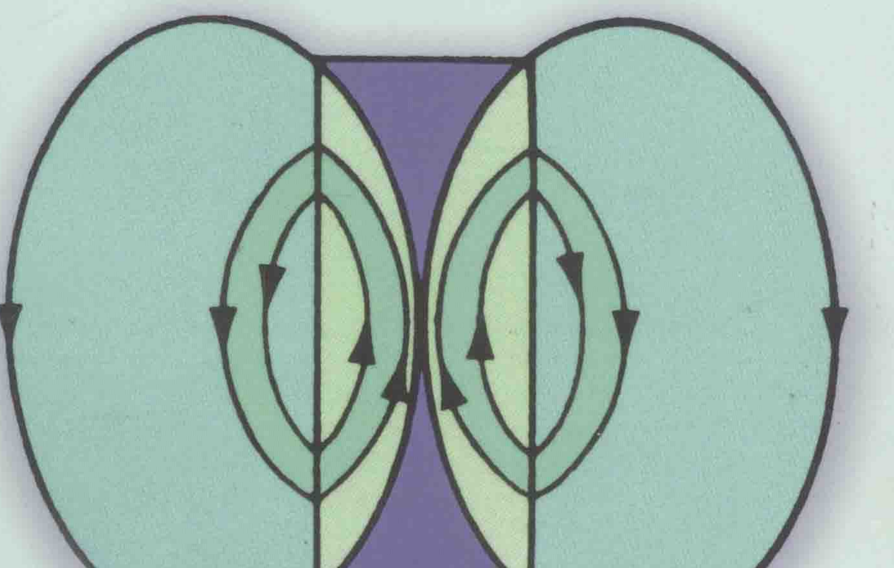
Melvin Schwartz



PRINCIPLES OF ELECTRODYNAMICS

Principles of Electrodynamics

MELVIN SCHWARTZ
Consulting Professor of Physics
Stanford University

DOVER PUBLICATIONS, INC.
New York

TO MY WIFE, MARILYN

Bibliographical Note

This Dover edition, first published in 1987, is an unabridged, corrected republication of the work first published by the McGraw-Hill Book Company, New York, 1972, in its International Series in Pure and Applied Physics.

Drawings by Judith L. McCarty

Library of Congress Cataloging-in-Publication Data

Schwartz, Melvin, 1932–
Principles of electrodynamics.

". . . unabridged, corrected republication of the work first published by the McGraw-Hill Book Company, New York, 1972 . . ."T.p. verso.
Includes index.
1. Electrodynamics. I. Title.
QC631.S38 1987 537.6 87-13607
ISBN-13: 978-0-486-65493-5
ISBN-10: 0-486-65493-1

Manufactured in the United States by Courier Corporation
65493112
www.doverpublications.com

Preface

Electromagnetic theory is beautiful! When looked at from the relativistic point of view where electric and magnetic fields are really different aspects of the same physical quantity, it exhibits an aesthetically pleasing structure which has served as a model for much of modern theoretical physics. Unfortunately this beauty has been all but buried as most textbooks have treated electricity, magnetism, Coulomb's law, and Faraday's law as almost completely independent subjects with the ground work always supplied by means of empirical or historical example. Occasionally a chapter is devoted to the relativistic coalescence of the various aspects of electromagnetism but use is rarely made of the requirement of Lorentz invariance in deriving the fundamental laws.

Our point of view here is quite different. Basically we have two purposes in mind—one is to exhibit the essential unity of electromagnetism in its natural, relativistic framework and the other is to show how powerful the constraint of relativistic invariance is. To these ends we shall show that all electromagnetism follows from electrostatics and the requirement that our laws be the simplest ones allowable under the relativistic constraint. The hope is that the student will make use of these new insights in thinking about theories that are as yet undeveloped and that the model we set here will be generally useful in other areas of physics.

A word about units. Unfortunately one of the results of the completely disconnected way in which electricity and magnetism have been taught in the past has been the growing acceptance of the mks over the cgs system of units. We have no special preference for centimeters over meters or of grams over kilograms. We do, however, require a system wherein the electric field **E** and the magnetic field **B** are in the *same* units. Using the mks system, as it is presently constituted, for electromagnetic theory is akin to using a meterstick to measure along an East-West line and a yardstick to measure along a North-South line. To measure **E** and **B** in different units is completely antithetical to the entire notion of relativistic invariance. Accordingly we will make use of the cgs (gaussian) system of units exclusively. Conversion to practical units where necessary can be carried out with no difficulty.

The author would like to express his most profound appreciation to Miss Margaret Hazzard for her patient and careful typing of the text.

MELVIN SCHWARTZ

Contents

1
Mathematical Review and Survey of Some New Mathematical Ideas

It would be delightful if we could start right out doing physics without the need for a mathematical introduction. Unfortunately though, this would make much of our work immeasurably more laborious. Mathematics is much more than a language for dealing with the physical world. It is a source of models and abstractions which will enable us to obtain amazing new insights into the way in which nature operates. Indeed, the beauty and elegance of the physical laws themselves are only apparent when expressed in the appropriate mathematical framework.

We shall try to cover a fair bit of the mathematics we will need in this introductory chapter. Several subjects are, however, best treated within the context of our physical development and will be covered later. It is assumed that the reader has a working familiarity with elementary calculus, three-dimensional vectors, and the complex number system. All other subjects will be developed as we go along.

1-1 VECTORS IN THREE DIMENSIONS: A REVIEW OF ELEMENTARY NOTIONS

We begin by reviewing what we have already learned about three-dimensional vectors. As we remember from our elementary physics, there are a large number of quantities that need three components for their specification. Position is, of course, the simplest of these quantities. Others include velocity and acceleration. Even though we rarely defined what was meant by a vector in mathematically rigorous terms, we were able to develop a certain fluency in dealing with them. For example, we learned to add two vectors by adding their components. That is, if $\mathbf{r}_1 = (x_1,y_1,z_1)$ and $\mathbf{r}_2 = (x_2,y_2,z_2)$ are two vectors, then

$$\mathbf{r}_1 + \mathbf{r}_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

If a is a number, then

$$a\mathbf{r}_1 = (ax_1,ay_1,az_1)$$

We also found it convenient to represent a vector by means of an arrow whose magnitude was equal to the vector magnitude and whose direction was the vector direction. Doing this permitted us to add two vectors by placing the "tail" of one at the "head" of the other as in Fig. 1-1. We also learned how to obtain a so-called scalar quantity by carrying out a type of multiplication with two vectors. If $\mathbf{r}_1 = (x_1,y_1,z_1)$ and $\mathbf{r}_2 = (x_2,y_2,z_2)$ are two vectors, then $\mathbf{r}_1 \cdot \mathbf{r}_2$ is defined by the equation

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = x_1x_2 + y_1y_2 + z_1z_2$$

It was also shown that $\mathbf{r}_1 \cdot \mathbf{r}_2$ could be obtained by evaluating $|\mathbf{r}_1|\,|\mathbf{r}_2|\cos\theta_{12}$, where $|\mathbf{r}_1|$ and $|\mathbf{r}_2|$ are, respectively, the magnitudes of $\mathbf{r}_1$ and $\mathbf{r}_2$ and θ_{12}

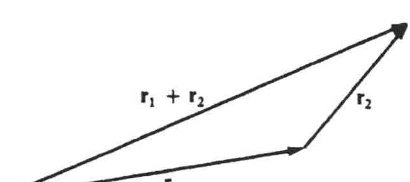


Fig. 1-1 The addition of two vectors can be accomplished by placing the "tail" of one at the "head" of the other.

is the angle between them. Another so-called vector was obtained by taking the cross product of $\mathbf{r}_1$ and $\mathbf{r}_2$. That is,

$$\mathbf{r}_1 \times \mathbf{r}_2 = (y_1 z_2 - y_2 z_1, z_1 x_2 - z_2 x_1, x_1 y_2 - y_1 x_2)$$

We shall have much more to say about the true nature of this beast very shortly. At the moment we just recall that it appears in some respects to be a vector whose magnitude is equal to $|\mathbf{r}_1||\mathbf{r}_2|\sin\theta_{12}$ and whose direction, at right angles to both $\mathbf{r}_1$ and $\mathbf{r}_2$, is given by a so-called **right-hand rule** in going from $\mathbf{r}_1$ to $\mathbf{r}_2$. If we look from the head toward the tail of $\mathbf{r}_1 \times \mathbf{r}_2$, we would see the shortest rotation from $\mathbf{r}_1$ to $\mathbf{r}_2$ to be in the counterclockwise direction.

Unfortunately, we shall have to relearn much of the above within a more abstract framework if we are to make any progress beyond this point. We shall have to go back to our basic notions and see if we can define what we mean by vector in a more suitable, less intuitive manner. Only by doing so will we be prepared to say clearly which combinations of three numbers are vectors and which are not. We will also be able to define scalar in a reasonable way and will then see our way clear to an understanding of higher-rank tensors.

1-2 THE TRANSFORMATION PROPERTIES OF VECTORS UNDER SPATIAL ROTATION

To open the way for a more rigorous definition of vector, we proceed a bit further with our old intuitive notions. Let us consider a so-called **position vector**, that is, a vector from the origin of our coordinate system to the point (x,y,z). If we draw a unit vector along each of the three axes as shown in Fig. 1-2 and call them $\hat{\mathbf{i}}, \hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$, respectively, we can write $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$. Now, we ask, what if we were to rotate our coordinate system to a new set of axes x', y', and z' with a new set of unit vectors $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, and $\hat{\mathbf{k}}'$? How would $\mathbf{r}$ be expressed now? We answer this question very simply by expressing $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ in terms of the new unit vectors $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, and $\hat{\mathbf{k}}'$. (This is possible

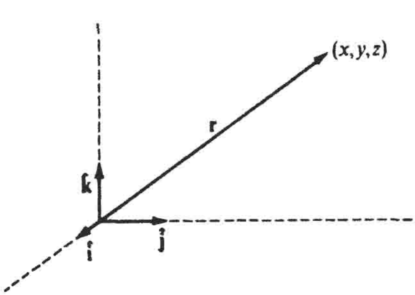


Fig. 1-2 The vector $\mathbf{r}$ can be expressed as $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are unit vectors along the x, y, z axes.

because any three-dimensional vector whatsoever can be expressed either in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ or in terms of $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, and $\hat{\mathbf{k}}'$.) We write

$$\begin{aligned}
\hat{\mathbf{i}} &= a_{11}\hat{\mathbf{i}}' + a_{21}\hat{\mathbf{j}}' + a_{31}\hat{\mathbf{k}}' \\
\hat{\mathbf{j}} &= a_{12}\hat{\mathbf{i}}' + a_{22}\hat{\mathbf{j}}' + a_{32}\hat{\mathbf{k}}' \\
\hat{\mathbf{k}} &= a_{13}\hat{\mathbf{i}}' + a_{23}\hat{\mathbf{j}}' + a_{33}\hat{\mathbf{k}}'
\end{aligned} \tag{1-2-1}$$

We note the obvious fact that

$$\begin{array}{lll}
\hat{\mathbf{i}}\cdot\hat{\mathbf{i}}' = a_{11} & \hat{\mathbf{i}}\cdot\hat{\mathbf{j}}' = a_{21} & \hat{\mathbf{i}}\cdot\hat{\mathbf{k}}' = a_{31} \\
\hat{\mathbf{j}}\cdot\hat{\mathbf{i}}' = a_{12} & \hat{\mathbf{j}}\cdot\hat{\mathbf{j}}' = a_{22} & \hat{\mathbf{j}}\cdot\hat{\mathbf{k}}' = a_{32} \\
\hat{\mathbf{k}}\cdot\hat{\mathbf{i}}' = a_{13} & \hat{\mathbf{k}}\cdot\hat{\mathbf{j}}' = a_{23} & \hat{\mathbf{k}}\cdot\hat{\mathbf{k}}' = a_{33}
\end{array}$$

This, of course, permits us immediately to express the unit vectors $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, and $\hat{\mathbf{k}}'$ in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$, viz.,

$$\begin{aligned}
\hat{\mathbf{i}}' &= a_{11}\hat{\mathbf{i}} + a_{12}\hat{\mathbf{j}} + a_{13}\hat{\mathbf{k}} \\
\hat{\mathbf{j}}' &= a_{21}\hat{\mathbf{i}} + a_{22}\hat{\mathbf{j}} + a_{23}\hat{\mathbf{k}} \\
\hat{\mathbf{k}}' &= a_{31}\hat{\mathbf{i}} + a_{32}\hat{\mathbf{j}} + a_{33}\hat{\mathbf{k}}
\end{aligned} \tag{1-2-2}$$

We realize that not all the nine quantities a_{ij} can be chosen independently. After all, only three angles are necessary to specify the rotation of one coordinate system into another. We expect then to have six equations linking the coefficients. We obtain these equations by requiring that $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, and $\hat{\mathbf{k}}'$ form an orthogonal set of unit vectors.

$$\begin{aligned}
\hat{\mathbf{i}}'\cdot\hat{\mathbf{i}}' &= 1 = a_{11}^2 + a_{12}^2 + a_{13}^2 \\
\hat{\mathbf{j}}'\cdot\hat{\mathbf{j}}' &= 1 = a_{21}^2 + a_{22}^2 + a_{23}^2 \\
\hat{\mathbf{k}}'\cdot\hat{\mathbf{k}}' &= 1 = a_{31}^2 + a_{32}^2 + a_{33}^2 \\
\hat{\mathbf{i}}'\cdot\hat{\mathbf{j}}' &= 0 = a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} \\
\hat{\mathbf{j}}'\cdot\hat{\mathbf{k}}' &= 0 = a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} \\
\hat{\mathbf{k}}'\cdot\hat{\mathbf{i}}' &= 0 = a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33}
\end{aligned} \tag{1-2-3}$$

Now to return to our original vector $\mathbf{r}$. We can write $\mathbf{r}$ in terms of its components in either of two ways:

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} \qquad \text{or} \qquad \mathbf{r} = x'\hat{\mathbf{i}}' + y'\hat{\mathbf{j}}' + z'\hat{\mathbf{k}}'$$

Making use of Eqs. (1-2-1), we find immediately that

$$\begin{aligned}
x' &= a_{11}x + a_{12}y + a_{13}z \\
y' &= a_{21}x + a_{22}y + a_{23}z \\
z' &= a_{31}x + a_{32}y + a_{33}z
\end{aligned} \tag{1-2-4}$$

We have traditionally used a right-handed coordinate system to specify the components of a vector. That is to say we have chosen $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ so that if we curl the fingers of our right hand from $\hat{\mathbf{i}}$ to $\hat{\mathbf{j}}$, our thumb will point along $\hat{\mathbf{k}}$. Expressing this in language somewhat more abstract and less anthropomorphic, we can say that $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$ in such a system. Obviously there is nothing in nature that requires us to limit ourselves to right-handed coordinate systems, and we might ask if there is anything special about the set of numbers a_{ij} if the primed system should happen to be a left-handed system. For a left-handed system we can write

$$(\hat{\mathbf{i}}' \times \hat{\mathbf{j}}') \cdot \hat{\mathbf{k}}' = -1 \tag{1-2-5}$$

Expressing $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, and $\hat{\mathbf{k}}'$ in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$, we can rewrite this equation as follows:

$$[(a_{11}\hat{\mathbf{i}} + a_{12}\hat{\mathbf{j}} + a_{13}\hat{\mathbf{k}}) \times (a_{21}\hat{\mathbf{i}} + a_{22}\hat{\mathbf{j}} + a_{23}\hat{\mathbf{k}})] \cdot (a_{31}\hat{\mathbf{i}} + a_{32}\hat{\mathbf{j}} + a_{33}\hat{\mathbf{k}}) = -1$$

Carrying out the indicated multiplications, we find

$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) = -1 \tag{1-2-6}$$

The expression on the left of Eq. (1-2-6) is called the **determinant** of the matrix of numbers a_{ij} or det a_{ij} for short. It is often written in the notation

$$\det a_{ij} = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

We see then that any transformation that takes us from a right-handed coordinate system to a left-handed coordinate system is characterized by having its determinant equal to -1. Indeed, as we can easily see, the determinant is equal to -1 whenever we change the handedness of our system and $+1$ if we keep it unchanged. By allowing transformations with either sign of determinant, we allow ourselves to deal with both rotations and reflections or with any combination of these transformations.

We have begun to think of our transformation as having an "identity" all its own. It is characterized by a set of nine numbers, which we have called a **matrix.** Furthermore we have seen in Eq. (1-2-4) that we can obtain the triplet (x',y',z') by "multiplying" the triplet (x,y,z) by this matrix, with the operation of multiplication being defined as

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{bmatrix} \tag{1-2-7}$$

We can represent the above operation symbolically by writing

$$\mathbf{r}' = \mathbf{a}\mathbf{r} \tag{1-2-8}$$

(In the future, a boldface sans serif symbol, such as **a**, will mean that the symbol is a matrix and not a number.)

Suppose now that we wish to undertake two successive transformations, the first characterized by **a** and the second by another matrix **b**. If we begin with the triplet **r**, then the first transformation leads to the triplet **r**′ and the second to the triplet **r**″. That is,

$$\mathbf{r}' = \mathbf{a}\mathbf{r}$$
$$\mathbf{r}'' = \mathbf{b}\mathbf{r}'$$

Alternatively, we might have gone directly from the unprimed to the double-primed coordinate system by means of a transformation **c**.

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

Writing out these transformations in detail will show that we could determine all the elements of **c** directly from **a** and **b** by means of the simple set of equations

$$c_{11} = b_{11}a_{11} + b_{12}a_{21} + b_{13}a_{31}$$
$$c_{12} = b_{11}a_{12} + b_{12}a_{22} + b_{13}a_{32}$$

or, in general,

$$c_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + b_{i3}a_{3j}$$

We abbreviate this in the customary way by writing

$$c_{ij} = \sum_{k=1}^{3} b_{ik}a_{kj} \tag{1-2-9}$$

Thus the element c_{ij} can be obtained by taking the "scalar product," so to speak, of the ith row in **b** with the jth column in **a**.

The operation which we have defined above in Eq. (1-2-9) is called the **product of two matrices a and b** and can be represented by the expression **c** = **ba**. Matrix multiplication, unlike the multiplication of two numbers, is *not* in general commutative, as the reader can very easily convince himself. That is to say the product **ab** is not in general equal to the product **ba**. Multiplication is, however, associative. This means that we can in general write, for three transformations **a**, **b**, and **c**,

$$\mathbf{a}(\mathbf{b}\mathbf{c}) = (\mathbf{a}\mathbf{b})\mathbf{c} \tag{1-2-10}$$

To complete our picture we should point out that one of the possible transformations is the identity transformation which leaves the coordinate system unchanged. We write this matrix as **1** with the observation that

$$\mathbf{1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1-2-11}$$

Returning back to Eqs. (1-2-1) and (1-2-2), we see that for every transformation $\mathbf{a}$ there is also an inverse transformation $\mathbf{a}^{-1}$ such that

$$\mathbf{a}\mathbf{a}^{-1} = \mathbf{a}^{-1}\mathbf{a} = \mathbf{1} \tag{1-2-12}$$

The inverse transformation is just given by the **transposed** matrix. That is to say

$$a_{ij}^{-1} = a_{ji} \tag{1-2-13}$$

(For those whose mathematical sophistication is just a bit above average, we might point out that the set of all transformations defined above constitute what is known in the trade as a **group.** The detailed properties of groups play an important role in the development of much of quantum mechanics and should be studied at the earliest possible moment by those who intend to extend their horizons in physics beyond the classical domain.)

We can now think in terms of the complete set of all transformations from one orthogonal coordinate system to another, including within our set both rotations (det $\mathbf{a} = +1$) and reflections (det $\mathbf{a} = -1$). The definition of scalar, vector, and various other entities is now best done in terms of this set of transformations.

Let us begin with what is intuitively the simplest of these entities, the scalar. Imagine that we are given a set of explicit instructions for determining some number. We follow these instructions scrupulously, coming up with a value for the number. We can now rotate our coordinate system or change its handedness (by means of the transformation $\mathbf{a}$). If the same set of rules for determining the number leads to the same result in the new system, regardless of the choice of rotation or reflection, then the number is a **scalar.**

Obviously there are innumerable trivial examples of scalars that we can readily cite. The number of cents in the dollar or the number of fingers on your hand have nothing to do with the coordinate system and hence are ipso facto scalars. Much less trivial, though, are numbers that are derived by means of rules which concern coordinates themselves. Let us take a simple example.

Suppose the rule tells us to take the x coordinate of a point, square it, add to that the square of the y coordinate of the same point, and add to the sum the square of the z coordinate of the point. We would have then a number equal to $x^2 + y^2 + z^2$. If we transform to a new system and follow the same prescription in the new system, we come up with $x'^2 + y'^2 + z'^2$. Unless we knew the Pythagorean theorem we would have no a priori expectation that the same rule applied in these different systems would give us the same result. Indeed it does because we have just determined the square of the distance from our point to the origin, and that quantity does not depend on the rotational orientation or the handedness of our system. Clearly then the number $x^2 + y^2 + z^2$ is a scalar.

Let us try a more difficult example now. Consider two points whose coordinates in one system are (x_1,y_1,z_1) and (x_2,y_2,z_2). We can form the expression $x_1x_2 + y_1y_2 + z_1z_2$ and evaluate it in this coordinate system. We can now transform coordinates and evaluate the same expression in the new system, obtaining $x'_1x'_2 + y'_1y'_2 + z'_1z'_2$. Again we have no a priori expectation that the two numbers will come out to be the same. Making use of Eqs. (1-2-4) and (1-2-3), the reader can easily convince himself that this is, however, the case—the numbers are the same and so the expression $x_1x_2 + y_1y_2 + z_1z_2$ is a scalar. (The result is not entirely unanticipated for we remember that this expression is the scalar product of $\mathbf{r}_1$ and $\mathbf{r}_2$ and can also be written as $|\mathbf{r}_1||\mathbf{r}_2|\cos\theta$. The latter formula does not depend on the coordinate system.)

There is a great temptation now to let every "constant" of nature, like charge and mass, be labeled a scalar. In fact we must be exceedingly careful since an attribute like charge is defined operationally in terms of forces by external fields, and we must investigate the behavior of the entire system under *both* rotation and reflection before we can conclude that the attribute is a scalar. We shall have more to say about this very shortly.

We go on now to the definition of another important entity, the **pseudoscalar.** The pseudoscalar differs from the scalar in only one important respect. The sign of the number we obtain by following our prescription in a left-handed coordinate system is *opposite* to that we obtain in a right-handed system. For pure rotations, scalars and pseudoscalars behave identically.

To find an example of a pseudoscalar is not difficult at all. Let us take three points in space which in one coordinate system have the components (x_1,y_1,z_1), (x_2,y_2,z_2), and (x_3,y_3,z_3). We can construct a determinant D out of these nine numbers:

$$D = \begin{Vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{Vmatrix} = x_1(y_2z_3 - y_3z_2) + y_1(z_2x_3 - z_3x_2) + z_1(x_2y_3 - x_3y_2) \qquad \text{(1-2-14)}$$

(It is quite clear that D is equal to $\mathbf{r}_1 \cdot (\mathbf{r}_2 \times \mathbf{r}_3)$ and has magnitude equal to the volume of the parallelopiped determined by $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$.) If we