9) hold for X-multiplication between elements of S_k and S_n. For clarification and reference, we restate these properties in their appropriate form:

(6′) $\mathbf{x} \times (\mathbf{y} \times \mathbf{A}) = (\mathbf{x} \cdot \mathbf{y}) \times \mathbf{A}$ and $(\mathbf{A} \times \mathbf{x}) \times \mathbf{y} = \mathbf{A} \times (\mathbf{x} \cdot \mathbf{y})$,

(7′) $\mathbf{x} \times (\mathbf{A} + \mathbf{B}) = (\mathbf{x} \times \mathbf{A}) + (\mathbf{x} \times \mathbf{B})$ and $(\mathbf{A} + \mathbf{B}) \times \mathbf{x} = (\mathbf{A} \times \mathbf{x}) + (\mathbf{B} \times \mathbf{x})$,

(8′) $(\mathbf{x} + \mathbf{y}) \times \mathbf{A} = (\mathbf{x} \times \mathbf{A}) + (\mathbf{y} \times \mathbf{A})$ and $\mathbf{A} \times (\mathbf{x} + \mathbf{y}) = (\mathbf{A} \times \mathbf{x}) + (\mathbf{A} \times \mathbf{y})$,

(9′) $\mathbf{i} \times \mathbf{A} = \mathbf{A}$ and $\mathbf{A} \times \mathbf{i} = \mathbf{A}$.

As the proofs to these properties follow quite readily from the definition, we give only one of them as an example.

Proof of (6′).

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{A}) = \mathbf{x} \times \begin{bmatrix} \mathbf{y} \cdot \alpha_{11} & \cdots & \mathbf{y} \cdot \alpha_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{y} \cdot \alpha_{m1} & \cdots & \mathbf{y} \cdot \alpha_{mm} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{x} \cdot (\mathbf{y} \cdot \alpha_{11}) & \cdots & \mathbf{x} \cdot (\mathbf{y} \cdot \alpha_{1m}) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{x} \cdot (\mathbf{y} \cdot \alpha_{m1}) & \cdots & \mathbf{x} \cdot (\mathbf{y} \cdot \alpha_{mm}) \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{x} \cdot \mathbf{y}) \cdot \alpha_{11} & \cdots & (\mathbf{x} \cdot \mathbf{y}) \cdot \alpha_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (\mathbf{x} \cdot \mathbf{y}) \cdot \alpha_{m1} & \cdots & (\mathbf{x} \cdot \mathbf{y}) \cdot \alpha_{mm} \end{bmatrix} = (\mathbf{x} \cdot \mathbf{y}) \times \mathbf{A}$$

Thus by possessing properties (1) - (5) and (6′) - (9′) S_n forms a left or right **module** [3] (a structure over a ring smiliar to that of a vector space over a field) over S_k. We find, however, that a property analogous to (14) is not valid due to the lack of commutativity under multiplication between elements of S_k. Although S_n does not form an algebra over S_k, we do have the associative property

(15) $\mathbf{x} \times (\mathbf{A} \cdot \mathbf{B}) = (\mathbf{x} \times \mathbf{A}) \cdot \mathbf{B}$ and $(\mathbf{A} \cdot \mathbf{B}) \times \mathbf{x}$
$= \mathbf{A} \cdot (\mathbf{B} \times \mathbf{x})$.

Therefore we conclude that the structural properties of S_n over S_k, while not sufficient for an algebra, are beyond the requirements for a module and an associative ring.

6. The determinant. It is a common result of matrix theory that $\det(\mathbf{A} \cdot \mathbf{B}) = (\det \mathbf{A})(\det \mathbf{B})$, where **A** and **B** are conformal square matrices. Another well-known characteristic is $\det(d\mathbf{A}) = d^n(\det \mathbf{A})$, where **A** is a square matrix of order n and d is a scalar. Similarly for X-multiplication we have

THEOREM 1. $\operatorname{Det}(\mathbf{x} \times \mathbf{A}) = (\det \mathbf{x})^m(\det \mathbf{A})$ *for* $\mathbf{x} \in S_k$ *and* $\mathbf{A} \in S_n$.

Proof. Let $\mathbf{X} = \operatorname{diag}(\mathbf{x}, \cdots, \mathbf{x})_{mk}$, $\mathbf{A} = [\alpha_{rs}]$, and $\mathbf{I} = \operatorname{diag}(\mathbf{i}, \cdots, \mathbf{i})_{mk}$ all be elements of S_n. Then form the matrix

$$\begin{bmatrix} \mathbf{X} & \mathbf{O} \\ -\mathbf{I} & \mathbf{A} \end{bmatrix}_{2mk}$$

and we get by Laplace's expansion

$$\det \begin{bmatrix} \mathbf{X} & \mathbf{O} \\ -\mathbf{I} & \mathbf{A} \end{bmatrix} = (\det \mathbf{X})(\det \mathbf{A}) = (\det \mathbf{x})^m(\det \mathbf{A}).$$

To complete the proof we shall show that

$$\det \begin{bmatrix} \mathbf{X} & \mathbf{O} \\ -\mathbf{I} & \mathbf{A} \end{bmatrix} = \det(\mathbf{x} \times \mathbf{A}).$$

We recall that the value of a determinant is unaltered when to any row or column we add a constant multiple of any other row or column [1]. Thus in

$$\det \begin{bmatrix} \mathbf{X} & \mathbf{O} \\ -\mathbf{I} & \mathbf{A} \end{bmatrix}_{2mk} = \det \begin{bmatrix} \mathbf{x} & & & \mathbf{o} & & \\ & \ddots & & & \mathbf{o} & \\ \mathbf{o} & & \mathbf{x} & & & \\ -\mathbf{i} & & \mathbf{o} & \boldsymbol{\alpha}_{11} & \cdots & \boldsymbol{\alpha}_{1m} \\ & \ddots & & \vdots & & \vdots \\ \mathbf{o} & & -\mathbf{i} & \boldsymbol{\alpha}_{m1} & \cdots & \boldsymbol{\alpha}_{mm} \end{bmatrix}_{2mk}$$

we can consider the determinant on the right as one of order $2m$ whose elements are matrices of order k. To the $(m + r)$-th column of this matrix we add the following terms: (column 1) $\cdot\, \boldsymbol{\alpha}_{1r}$, (column 2) $\cdot\, \boldsymbol{\alpha}_{2r}$, $\cdots$, (column m) $\cdot\, \boldsymbol{\alpha}_{mr}$, for $r = 1, 2, \cdots, m$. The determinant is then transformed into

$$\det \begin{bmatrix} \mathbf{X} & \mathbf{x} \times \mathbf{A} \\ -\mathbf{I} & \mathbf{O} \end{bmatrix}_{2mk} = (-1)^{\alpha} \det(\mathbf{x} \times \mathbf{A})\det(-\mathbf{I}),$$

where the equality results from a Laplace expansion around the last n columns and the first n rows of the determinant. Thus our goal is reached provided the sign is correct. The sign of this expansion depends on α and $\det(-\mathbf{I})$. Since α is the sum of the first $2n$ integers, $\alpha = 2n(2n + 1)/2 = mk(2mk + 1)$. For $\det(-\mathbf{I})$ we find by repeating Laplace's expansion that $\det(-\mathbf{I}) = [\det(-\mathbf{i})]^m$. Since $-\mathbf{i} = \text{diag}\,(-1, \cdots, -1)_k$, we have that $[\det(-\mathbf{i})]^m = (-1)^{mk}$. Therefore the sign associated with the expansion is $\delta = (-1)^{\alpha}(-1)^{mk} = (-1)^{mk(2mk+2)}$ which is positive for any choice of m and k since the exponent is always even. Thus det $(\mathbf{x} \times \mathbf{A}) = (\det \mathbf{x})^m(\det \mathbf{A})$.

7. Multiplicative inverse elements. The multiplicative inverse of any nonsingular matrix $\alpha \varepsilon S_k$ is denoted by α^{-1}; it is such that $\alpha^{-1} \cdot \alpha = \alpha \cdot \alpha^{-1} = \mathbf{i}$. A similar characteristic holds for any nonsingular matrix $\mathbf{A} \varepsilon S_n$.

Suppose we let the matrix $\mathbf{e} \varepsilon S_k$ be such that

$$\mathbf{e} \times \mathbf{A} = \mathbf{I}, \text{ or } \begin{bmatrix} \mathbf{e} \cdot \alpha_{11} & \cdots & \mathbf{e} \cdot \alpha_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{e} \cdot \alpha_{m1} & \cdots & \mathbf{e} \cdot \alpha_{mm} \end{bmatrix} = \text{diag } (\mathbf{i}, \cdots, \mathbf{i})_{mk},$$

under X-multiplication for $\mathbf{A} \varepsilon S_n$. From this equation we get the equations

$$\mathbf{e} \cdot \alpha_{rs} = \mathbf{i}, \text{ when } r = s, \text{ and}$$

$$\mathbf{e} \cdot \alpha_{rs} = \mathbf{o}, \text{ when } r \neq s.$$

These imply that $\alpha_{rs} = \mathbf{o}$ for $r \neq s$, and all α_{rs} must be equal and nonsingular if $r = s$. That is, $\mathbf{A}$ must have the form $\mathbf{A} = \text{diag } (\alpha_{11}, \cdots, \alpha_{11})_{mk}$, where det $\alpha_{11} \neq 0$. Under these conditions it can be readily shown by the use of Laplace's expansion that $\mathbf{A}$ is nonsingular. Furthermore $\mathbf{e} = \alpha_{11}^{-1} = \cdots = \alpha_{mm}^{-1}$.

We now prove a theorem analogous to the well-known result that $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$, where $\mathbf{A}$ and $\mathbf{B}$ are conformal square matrices.

THEOREM 2. **Let $\mathbf{x} \varepsilon S_k$ and $\mathbf{A} \varepsilon S_n$ be nonsingular matrices.** Then $(\mathbf{x} \times \mathbf{A})^{-1} = \mathbf{A}^{-1} \times \mathbf{x}^{-1}$.

Proof. From Theorem 1 we have that if $\mathbf{x}$ and $\mathbf{A}$ are nonsingular then $(\mathbf{x} \times \mathbf{A})$ is nonsingular. Thus by the n by n matrix $(\mathbf{x} \times \mathbf{A})$ possess an inverse and hence we can write

$$(\mathbf{x} \times \mathbf{A})^{-1} \cdot (\mathbf{x} \times \mathbf{A}) = \mathbf{I}.$$

Thus $\quad (\mathbf{x} \times \mathbf{A})^{-1} \cdot (\mathbf{x} \times \mathbf{A}) \cdot \mathbf{A}^{-1} = \mathbf{I} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1},$

and

$$(\mathbf{x} \times \mathbf{A})^{-1} \cdot (\mathbf{x} \times \mathbf{I}) = (\mathbf{x} \times \mathbf{A})^{-1} \cdot \{\text{diag}\,(\mathbf{x}, \cdots, \mathbf{x})_{mk}\} = \mathbf{A}^{-1}.$$

Then $(\mathbf{x} \times \mathbf{A})^{-1} \cdot \{\text{diag}\,(\mathbf{x}, \cdots, \mathbf{x})_{mk}\} \cdot \{\text{diag}\,(\mathbf{x}^{-1}, \cdots, \mathbf{x}^{-1})_{mk}\}$

$$= \mathbf{A}^{-1} \cdot \{\text{diag}\,(\mathbf{x}^{-1}, \cdots, \mathbf{x}^{-1})_{mk}\},$$

or

$$(\mathbf{x} \times \mathbf{A})^{-1} \cdot \mathbf{I} = \mathbf{A}^{-1} \cdot (\mathbf{I} \times \mathbf{x}^{-1}),$$

and

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

In a similar manner we can show that $(\mathbf{A} \times \mathbf{x})^{-1} = \mathbf{x}^{-1} \times \mathbf{A}^{-1}$.

8. The transpose. We now wish to prove a theorem analogous to $(\mathbf{A} \cdot \mathbf{B})' = \mathbf{B}' \cdot \mathbf{A}'$, where $\mathbf{A}$ and $\mathbf{B}$ are conformal square matrices.

THEOREM 3. For any $\mathbf{x} \in S_k$ and $\mathbf{A} \in S_n$, $(\mathbf{x} \times \mathbf{A})' = \mathbf{A}' \times \mathbf{x}'$.

Proof. Let $(\mathbf{x} \times \mathbf{A}) = \mathbf{B} \in S_n$. Let $\mathbf{B}' = \mathbf{D} \in S_n$, where $\mathbf{d}_{rs} = (\mathbf{b}_{sr})'$. Now $\mathbf{b}_{rs} = \mathbf{x} \cdot \mathbf{a}_{rs}$. Hence, $\mathbf{b}_{sr} = \mathbf{x} \cdot \mathbf{a}_{sr}$ and $(\mathbf{b}_{sr})' = (\mathbf{a}_{sr})' \cdot \mathbf{x}'$. Thus $(\mathbf{x} \times \mathbf{A})' = \mathbf{D}$, where $\mathbf{d}_{rs} = (\mathbf{a}_{sr})' \cdot \mathbf{x}'$, and we have $(\mathbf{x} \times \mathbf{A})' = \mathbf{A}' \times \mathbf{x}'$.

In a similar manner we may prove that $(\mathbf{A} \times \mathbf{x})' = \mathbf{x}' \times \mathbf{A}'$.

9. Special cases of X-multiplication. Throughout this paper we have been concerned with two sets of matrices S_n and S_k, where $n = mk$ for m and k positive integers. For the sake of generality we have required m and k to be other than one. The removal of this restriction results in two special cases of X-multiplication, the multiplication of a matrix by a scalar and the ordinary product of square matrices.

Case I. If $k = 1$, the product $(\mathbf{x} \times \mathbf{A})$ becomes simply the multiplication of an m by m matrix $\mathbf{A}$ by a scalar $\mathbf{x}$, and the usual results are obtained. That is, S_m forms an algebra over $\mathfrak{F}$ since properties (1) - (14) are satisfied.

Case II. If $m = 1$ then $n = k$. We recall that X-multiplication is executed on a matrix $\mathbf{A}$ partitioned into submatrices of order k. Since $\mathbf{A}$ itself is now its only k by k submatrix, we have ordinary multiplication between two matrices of order k. Properties (6′), (7′), and (8′) become properties (11), (12), and (13) respectively. The $\mathbf{i}$ and $\mathbf{A}$ of property (9′) are now of the same order. A property analogous to (14) still does not hold, of course, since it did not hold for X-multiplication.

(Continued on page 128.)

"Prime", "Elementary", and "Fundamental" Comparisons

STEPHEN I. BROWN
Harvard Graduate School of Education

I. Introduction

That there exist multiplicative semigroups which do not enjoy the unique prime factorization property (The Fundamental Theorem of Arithmetic), is well known [1], [2], [3, p. 31]. So, for example, let us denote the set of natural numbers by N, and confine our attention to the subset $E = \{1\} \cup \{x : x = 2n \text{ for } n \in N\}$ — 1 together with all the even elements of N. If we once again define a prime as *a number with exactly two different factors* (both of which must belong not only to N but to E this time), then we can show that some numbers in E are factorable into a product of primes in two different ways. Thus, $36 = 6 \cdot 6 = 2 \cdot 18$, $60 = 6 \cdot 10 = 2 \cdot 30$, $100 = 10 \cdot 10 = 2 \cdot 50$. It is not difficult to show that any composite in E which is divisible either by two primes (in E) both of which are greater than 2, or by some p^2 where p is a prime (in E) greater than 2, must violate The Fundamental Theorem of Arithmetic in E.

Some numbers possess *three* different prime factorizations in E. So for example, $216 = 6 \cdot 6 \cdot 6 = 2 \cdot 6 \cdot 18 = 2 \cdot 2 \cdot 54$. As a matter of fact, an obvious generalization of the case for two distinct factorizations, enables one to find elements of E having as many different prime factorizations as he wishes. Note that we have included 1 in E (at the expense of closure under addition) in order to permit the same definition of prime in both N and E.

As expressed in [2], a strong pedagogical reason for investigating such a system is that it sheds further light on what seems on first glance to be an obvious and trivial theorem in N. The fact that the theorem fails in E suggests that there must in fact be a profound conceptual difference between the two systems. We are thus motivated to re-examine the proof in N with a new pair of lenses — one which encourages us to syphen out relevant aspects of the proof which justify the existence of a theorem in one set and its failure in another.

In studying certain aspects of elementary number theory it is profitable to extend the comparisons (between the natural numbers and similar conjectures in particular subsets) substantially beyond The Fundamental Theorem of Arithmetic, and it is a consideration of these problems (in II) that represents the focus of this article. It is a considerable elaboration and refinement of comparisons of this sort made elsewhere ([4]).

Many of the *unsolved* problems in N have trivially simple solutions in E. Problems that are *elementary* in N only in the technical sense that they do not invoke concepts of a complex variable or of continuity principle for the real numbers, become truly elementary in the common-sensical use of the term when investigated in E. We shall proceed by enumerating several unsolved or difficult famous prime number problems in N, and will suggest the answers in E. In most cases the proofs in E are simple and will be omitted. In all cases the proofs in N (if they exist) are far from trivial, and can be found in most introductory number theory texts.

We could of course disguise the fact that we are working in E (as is done in [2]) by employing the isomorphic system of two by two matrices with equal integral entries. For ease of exposition, however, we shall not do so here.

II. Some Comparisons

A—On a formula to generate primes:

What is the nth prime? The fact that the number of primes in N is infinite (proven so elegantly by Euclid over 2000 years ago) but countable, assures us that the question makes sense and has a definite answer for each n. That it is theoretically answerable by a "brute force" application of the definition of prime (or some consequence of the definition) in conjunction with the sieve technique of Eratosthenes, however, does not guarantee (even with the aid of high speed computers) that the answer for large n may not take centuries to compute. A formula that would generate all primes in succession would represent a quite adequate solution to the problem. Such solution, however, has eluded the grasp of research mathematicians for centuries, and a liberal revision of the question which (though still requiring an infinite number of primes and only primes) allows for the omission of many primes from the sequence has also met with no success in N.

The functions of Fermat ($F(n) = 2^{2^n} + 1$) and Mersenne ($M(p) = 2^p - 1$ for p a prime belonging to N) are well known

17th century attempts to answer the revised question. The inefficiency of "brute force" in determining primality is illustrated by the fact that a century elapsed before a mathematician (and it took the genius of Euler in 1732) furnished a counter-example to Fermat's conjecture. Most revealing is the fact that $F(n)$ breaks down at an embarrassingly early stage, for $F(5)$ is divisible by 641. Similarly $M(11)$ is not prime since it equals $23 \cdot 89$.

Though the conjectures as they stood were both false, Mersenne and Fermat have in a sense been vindicated by the following two theorems (see [3, p. 50]) which represent modified converses of the proposed conjectures:

1—If $a^n + 1$ is a prime (in N) where $a > 1$, then a is even and $n = 2^r$.

2—If $a^n - 1$ is a prime (in N) and $n > 1$, then $a = 2$ and n is prime.

There exist many other simple formulas that generate a significant number of primes before "breaking down". Though it succeeds for the first forty substitutions, Euler's famous $x^2 + x + 41$ obviously fails to yield a prime for $x = 41$. Similarly $x^2 - 79x + 1601$ works for the first eighty substitutions before yielding a composite. In general, it can be shown that there exists a polynomial function for any n which will yield at least n successive primes. Eventually though, the formula breaks down, and whether or not the number of primes generated after such failure is infinite, is still an unsolved problem. An attempt is made in [5] to attack this problem in a probabilistic sense. Using a computer, they showed that there exist some functions that are quite rich in primes. Thus for primes in the Euler form $n = x^2 + x + 41$, they found the ratio r of these to all numbers of this form n up to ten million to be $5 = .475 \cdots$.

In 1947 W. H. Mills proved that there exists an exponential type function which does in a peculiar sense solve the problem of generating an infinite number of primes in N (allowing for omissions). He has shown that there exists a real number A having the property that $[A^{3^n}]$ is prime for every natural number n. The interesting twist in his "solution" however, is that the proof is purely existential, i.e., nothing is known about the actual value (nor even of the order of magnitude) of A.

Let us now consider the prime generator question in E. The first few primes are: 2, 6, 10, 14. Since any odd element of N can have only odd factors, it is not difficult to see that any element of the form $2 \cdot (2n - 1)$ for any $n \epsilon N$ must be prime in E. Furthermore successive substitution of elements of N yields all the primes of E in order. For aesthetic reasons, we may prefer a formula whose domain is also E instead of N as in the above case. Then we could characterize a prime as any number expressible in the form $2e - 2$ for $e \epsilon E - \{1\}$, and the nth prime would be equal to $2(2n) - 2$. Notice that in E, since we readily meet with success, there is no need to revise the original question (searching for a formula that produces an infinite number of primes allowing for a countable number of omissions).

B—On the infinitude of primes belonging to arithmetic progressions:

In the previous section we indicated that for many polynomial functions which are rich in primes, we cannot ascertain whether or not they generate an infinite number, granted a countable number of failures. This of course does not imply that (at this time) the answer is undecidable for all polynomial functions. Two obvious formulas which generate an infinite number of primes (and incidentally an infinite number of composites also) are: n and $2n - 1$ for all $n \epsilon N$.

That we can generalize from the above instances to include other arithmetic progressions, is a conclusion that was proved by Dirichlet (1805-1859). Let us denote arithmetic progressions by $a + bn$, where a and b are fixed and n varies in N. Though it is obvious that the only candidates which have a chance of succeeding are those for which a and b are relatively prime (for if they have a factor in common greater than 1, each number will be composite) it does not follow that the condition is a sufficient one. One can gain an intuitive appreciation that this is the case for any a and b, by constructing a sieve (of Eratosthenes) with b columns, and noting that there are many primes in all columns for which a and b are relatively prime.

It is also possible in a few cases (like progressions of the form $3 + 4n$ or $5 + 6n$) to provide a simple deductive proof of the infinitude of primes belonging to arithmetic progressions, and the proof is very similar in spirit to Euclid's original proof of the infinitude of primes. Dirichlet's general proof is non-elementary, in that it requires concepts from complex number theory, and is beyond the scope of an introductory course in elementary number theory.

It was suspected for over a century that no "elementary" proof of Dirichlet's theorem existed. In 1949, however, Atle Selberg proved it without invoking complex variables or the principle of the continuity of the real numbers. [6].

When we turn to E, we again appreciate that the characterization of all arithmetic progressions which generate an infinitude of primes is easily discovered and proved. That relative primeness is neither necessary, nor sufficient, is demonstrated by the fact that $6 + 12n$ succeeds while $4 + 6n$ fails (and for purposes of analogy, we assume that the domain of n is E). It is not difficult to see that $a + bn$ generates an infinitude of primes in E (a, b, and n, ε E) if and only if a is prime in E. Furthermore, these generators succeed in a way that the functions of Fermat and Mersenne fail, i.e., if a is prime, then *all* elements generated by $a + bn$ are prime.

C—On determination of primality:

That the problems of determining primality for any number, and of producing a formula to generate all primes, may not be equivalent is revealed by the fact that the former request is theoretically answerable (though perhaps very time consuming) for any element of N, while the latter has not been solved, and as a matter of fact may be unanswerable [7].

Evidence of the practical difficulty of determining primality has already been cited with regard to conjectures such as those of Mersenne and Fermat. Though the problem is conceptually the same today as it was when Euler disproved Fermat's conjecture for a *ten* digit number, the existence of computers has provided additional motivation to dichotomize "superastronomically" large numbers according to primality. Thus, Gillies in 1963 demonstrated the primality of $M(11213)$, one of the largest known primes. Lest we conceive of the task as insignificant, let us recall that the number has approximately 3375 digits in decimal form.

Though several theorems provide us with a drastic reduction in the number of divisions necessary to determine primality, they do so at an expense that only the highest of high speed computers can afford. It is easily seen that to determine the primality of n, we need not divide by numbers which exceed $\sqrt{n}$. There are other theorems which, though they require only one division, do involve much larger dividends than is the case when we apply either the definition or the square root law. Wilson (1741-1793) for example established that a number n is prime if and only if $(n - 1)! \equiv -1 \pmod{n}$. Also Proth in 1878 discovered a way of determining primality for numbers of the form $n = k2^m + 1$ for $k < 2^m$. If b has the property that there exists no s such that $s^2 \equiv b \pmod{n}$, then n is prime if and only if $b^{\frac{b-1}{2}} \equiv -1 \pmod{n}$. Notice that the implied division involves such huge dividends that to determine the primality of 103 by Wilson's formula would be a formidable hand calculation.

In E on the other hand, the existence of a formula to generate all primes, also solves the very practical problem of determining primality for any number. Since all primes of E are generated by $2e - 2$, and all composites by $2e$ (for $e \in E - \{1\}$) it follows that divisibility by 2 (in E) is the simple test for primality.

D—On distribution of primes:

Granted that we lack a simple formula for the generation of primes in N, is there some way of determining their density, or how they tend to cluster? Observation of a table of primes suggests that, though the number of primes is infinite, as we progress through the natural numbers, the number of primes within a fixed span tends to decrease. So, in the first five groups of 5000 elements of N, we have the following number of primes (respectively): 168, 135, 127, 120, 119.

The search for a simple formula to approximate the number of primes $\leq x$ for any x belonging to N (and $\pi(x)$ denotes that number), captured the imagination of Gauss and Legendre in the late 18th century. Over 100 years later, in 1896 two independent proofs (known now as the Prime Number Theorem) by Vallee-Poussin and T. Hadamard verified the accuracy of the 18th century approximations.

Using concepts that depend upon functions of a complex variable (and it took over fifty years to find an "elementary" — in the technical use of the word — proof), Vallee-Poussin and Hadamard asserted and proved the following:

$$\lim_{x \to \infty} \frac{\pi(x)}{\frac{x}{\ln x}} = 1.$$

Thus

$$\pi(x) \sim \frac{x}{\ln x}.$$

Though this formula tells us how to estimate the number of primes in an interval, it does not suggest a way of deciding where in the interval those primes lie.

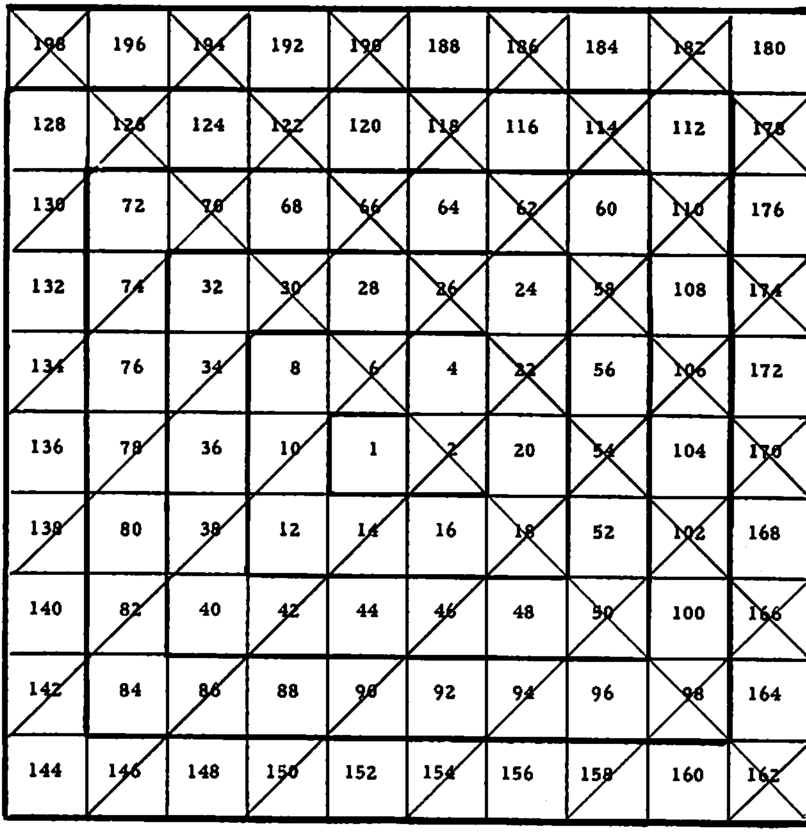
In [5], an attempt is made to determine to some extent paths along which these primes tend to cluster. If points in a plane are numbered in a counter-clockwise spiral sequence from the origin, [5, fig. 1] and [8, cover], then if we examine those lattice points which represent a prime number in the sequence, it seems to be the case that they tend to cluster along straight line segments. Such indices consist of values of a quadratic form, since the nature of the spiral determines that the second differences between such points be constant. In particular, Euler's form is one of the diagonals.

Both $\pi(x)$ and various spirals (only one of which we have considered in the above discussion) lead to rather interesting comparisons in *E*. Since elements of *E* are double those of *N* (with the exception of 1), and since every other element is prime, it follows that for any $x \varepsilon E$, $\pi(x)$ is approximately $\frac{1}{4} \cdot x$. The formula is exactly correct for any composite in *E*, and misses by $\frac{1}{2}$ if the number in *E* is prime. It is easy to see that the "greatest integer" function enables us to produce a single concise formula that covers all the cases (including $\pi(1)$). Thus

$$\left[\frac{x+2}{4}\right]$$

not only approximates $\pi(x)$, but furnishes us with an *exact* value in each case. Note that the approximation due to the Prime Number Theorem in *N* can *never* be exact for *any* element in the sequence since it can be proven that $\ln x$ is irrational for all integral *x*.

In the case of a spiral in *E*, we need only double each index (with the exception of 1) that appears in spiral of N. Thus we arrive at figure 1.



A spiral grid indicating the distribution of "primes" in the set E

Fig. 1

It is easy to demonstrate that the tendency which is suggested in the spiral of *N* becomes fully realized in *E*. Since any two elements along a diagonal differ by a multiple of 4 (as they do by a multiple of 2 in *N*) — with the obvious exception of 1 along its diagonal — it is easy to see that the primality of any element along a diagonal, determines the primality of all elements that lie along it. Also since any two adjacent elements (horizontal) differ by a multiple of 2 (again excluding 1) but not of 4, it is apparent why the primes and composites align themselves along alternate diagonals as they do.

We see once again that since the second differences are constant along any diagonal, formulas of a quadratic nature will generate these elements that are "super-rich" in primes of E.

E—On Odds and Evens:

The (still unproven) conjecture of Goldbach (1742) that any even $n \varepsilon N$, greater than 2 can be represented as the sum of two primes, is an obvious parity problem. In considering that same problem in E, we must first define what we mean by "even" in that set. If we employ the same definition as in N (divisible by 2), we realize that, though all elements of $E - \{1\}$ are even in N, they are not all even in E. We recall that none of the primes are divisible by 2 in E, and that all composites are. Therefore, unlike N, all the composites of E are even, but just as in N, all the primes in E (greater than 2) are odd. Goldbach's gnawing conjecture then becomes a trivial one line theorem in E: $2e = (2e - 2) + 2$.

The concept of parity tends to enrich another comparison between N and E. We know that another unsolved famous prime number conjecture is that there exists an infinite number of twin primes. But what do we mean by "twin primes"? One obvious interpretation is that in order to qualify, both p and $p + 2$ must be prime. If we thus literally interpret the conjecture in E, then obviously there are no twin primes in that set, for all primes must differ by 4.

Another way of looking at the problem, however, is to realize that in N we are searching for twin primes from the set of successive odd elements. It so happens that in N, odd elements differ by 2, but we need not choose this literal interpretation for the purpose of generalizing to other domains. If we search for twin primes in E from among the successive odd elements, then it is clear that the twin prime conjecture becomes a bona fide theorem in that set, since for any $e \varepsilon E - \{1\}$, it is clear that $2e - 2$ and $2e + 2$ are both primes.

III. Conclusions

We have only begun to suggest the kinds of comparisons that can be made among different domains. The pedagogical relevance of such an approach is that it encourages us to distill those abstract properties (for example, mathematical induction) that do in fact distinguish one domain from another. In cases where relatively simple and "elementary" proofs exist in N (such as the Fundamental Theorem of Arithmetic), the consideration of another domain in which either the conclusion is different or the proof is more trivial, goads the student to make more of proof than the digestion of someone else's thinking.

Proofs that may be too difficult to include in an elementary course (as Dirichlet's Theorem, or The Prime Number Theorem), can at least be discussed in historical context, and can also be made more significant for future encounter by whetting the appetite with analogous problems in other domains that can (at the time) be tackled. Unsolved problems (as Goldbach's conjecture, the twin prime conjecture, production of a formula to generate all primes) gain stature as one becomes aware of the fact that the obvious techniques to solve the problem in other domains break down when applied to N.

As we have indicated, there is much in the literature on comparisons of domains with regard to The Fundamental Theorem of Arithmetic, and we have, therefore, not emphasized comparisons of this nature. We note in passing that the significance of the theorem can be further strengthened by considering the pervasiveness of consequences in domains where the theorem fails. The number-theoretic functions like φ, τ, σ cannot be so simply expressed as they are in N, for non-unique factorization destroys the multiplicative nature of these functions. Many fractions cannot be "reduced to lowest terms" (in the sense of getting *an* answer), for there may be more than one way of expressing an equivalent fraction so that the numerator and denominator are relatively prime. Thus, for example,

$$12/36 = \frac{6 \cdot 2}{6 \cdot 6} = 2/6; \text{ but also } 12/36 = \frac{6 \cdot 2}{18 \cdot 2} = 6/18.$$

There are many other simple divisibility properties that fail in E. For example, the greatest common divisor theorem in N asserts that every common divisor of a and b must divide the greatest common divisor. The following counter-example suggests that the proof depends upon properties that are not shared by N and E: The greatest common divisor of 36 and 60 is 6. Also, 2 divides 36 and 2 divides 60, yet 2 does not divide 6.

The choice of E in order to make such comparisons has been convenient but somewhat arbitrary, and we have found it profitable to consider many other subsets of N — such as 1 together with multiples of 3. The kinds of comparisons we have suggested in this article gain in significance when it becomes necessary for us to generalize concepts such as even and odd (for the purpose of considering Goldbach's conjecture, or the infinitude of twin primes for example) to sets that lack 2 as an element.

REFERENCES

1. Newcomb Greenleaf and Robert J. Wisner, "The Unique Factorization Theorem," *Mathematics Teacher,* 52 (1959) 600-603.
2. Bernard Jacobson and Robert J. Wisner, "Matrix Number Theory: An Example of Non-Unique Factorization," *The American Mathematical Monthly,* 72 (1965) 399-402.
3. Calvin T. Long, *Elementary Introduction to Number Theory,* D. C. Heath and Company, Boston, 1965.
4. Stephen I. Brown, "Of 'Prime' Concern: What Domain?", *Mathematics Teacher,* 58, (1965) 402-407.
5. M. L. Stein, S. M. Ulam, and M. B. Wells, "A Visual Display of Some Properties of the Distribution of Primes," *The American Mathematical Monthly,* 71 (1964) 515-520.
6. Atle Selbert, "An Elementary Proof of Dirichlet's Theorem about Primes in an Arithmetic Progression," *Annals of Mathmatics,* (2) 50, 1949, 297-304.
7. David Hawkins, "Mathematical Sieves," *Scientific American,* Dec. 1958 105-112.
8. Martin Gardner, "Mathematical Games," *Scientific American,* March, 1964 120-128 and cover.

Flexing Rings of Regular Tetrahedra

Douglas A. Engel
Student, Fort Hays Kansas State College

A tetrahedral ring is formed by connecting opposite edges of several tetrahedra to form a circular, hinged, chain [1]. The purpose of this paper is to show how rings of regular tetrahedra can be formed with cyclic properties of movement. If an element (regular tetrahedron) of a ring is held non-rotatable in space and the ring can be rotated about it any number of times, the ring is *flexible*. A ring of eleven elements, with a ¾ twist between its ends, has an interesting cycle of flexations and a unique symmetry.

Two chains of regular tetrahedra can be wound about one another to form a solid core (Fig. 1). A chain that is given a simple

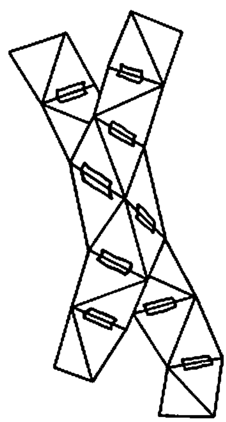


Fig. 1

twist forms a core with a shape congruent to that of two chains wound about one another, but it is not rigid. Four of the rigid cores can be connected at an intersection to give a structure that flexes by winding the opposite cores into the adjacent cores (Fig. 2). This

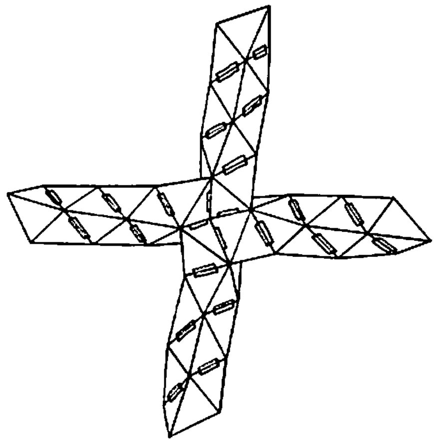


Fig. 2

structure alone is not a true flexible solid, since any number of rotations of the ring about an element are not possible. By leaving a loose portion in each of the four arms they can be reconnected into a four armed structure (Fig. 3). The resulting structure is flexible.

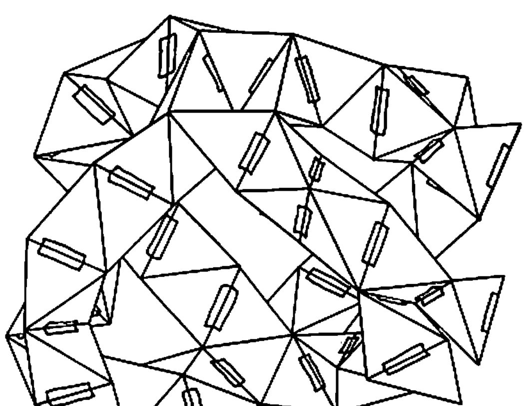


Fig. 3

A flexing structure type can be formed by winding two chains about one another for a short ways and then winding one of the chains out, around the other, before continuing the winding process (Fig. 4). The flexing motion is a simple rolling of the

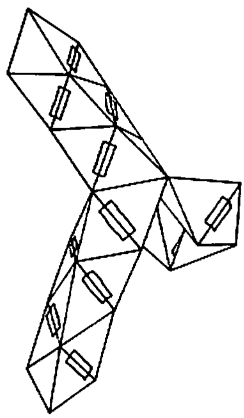


Fig. 4

tetrahedra in the protruding portion, in either direction, along the ring. The protruding structure bends the cores about 30°, allowing a circular chain of a number of protruding structures which is a true flexing solid. The loose portions in the structure in Fig. 3 can be replaced by a number of the structures in Fig. 4, giving a flexible solid that is rigid except for its flexing movement. More complex solids can be formed by having more four armed structures in the ring. The number of twists between the ends of two chains wound about one another is approximately $\frac{n}{4}$, where n is the total number of elements. Regular tetrahedra in a linear core can be flattened to a layer of squares. The four armed structure can be reduced to a general type of tetraflexagon by flattening all four cores [2] [3].

(Continued on page 122.)

The Problem Corner

EDITED BY H. HOWARD FRISINGER

The Problem Corner invites questions of interest to undergraduate students. As a rule the solution should not demand any tools beyond calculus. Although new problems are preferred, old ones of particular interest or charm are welcome provided the source is given. Solutions of the following problems should be submitted on separate sheets before October 1, 1967. The best solutions submitted by students will be published in the Fall 1967 issue of *The Pentagon*, with credit being given for other solutions received. To obtain credit, a solver should affirm that he is a student and give the name of his school. Address all communications to Professor H. Howard Frisinger, Department of Mathematics and Statistics, Colorado State University, Fort Collins, Colorado 80521.

PROPOSED PROBLEMS

201. *Proposed by William Mikesell, Indiana University of Pennsylvania, Indiana, Pennsylvania.*

Prove the following statement: In the set of regular polygons only three, the triangle, square, and the hexagon are such that they can fit together exactly without any gaps or overlaps.

202. *Proposed by R. S. Luthar, Colby College, Waterville, Maine.*

Show that there are infinitely many primes of the form:

$$x^3 + y^3 + z^3 + u^3 + t^3.$$

203. *Proposed by Layne Watson, Evansville College, Evansville, Indiana.*

Prove that the sum of N vectors of equal length radiating from a point P is zero, where the angle between a vector and the preceding one is $\frac{2\pi}{N}$. Use this result to prove that $\sum_{n=0}^{N-1} \cos \frac{2\pi n}{N} = 0$ and $\sum_{n=0}^{N-1} \sin \frac{2\pi n}{N} = 0$, where N is an integer > 1.

204. *Proposed by the Editor.*

Show that among any ten consecutive positive integers, at most five can be primes, and that five actually occur in only one case. Must at least one of any ten consecutive positive integers be prime?

205. *Proposed by Thomas P. Dence, Bowling Green State University, Bowling Green, Ohio.*

The Fibonacci sequence $\{F_n\}$ is defined as

$$F_0 = 0, F_1 = 1, \cdots, F_k = F_{k-1} + F_{k-2} \text{ for } k \geqq 2.$$

Now let $f(x)$ represent the continued fraction

$$f(x) = F_0 + \cfrac{1}{F_1 + \cfrac{1}{F_2 + \cfrac{1}{F_3 + \cfrac{1}{F_4 + \ddots + \cfrac{1}{F_n}}}}}$$

and let $g(x)$ represent the continued fraction

$$g(x) = F_0 + \cfrac{F_1}{F_2 + \cfrac{F_3}{F_4 + \cfrac{F_5}{F_6 + \ddots + \cfrac{F_{n-1}}{F_n}}}}$$

Determine, whether possible or not, and if so, an exact value for $\lim\limits_{n \to \infty} f(x)$ and $\lim\limits_{n \to \infty} g(x)$.

SOLUTIONS

196. *Proposed by William K. Sjoquist, University of California at Berkeley, Berkeley, California.*

If $y = uv$ where u and v are functions of x, prove that the nth derivative of y with respect to x is given by

$$y^{(n)} = uv^{(n)} + nu'v^{(n-1)} + n(n-1)\, u''v^{(n-2)}/2!$$
$$+\ n(n-1)(n-2)\ u'''\ v^{(n-3)}/3! + \cdots + u^{(n)}v.$$

Solution by Lyne H. Carter, University of Southern Mississippi, Hattiesburg, Mississippi.

Assuming that $v^0 = v$, we need to show that

$$(^*)y^{(n)} = uv^{(n)} + nu'v^{(n-1)}\ n(n-1)u''v^{(n-2)}/2!$$
$$+\ n(n-1)(n-2)u'''\ v^{(n-3)}/3!$$
$$+ \cdots + \frac{n(n-1)(n-2)\cdots 2u^{(n-1)}v'}{(n-1)!}$$
$$+\ u^{(n)}v.$$

Note that the expansion for $y^{(n)}$ terminates at the $(n+1)$th term. A simple induction proof should suffice; let

$$S = \{x \mid x \text{ is a positive integer, } (^*) \text{ holds for } n = x\}.$$

Clearly $1 \varepsilon S$, since from the product differentiation formula (and if y' exists),

$$y^{(1)} = y' = u'v + uv' = uv^{(1)} + (1)u^{(1)}v^{(1-1)},$$

which is in the form of $(^*)$ with $n = 1$. Now assume that $k \varepsilon S$ for some positive integer k. Then $(^*)$ takes the form

$$^{(I)}y^{(k)} = uv^{(k)} + ku'v^{(k-1)} + k(k-1)u''v^{(k-2)}/2!$$
$$+\ k(k-1)(k-2)u'''\ v^{(k-3)}/3!\ \cdots$$
$$+\ k(k-1)(k-2)\cdots 2u^{(k-1)}v'/(k-1)!$$
$$+\ u^{(k)}v.$$

Assuming that $y^{(k+1)}$ exists, take the derivatives of the functions on each side of the equation and obtain

$$^{(II)}y^{(k+1)} = uv^{(k+1)} + u'v^{(k)} + ku'v^{(k)} + ku''v^{(k-1)}$$
$$+\ k(k-1)u''v^{(k-1)}/2! + k(k-1)u'''v^{(k-2)}/2!$$
$$+\ k(k-1)(k-2)u'''\,v^{(k-2)}/3! + \cdots$$
$$+\ k(k-1)(k-2)\cdots 2u^{(k)}v'/(k-1)! + u^{(k+1)}v + u^{(k)}v'.$$

$\therefore y^{(k+1)} = uv^{(k+1)} + (k+1)u'v^{(k)} + (k+1)ku''v^{(k-1)}/2!$
$+ (k+1)k(k-1)u'''v^{(k-2)}/3! + \cdots + (k+1)(k-1)k$
$\cdots 2u^{(k)}v'/(k-1)! + u^{(k+1)}v.$

This expression is clearly (*) with $n = k + 1$, so since $k \in S$ implies that $k + 1 \in S$, the principle of mathematical induction tells us that S is the set of all positive integers. Therefore (*) holds for all positive integers and the theorem is proved.

Also solved by Douglas Lind, University of Virginia, Charlottesville, Virginia; Layne Watson, Evansville College, Evansville, Indiana.

197. *Proposed by Thomas P. Dence, Bowling Green State University, Bowling Green, Ohio.*

Let us denote a set of sequences $\{X_{m,n}\}$ by

$\{X_{1,n}\} = (1, 1, 2, 3, 5, \cdots)$ where $X_{1,1} = 1$
$X_{1,2} = 1$
$X_{1,3} = 2$
•
•
•
$X_{1,k} = X_{1,k-1} + X_{1,k-2}$

$\{X_{2,n}\} = (1, 3, 4, 7, 11, \cdots)$ where $X_{2,1} = 1$
$X_{2,2} = 3$
$X_{2,3} = 4$
•
•
•
$X_{2,k} = X_{2,k-1} + X_{2,k-2}$

$\{X_{3,n}\} = (1, 4, 5, 9, 14, \cdots)$ for similar definitions of $X_{3,n}$
•
•
•

$\{X_{i,n}\} = (1, i+1, i+2, 2i+3, \cdots)$ for $i > 1$, and where the X's are defined by the same recurrence relation as before.

Express the nth term of the ith sequence in terms of the nth *term of the first sequence,* where the first sequence is actually the Fibonacci sequence.

No solution. Note that $X_{i,n} = X_{1,n} + X_{1,n-i}$. The problem is to express $X_{1,n-1}$ in terms of $X_{1,n}$.

198. *Proposed by the Editor.*

For what values of n is $(11 \times 14^n) + 1$ prime?

Solution by R. S. Luthar, Colby College, Waterville, Maine.

1. Suppose n is odd.

$$\begin{aligned} 11 &\equiv 1(5) \\ 14^n &\equiv -1(5) \\ \therefore \quad 11 \times 14^n &\equiv -1(5) \end{aligned}$$

Hence, $11 \times 14^n + 1$ is divisible by 5.

2. Suppose n is even, say $2k$

$$\begin{aligned} 11 &\equiv -1(3) \\ 14^n = 14^{2k} &= 196^k \equiv 1(3) \\ \therefore \quad 11 \times 14^n &\equiv -1(3) \end{aligned}$$

Hence, $11 \times 14^n + 1$ is divisible by 3.

$\therefore$ $11 \times 14^n + 1$ is a composite number and does not give prime for any value of n.

Also solved by Thomas P. Dence, Bowling Green State University, Bowling Green, Ohio; Douglas Lind, University of Virginia, Charlottesville, Virginia; Layne Watson, Evansville College Evansville, Indiana.

199. *Proposed by R. S. Luthar, Colby College, Waterville, Maine.*

Let ΔABC be a right triangle with right angle at A. Construct regular n-gons on $\overline{AB}$, $\overline{AC}$, and $\overline{BC}$ with respective areas α, β, γ. Prove $\alpha + \beta = \gamma$.

Solution by Layne Watson, Evansville College, Evansville, Indiana.

The area of a regular n-gon with side s is given by

$$A = \frac{s^2 n}{4} \cot \frac{\pi}{n}. \quad \alpha = \frac{\overline{AB}^2 n}{4} \cot \frac{\pi}{n},$$

$$\beta = \frac{\overline{AC}^2 n}{4} \cot \frac{\pi}{n}, \ \gamma = \frac{\overline{BC}^2 n}{4} \cot \frac{\pi}{n}$$

and since ABC is a right triangle $\overline{AB}^2 + \overline{AC}^2 = \overline{BC}^2$. Multiplying by

$$\frac{n}{4} \cot \frac{\pi}{n} \text{ gives } \frac{n}{4} \cot \frac{\pi}{n}\, \overline{AB}^2 + \frac{n}{4} \cot \frac{\pi}{n}\, \overline{AC}^2 = \frac{n}{4} \cot \frac{\pi}{n} \overline{BC}^2$$

which is $\alpha + \beta = \gamma$.

Also solved by William Mikesell, Indiana University of Pennsylvania, Indiana, Pennsylvania.

200. *Proposed by E. R. Deal, Colorado State University, Fort Collins, Colorado.*

In the Fall 1966 issue of *The Pentagon,* this problem was incomplete. The complete problem should read as follows:

"Are those your children I hear playing in the garden?" asked the visitor.

"There are really four families of children," replied the host. "Mine is the largest, my brother's family is smaller, my sister's is smaller still, and my cousin's is the smallest of all. They are playing drop the handkerchief," he went on. "They prefer baseball but there are not enough children to make two teams." "Curiously enough," he mused, "the product of the members in the four groups is my house number, which you saw when you came in."

"I am something of a mathematician," said the visitor. "Let me see whether I can find the number of children in the various families." After figuring for a time, he said, "I need more information. Does your cousin's family consist of a single child?" The host answered his question, whereupon the visitor said, "Knowing your house number and knowing the answer to my question, I can now deduce the exact number of children in each family."

How many children were there in each of the four families?

Kappa Mu Epsilon News

EDITED BY J. D. HAGGARD, Historian

Alabama Beta, Florence State College, Florence

At an initiation banquet in April, 1966, Alabama Beta initiated thirteen new members: Eddy Joe Brackin, Mary Virginia Darby, David Hammond, Joyce Hargrave, Ronald Killen, Martina Lamb, Frank Lee, Pamela Sams, Mary Emma Wakefield Albert Wallace, Michael Weston, Linda White, and Barbara Wright. The guest speaker was James Hooper, a 1958 initiate of Alabama Beta.

Other programs in the spring of 1966 included lectures by Dr. J. W. Wesson of Vanderbilt University on linear algebra and on projective geometry, and a student program directed by Marjory Johnson.

Six of our 1966 graduates have received graduate scholarships and assistantships: Norman Cooper, N.A.S.A. Trainee Grant to the University of Mississippi; Marjory Johnson, N.D.E.A. Fellowship to the State University of Iowa and honorable mention for the Woodrow Wilson Fellowship; Cecilia Holt, N.D.E.A. Fellowship to the University of Kentucky; Bettye Bergin, NSF Graduate Fellowship to the University of Alabama; Janice Cox, teaching assistantship to Sanford University; Harold Darby, teaching assistantship to Tennessee Technological University. Bettye Bergin, Cecilia Holt, Hugh Huffman, and Patricia Powell were named to Who's Who in American Colleges.

An alumnus, John Finley, received the annual "Faculty Member of the Year" award.

A reception for all freshmen who are interested in majoring or minoring in mathematics was held to acquaint them with the purposes, advantages, and requirements for membership in Kappa Mu Epsilon.

A coffee hour at the 1966 homecoming was attended by sixty-two alumni and guests from nineteen different years of initiation. Alabama Beta has initiated 410 members since its installation in May, 1935.

Alabama Epsilon, Huntingdon College, Montgomery

Again this year the chapter is sponsoring a tutoring service for students in high schools or nearby colleges. The participation in this program has been excellent. We are planning an initiation for later in the semester.

Illinois Beta, Eastern Illinois University, Charleston

We now have sixty-seven members on campus, twenty-six of these were initiated last spring, and twenty are faculty. Only twenty-one members were participants last year, but this limited number of "two-year" people will be changed in the future by having initiation of pledges twice a year. Those who meet the requirements during spring or summer quarters will be initiated in the fall; those meeting requirements during fall or winter quarters will be initiated in the spring. There will continue to be only one banquet held in conjunction with the spring initiation. We hope that membership in Kappa Mu Epsilon for more than one year will help our organization become more active.

Four of our members attended the Regional Convention at Mount Mary College, Milwaukee, Wisconsin. Expenses were shared equally by the organization treasury and those who attended. There were very favorable reports of the convention and requests for a similar trip next spring.

Indiana Alpha, Manchester College, North Manchester

The theme for the programs this year is "Would You Believe, Mathematics Is Everything?" Each program will be about applications in or relations to other fields. We are also planning field trips to the National Council of Teachers of Mathematics meeting in Cincinnati and to a planetarium.

Indiana Gamma, Anderson College, Anderson

Last year's chapter president, James French, is now a graduate assistant in mathematics at the University of Nebraska.

Kansas Alpha, Kansas State College, Pittsburg

Harold Thomas replaced J. Bryan Sperry as sponsor since Mr. Sperry is currently enrolled in graduate work at the University of Kansas.

Louise Gomer received the Mendenhall Memorial Award as the outstanding senior majoring in mathematics.

Programs for the year included: "Conic Sections with Circles as Focal Points" by Tom Potts; "The Five-Color Problem" by Dr. Arthur Bernhart, University of Oklahoma; "The Planimeter" by Roger Christian; "Game Theory" by Tony Dousette; "Mathematical

Fallacies" by Curtis Woodhead; "Mathematics In Industry" by John Brunet and Edd Grigsby, Phillips Petroleum Co., Bartlesville, Oklahoma. Both of these men are former KSC students and members of Kansas Alpha.

Kansas Epsilon, Fort Hays Kansas State College, Hays

Initiation of new members was held at the monthly meeting on November 15, 1966.

On February 14, 1967, Lt. Dee E. Kimbell, a member of Kansas Epsilon, was awarded the Department of Commerce Silver Medal for contributions of unusual value to the department. Kimbell is Chief of the Support and Maintenance Branch of the Satellite Triangulation Division of the Geodetic Survey, an agency of the Commerce Department's Environmental Science Services Administration, with headquarters in Rockville, Maryland.

Kansas Gamma, Mount St. Scholastica College, Atchison

Norma Henkenius is editor of the quarterly publication of the Kansas Gamma Chapter entitled "The Exponent."

Some of the interesting programs of the chapter during this past year have been: A Wassial Bowl Party which follows a custom common in England at Christmas time during past centuries and represents one of the unique ways of expressing "Good health" to one's fellow man; "Geometric Transformations" by Pat Moran; "Mathematical Paradoxes" by Roberta Robinson; "Fundamental Groups of Topology" by Joe Ingle; "Introduction to Geometric Models Based on Axiomatic Systems" by Leora Ernst; "Twin Primes" by Elizabeth Murphy; "Hilbert Spaces" by Anna Agnew; "Equidecomposible Figures" by Andrea Meyer; and "An Introduction to Fibonacci Numbers" by Bernita Meyers.

Maryland Alpha, College of Notre Dame of Maryland, Baltimore

All of the meetings this year (from October through March) are being devoted to a study of computers: — their history, uses, and languages. Each KME member is choosing a problem and a computer language; then she will attempt to program and solve her problem.

In April a specialist in the mathematics of the elementary school will acquaint the members with the current curricular changes.

The May meeting will be the formal initiation.

Michigan Alpha, Albion College, Albion

This fall's meetings have included the following lectures:

"Laplace Transforms in Industry;" "Mathematics as a Career;" and the "Computer."

This spring in addition to our regular presentation of pledge papers we hope to obtain a visiting lecturer from the Mathematical Association of America Visiting Lecture Program.

Mississippi Alpha, Mississippi State College for Women, Columbus

The annual initiation ceremonies and banquet were held October 27, 1966. There were twenty-six new initiates this year. The unusually large number of new members is the result of changes in the curriculum, which permit a student to be eligible for membership earlier in her college career.

A variety of programs has been planned for this year utilizing students and faculty of MSCW and also invited speakers from other institutions. Dr. Robert Plemmons of the University of Mississippi is to present the program for November.

Our program usually centers around graduate study in mathematics. Last spring, Dr. Roy Sheffield of Mississippi State University described their graduate program for our members.

Mississippi Gamma, Mississippi Southern College, Hattiesburg

The first meeting was held on October 19, 1966, to reorganize and set up a program for the 1966-67 year.

On April 13, 1965, the members of Kappa Mu Epsilon enjoyed a steak cook-out and initiated seven new members. This meeting was at the home of Jack D. Munn, corresponding secretary.

Missouri Alpha, Southwest Missouri State College, Springfield

Meetings are held monthly on the third Tuesday of each month. We have fifty active members this semester. Students are preparing papers for the national convention to be held at Kansas Gamma in April.

Missouri Beta, Central Missouri State College, Warrensburg

Missouri Beta Chapter of Kappa Mu Epsilon recently held its first meeting of the academic year 1966-67 and at that time initiated six students into the organization.

The chapter also presented an award to the outstanding freshman in mathematics at the initiation ceremony. At the annual spring banquet held in conjunction with Sigma Zeta, the honorary physics organization, the chapter presents its award to the outstanding senior student in mathematics.

Plans are under way to have papers presented for consideration at the National Convention. In addition, plans are underway to provide a mathematics library for interested students. The library will be stocked with textbooks available from the mathematics staff at the college. The club believes such a library will be quite beneficial. The Missouri Beta Chapter is looking forward to a very successful year.

New York Beta, State University of New York, Albany

The New York Beta Chapter at the State University of New York at Albany held its annual Christmas party on December 14 using mathematical games for entertainment.

On the more serious side, the February meeting was our annual Mathematics Evening, open to all members of the University. This year we held a panel discussion of the various career opportunities open to mathematics majors. The members of the panel included an economist, a statistician, an actuary, and a computer programmer.

In April, at the spring banquet new officers were elected and awards were presented for the best mathematics papers submitted during the year.

New York Gamma, State University College, Oswego

Our first meeting was held September 29, 1966, at which time we organized the club and discussed old business. At our second meeting held October 27, 1966, we discussed the Mathematics Honors Program just started this year. Also, along with Sigma Zeta, the science honorary society, we sponsored a faculty-student tea held November 10, 1966.

Ohio Alpha, Bowling Green State University, Bowling Green

The Ohio Alpha Chapter of Kappa Mu Epsilon is conducting once-a-week help sessions for beginning mathematics students. To also help new students in their selection of courses in the future, the chapter is planning an open meeting of mathematics course reviews presented by various members of the mathematics faculty.

Ohio Gamma, Baldwin-Wallace College, Berea

During the past quarter several interesting papers were presented by Kappa Mu Epsilon members. In April, Camille Falcone presented a paper on "Topology," and Bill Chen presented a paper on "Nomography."

In March, Mr. Paul Diedrich from Clevite Corporation was our guest speaker. He spoke on "The Care and Feeding of Burroughs Computers." Mr. Wayne Heritage from Sohio Company was our guest speaker in May. His topic was "Fibonacci Numbers and Computers." The election of officers was also held in May.

Art Davies, Terry Furman, and Bill Achberger accompanied the staff to the spring meeting of the Ohio section of the Mathematical Association of America at Ohio Wesleyan University, Delaware, Ohio.

In October, we initiated twenty new members, bringing our active total to fifty student members and four faculty members and the all time list of members to two hundred ninety-four initiated since installation in 1947. Several films from the Mathematical Association of America Committee on Educational Media are being shown this quarter. Those presented to date are: "Predicting at Random," "Let Us Teach Guessing," "Topology," "Pits, Peaks, and Passes," "Fixed Points," and "Challenge in the Classroom."

Pennsylvania Beta, La Salle College, Philadelphia

Whereabouts of last year's Kappa Mu Epsilon members: Bro. Jos. Braceland, teacher at Bishop O'Connell High School, Arlington, Virginia; Jos. Chamber, NSF trainee, Georgetown University, Washington, D.C.; Richard Clancy, teacher at Philadelphia High School for Girls; Bro. Foss, teacher at South Hills Catholic High School, Pittsburgh, Pennsylvania; Daniel Gallo, intern teaching program at Temple University, Philadelphia, Pennsylvania; Frank Gutekunst, Pennsylvania Power Co., Allentown, Pennsylvania; Robert McCormick, NSA, Washington, D.C.; George Murr, General Electric, Norristown, Pennsylvania; Robert Rigolizzo, assistantship, Villanove, Pennsylvania.

Future seminar papers include: "Quadratic Forms and Invariant Theory," Robert Minder; "Reduction of the General Quadratic," Michael Young; "Affine Geometry," William Becker; "Fourier Series," Adrian Karsh; "Stereographic Projection," James Crockett; and "Four Color Problem," Edward Keppel.

Pennsylvania Delta, Marywood College, Scranton

New members initiated on May 12, 1966, were: Mary A. Baldini, Ann M. Caporelli, Linda Connor, Mary T. Grace, Suzanne Klassner, Barbara Lachowicz, Linda McDonnell, Ann Meagher, Mary Karen Merkel, Agnes Mullally, Mary L. Palla, Joan Jenkins, and Ann Marie Solancis.

Lectures given: "When Computers Are Useful" by Mr. Lawrence Wheatley of Savannah River Laboratory, South Carolina; "Computer Solution of Matrix Eigen-systems" by Mr. Robert Funderlic of Oak Ridge Laboratory, Tennessee.

A series of six films will also be shown during the current year.

Tennessee Beta, East Tennessee State University, Johnson City

Tennessee Beta Chapter began the year with three well-attended and interesting meetings. The president, Harold Bullock, presided at the business sessions. At the October meeting Professor Tai-il Suh spoke on "Some Ideas of the Projective Plane;" at the November meeting Director of the Computer Center, Stanford H. Johnson, spoke on "Job Opportunities in Data Processing;" at the January meeting Professor Charles Taylor gave a talk on "Meteors." Members enjoyed a snack before each meeting.

The chapter had a spaghetti dinner party on February 9 at the home of Mrs. Lora McCormick, corresponding secretary. Thirty-five members were present and, following the dinner, enjoyed group singing and a program of guitar music and songs presented by Janette Gass and Harold Bullock.

Texas Beta, Southern Methodist University, Dallas

Texas Beta Chapter started this year with a luncheon meeting featuring a talk on "The Calculus of Variations" by Dr. Richard Williams. We had thirty-seven in attendance. We have also had a Halloween Party and an "Information about Graduate School and Coffee." We are planning a meeting in November which will feature a talk by Dr. W. L. Ayres on "The Map Coloring Problem." We have plans to send a delegation to the convention in April.

Texas Epsilon, North Texas State University, Denton

The Texas Epsilon Chapter received twenty applications for membership for the Fall 1966 semester. The pledges have been appointed to various committees to perform useful and needed duties in the Mathematics Building. Among these are responsibilities for the Mathematics Reading Room and constructive bulletin boards.

Wisconsin Alpha, Mount Mary College, Milwaukee

Sandra Mertes, Kappa Mu Epsilon member graduate of 1966, received a National Science Foundation Traineeship for study of mathematics at the Catholic University of America, Washington, D.C.

During the year the following programs were presented: "Number Bases" by Rosemary Reuille; "The Trachtenberg System of Basic Mathematics" by Sue O'Connor; "Panel on Three Mathematicians: Pythagoras, Euclid, and Archimedes" by Karen Mauro, Shirley Bruder, Ellen Levine; "Demonstration on Curve Stitching" by Delores Peirick; "Panel on Computers: Brief History, How They Work, and Applications" by Jerrilyn Foster, Judy Pokrop, and Grace Makarewicz; "Progressions: Arithmetical and Geometric" by Tiia Ostrovskis and Maureen O'Donnell.

Mount Mary College also played host to the Mu Alpha Theta regional meeting on January 7, 1967. The principal speakers were Professor Hannekan, Department Chairman from Marquette University, and William Golomski from industry.

Nine students and one faculty member are planning to attend the National KME Convention, and Karen Johnson is preparing a paper.

(Continued from page 108.)

REFERENCES

1. W. W. R. Ball, *Mathematical Recreations and Essays*, ed. H. S. M. Coexeter (New York: Macmillan, 1939), pp. 153-154.
2. Martin Gardner, "Mathematical Puzzles and Pastimes," *Scientific American*, CCII (May, 1958), pp. 122-123.
3. D. A. Engel, "Flexahedrons," *Recreational Mathematics Magazine*, No. 11 (Oct., 1962), pp. 3-5.

Acknowledgment: Professor Everett Marshall.

The Mathematical Scrapbook

EDITED BY GEORGE R. MACH

Readers are encouraged to submit Scrapbook material to the editor. Material will be used where possible and acknowledgement will be made in *The Pentagon.*

=Δ=

In one of its popular forms, the chain letter operates by having the originator send a fixed number of letters (say five), each receiver copies and sends five, and so on. If it were possible that no person received two such letters, how many "rounds" would it take for every American to receive one? Averaging one "round" a week, could it be done in a year or two?

We want the sum of the series $5^1 + 5^2 + 5^3 + \cdots + 5^n$ to equal our population, say 200,000,000.

$$S = \frac{a(r^n - 1)}{r - 1} = 200{,}000{,}000$$

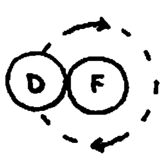
$$\frac{5(5^n - 1)}{4} = 200{,}000{,}000$$

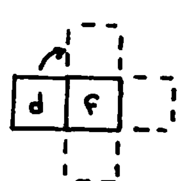
$$5^n - 1 = 160{,}000{,}000$$

Forgetting about the 1 and taking the logarithm of both sides we soon see that three months would be plenty of time to circulate the letters to everyone.

=Δ=

How many revolutions about its own center will a circular disk (D) have made when it rolls around an identical fixed disc (F)? Since the circumferences are the same and the circumference of (D) is rolled out once along that of (F), the usual answer is one revolution, but it is incorrect.





You can try it with coins or easily visualize a similar situation with squares (*d*) and (*f*). Note that (*d*) rotates half a revolution about its own center in just getting the first notch, or one-fourth of the way around (*f*).

Do these two examples have the same answer? Why? What about *n*-sided regular polygons? What about dissimilar figures, for example a square and a hexagon, with the same perimeter? Can any general statements be made about a situation in which one perimeter is an integral multiple of the other?

=△=

How many pairs of rational numbers, (x,y), are there for which $x^y = y^x$? Well, if $x = y$ it is always true and that's not very interesting. Most people quickly see that $2^4 = 4^2$. Are there any other pairs of unequal rational numbers? Is there a finite or an infinite number?

To simplify and analyze the problem in one way, suppose we let $y = rx$, where r is rational and $r \neq 0$. Then,

$$(x)^{rx} = (rx)^x$$

$$(x^r)^x = (rx)^x$$

$$x^r = rx$$

$$x^{r-1} = r$$

$$x = r^{\frac{1}{r-1}}$$

Now, x will be rational whenever $\frac{1}{r-1}$ is an integer, n. Then,

$$n = \frac{1}{r-1}$$

$$r - 1 = \frac{1}{n}$$

$$r = 1 + \frac{1}{n}, \quad n \neq -1, \text{ since } r \neq 0.$$

So, a number pair, (x,y), can be found if $x = (1 + \frac{1}{n})^n$, and $y = (1 + \frac{1}{n})^{n+1}$, for each $n = +1, +2, +3, \cdots$. When

$n = 1$, $(x,y) = (2,4)$; $n = 2$, $(x,y) = \left(\frac{9}{4}, \frac{27}{8}\right)$; etc. This analysis does not give all of the unequal rational number pairs but it does give at least an infinite number of them. Notice that the obvious $x = y$ does not fit this pattern. Can you find additional pairs of rational numbers which do not fit the above pattern for some integer n?

Now, suppose that we want to find pairs of *integers*, (x,y). Using the same approach as before let $y = rx$, where r is an integer and $r \neq 0$. As before, $x = r^{\frac{1}{r-1}}$. Now, x will be an integer whenever $\frac{1}{r-1}$ is an integer, n. As before, $r = 1 + \frac{1}{n}$. But now r and n must both be integers. Apparently $n = 1$, $r = 2$ is the only possibility. This yields the pair $(2,4)$, which we already had. Can you find a pair of integers which does not fit the above pattern? Try $(-2, -4)$. Is there a better approach to the problem than this one?

=Δ=

Editor's note: The following was submitted by Rex L. Hutton, formerly a student member of Ohio Gamma Chapter and now a faculty member of California Gamma Chapter.

A Numeration System and a Problem

We are fairly familiar with various numeration systems including, of course, base 10 and base 2. With just two tokens at our disposal and no digits like the 0, 1 of the binary system, is it possible to count and have a numeration system? Well, one token can be put down for "one" and both for "two." Then it is apparent that some positional scheme must be used. We could denote "three" by picking up both tokens and putting one down in a second position, like the ten's position in the second column on an abacus. This could be done with tally marks as well as tokens.

Consider this mixture of tally and positional numeration. New numerals are subscripted with H.

Base 10 Numeral	Tally					New Numeral
1					/	1_H
2					//	2_H
3				/		10_H
4				/	/	11_H
5				//		20_H
6			/			100_H
7			/		/	101_H
8			/	/		110_H
9			//			200_H
10		/				1000_H

Now, extending the numeration system we would continue:

$11 = 1001_H$	$14 = 2000_H$	$17 = 10010_H$	$20 = 20000_H$
$12 = 1010_H$	$15 = 10000_H$	$18 = 10100_H$	$21 = 100000_H$
$13 = 1100_H$	$16 = 10001_H$	$10 = 11000_H$	etc.

Note that the use of the digits, 0 and 1, is not required at all but is just a convenient way to denote the tally.

Observe that the numerals

$$1_H,\ 10_H,\ 100_H,\ 1000_H,\ 10000_H,\ 100000_H,\ \cdots$$

represent numbers of the form

$$\frac{n(n+1)}{2},$$

where n is a natural number.

Further, observe the correspondence between n and the position of the 1 in the H numeral.

Also notice, in a numeral such as 10010_H the rightmost 1 acts as a simple tally mark, its position giving its value.

These observations make it relatively easy to convert an H numeral to base 10. Examples:

$$10010_H = \frac{(5)(6)}{2} + 2 = 17$$

$$1001_H = \frac{(4)(5)}{2} + 1 = 11$$

$$1001000_H = \frac{(7)(8)}{2} + 4 = 32$$

$$200_H = \frac{(3)(4)}{2} + 3 = 9$$

$$20000_H = \frac{(5)(6)}{2} + 5 = 20$$

Now, consider writing the H numeral for a certain natural number, $k > 0$. The leftmost digit of our H numeral represents $\frac{n(n + 1)}{2}$ for some natural number n. We desire to determine the largest natural number n such that

$$\frac{n(n + 1)}{2} \leqq k.$$

We see that

$$n^2 + n \leqq 2k$$
$$n^2 < 2k,$$

thus $\sqrt{2k}$ is an estimate for n.

To write 56 as an H numeral, we consider $\sqrt{(2)(56)} = \sqrt{112} \approx 10$. We try $n = 10$.

$$\frac{(10)(11)}{2} = 55$$

and $55 + 1 = 56$, so that H numeral is 1000000001_H. Note that in some cases $\sqrt{2k}$ may be too large and so we try $\sqrt{2k} - 1$.

Problem: How can we compute with H numerals? For addition, the problem is to write a sum such as

$$\frac{n(n + 1)}{2} + m + \frac{k(k + 1)}{2} + j,$$

where $m \leqq n$ and $j \leqq k$, in the form

$$\frac{t(t+1)}{2} + s,$$

where $s \leq t$.

Example: $1001_H + 1100_H = \frac{4(4+1)}{2} + 1 + \frac{4(4+1)}{2} + 3$

$$= 10 + 1 + 10 + 3$$

$$= 24$$

$$= 21 + 3$$

$$= \frac{6(6+1)}{2} + 3$$

$$= 100100_H.$$

How can this and other sums be obtained without base 10 computations? How can system H products be computed?

What kind of a numeration system can be devised with three tokens?

(Continued from page 94.)

10. Conclusion. In this paper we have attempted to demonstrate how the multiplication of a matrix by a scalar can be generalized in certain cases of non-conformal matrices. The final section shows how the defined X-multiplication reduces to scalar multiplication of a matrix on one hand and the ordinary matrix multiplication on the other. A few analogies of standard matrix operations have been given as examples of X-multiplication — the reader may be interested in examining others.

REFERENCES

1. A. C. Aitken, *Determinants and Matrices,* Oliver and Boyd, Edinburgh, 1939.
2. Richard Bellman, *Introduction to Matrix Analysis,* McGraw-Hill, New Work, 1960.
3. I. N. Herstein, *Topics in Algebra,* Blaisdell, New York, 1964.