Mathematical Spectrum

Volume 3 1970/71 *Number 1*

A Magazine of *Contemporary Mathematics*

Published by *Oxford University Press*

Mathematical Spectrum is a magazine for the instruction and entertainment of student mathematicians in schools, colleges, and universities. It is published by Oxford University Press on behalf of the Applied Probability Trust, a non-profit making organization established in 1963 with the support of the London Mathematical Society. The object of the Trust is the encouragement of study and research in the mathematical sciences.

Volume 3 of *Mathematical Spectrum* (1970/71) will consist of two issues, the second of which will be published in the Spring of 1971.

Articles published in *Mathematical Spectrum* deal with the entire range of mathematical disciplines (pure mathematics, applied mathematics, statistics, operational research, computing science, numerical analysis, biomathematics). Both expository and historical material may be included, as well as elementary research and information on educational opportunities and careers in mathematics. There is also a section devoted to problems.

Articles are normally commissioned by the Editors; however, the Editorial Committee welcomes the submission of suitable material, including correspondence, queries and solutions to problems, for publication in *Mathematical Spectrum*. All correspondence about the contents should be sent to:

The Managing Editor, *Mathematical Spectrum*,
Hicks Building, The University, Sheffield S3 7RH.

Statistical Inference

J. KIEFER[1]

Cornell University, Ithaca, N.Y.

The word 'statistics' conjures up many meanings. It suggests the state of the economy, Detroit's quality-control methods, medical evidence linking lung cancer with smoking, and the physicist's configurations in statistical mechanics. In this paper, our interest is limited to *statistical inference*. This is the study of various possible procedures for analyzing data in order to guess the nature of the physical or biological mechanism that produced these data. It is an applied area of mathematics with particular conceptual difficulties not found in other applied areas. The mathematical theory tries to give a rationale for selecting a procedure for analyzing the data rather than relying on intuition or rule of thumb. We cannot hope to develop any deep theoretical results of the subject here. We hope to treat only a few examples that may give you some feeling for subtleties in even the simplest statistical problems where practitioners tread—usually fearlessly, frequently by intuition, and quite often disastrously.

How does statistical inference differ from probability theory? In probability theory one specifies a model (chance mechanism) and studies its consequences. For example, suppose a coin has probability p of coming up 'heads' on a single toss, and $1-p$ of coming up 'tails'. Here p is a number between 0 and 1 that reflects the physical makeup of the coin; roughly speaking, it is the approximate proportion of heads you would expect in a long series of tosses, and it would be $1/2$ for a 'fair coin'. Our model is to flip the coin five times, with independent flips. 'Independence' means that successive flips have no probabilistic effect on each other, so that the probability of obtaining a particular sequence of outcomes on the five tosses is computed by multiplying the probabilities of the outcomes for the individual tosses. Thus, the probability of obtaining 'heads, tails, heads, tails, tails', which we hereafter abbreviate HTHTT, is

$$p \times (1-p) \times p \times (1-p) \times (1-p) = p^2(1-p)^3. \tag{1}$$

Clearly, we obtain the same probability $p^2(1-p)^3$ for any other sequence of five tosses containing exactly two H's, for example TTHHT. There are ten different

[1] Reprinted from *The Mathematical Sciences*, a Collection of Essays edited by the Committee on Support of Research in the Mathematical Sciences with the Collaboration of George A. W. Boehm, by permission of The M.I.T. Press, Cambridge, Massachusetts.

arrangements of two H's and three T's, and the event 'two heads in five tosses' occurs if the five flips come up in any of these ten arrangements, each of probability $p^2(1-p)^3$. Thus, the event 'two heads in five tosses' has probability $10p^2(1-p)^3$. Similarly, the probability of obtaining exactly k heads in five tosses, hereafter denoted $b_p(k)$, is

$$b_p(k) = \frac{5!}{k!(5-k)!}p^k(1-p)^{5-k} \tag{2}$$

if $k = 0, 1, 2, 3, 4$, or 5. (We have used the convention $0! = 1$.) This is the well-known binomial probability law, and of course $b_p(0)+b_p(1)+\ldots+b_p(5) = 1$ since this sum includes the probabilities of all possible sequences of five flips. The development leading to equation 2, although very elementary, illustrates the computation of a consequence of a simple probability model, which typifies probability theory.

In contrast, in a statistical-inference model the chance mechanism is not completely known. Rather, the problem is to guess some of its features. A simple example is that of our coin when the value of p is unknown. The mint, it may be supposed, produces coins which, because of inhomogeneous makeup and irregular shape, range from those that almost always produce heads when flipped (p near 1) through those that produce about half heads (p near 1/2) to those that almost always produce tails (p near 0). It is not known which of these types of coin is the one we actually have, that is, which value of p between 0 and 1 characterizes the coin.

The emphasis of statistical inference is on choosing a procedure for using the observed outcome of the chance mechanism to make a guess regarding which of the *possible* mechanisms is the *actual* one at hand. In our example, this means using the observed sequence of five flips to make a guess as to the unknown value of p for our particular coin. In some reasonable sense, this guess should be as accurate as possible. However no procedure can guarantee to guess correctly. Any procedure we decide to use to make a guess about p will make the same guess in every experiment of five flips with outcome HTHTT (though perhaps some other guess whenever the outcome is TTHHH). Whatever the actual p is, the outcome HTHTT, and thus the guess obtained from it, really can arise. But it is easily seen that no meaningful statement made as this guess can be correct for all possible p. For example, if from HTHTT with two H's in five flips, we stated 'I guess that p is 2/5', this would be correct if p were actually 2/5 and would be reasonably accurate if p were close to 2/5, but it would be an erroneous guess, for most practical purposes, if the coin were actually one for which $p = \cdot 9$. Similarly, any other procedure we could use can sometimes make bad guesses. As we shall see, the accuracy of different guessing procedures is a probabilistic computation.

What's the question?

In precisely formulating a statistical problem about our coin, we begin by listing the possible relevant statements we could make about the coin. We shall decide on one of these statements after observing the outcome of the five tosses. This list can be regarded as the list of possible answers to a question about the nature of the coin.

For example, one might ask (a) 'What value of p characterizes the coin at hand?' The possible answers are all real numbers between 0 and 1, and a statistical procedure for this question is a rule which associates, with each of the 32 possible sequences of five tosses, a corresponding guess as to the value of p.

A question requiring a less precise statement about p is (b) 'Is the coin fair?' A statistical procedure for this question does not assert a numerical guess as to the value of p but, rather, makes one of the statements: 'The coin is fair ($p = 1/2$)' or 'The coin is unfair ($p \neq 1/2$)'. If the mint wanted to test each coin it produced, remelting unfair ones to preclude their falling into the hands of scheming gamblers, this second question would be of interest; a more precise guess about the value of p for a given coin would be irrelevant. On the other hand, if the H and T of the coin are replaced by the life or death of a dying patient to whom a new drug is administered, then a numerical guess as to the probability p that the drug can save a life—an answer to question (a)—would often be called for.

You can think of other possible relevant questions. For example, a gambler would want to ask (c) 'Is the coin fair, biased in favor of H, or biased in favor of T?' The answer tells him what use he can profitably make of a given coin: He can bet on the favored outcome in gambling games with unsuspecting victims if the coin is biased, or spend it if it is fair and, thus, has no such nefarious use. Of course, there are many similar examples of greater practical importance.

It is remarkable that, as natural as it may be to ask questions like (c), the classical development of statistics neglected such questions almost completely until Abraham Wald began his work on 'statistical decision theory' in 1939 and emphasized the completely general nature of the questions one could ask. Prior to that, statisticians tended to try to push every problem into the formulation (a), called *point estimation*, or (b), called *hypothesis testing*, whether or not one of these formulations fitted the need of the customer. (A variant of (a), called *interval estimation*, was also sometimes used; here the guess took the form 'I guess p to be in the interval from ·37 to ·45' instead of 'I guess p to be ·4'.)

Here is a simple example of what happened when the wrong question was asked. An experimenter might be testing the productivity of several varieties of grain, wanting to decide what progress he had made in developing a new strain. The actual yield from any plant differs by a chance value, because of such effects as soil variation, from the 'expected yield' for that variety, which is approximately the average yield per plant one would observe over a great many plants. The experimenter who felt he had to ask a question of type (a) or (b) would usually choose the latter and ask, 'Are any of these varieties better than the standard variety most farmers now use?' If the data made him answer 'No', he would return to the laboratory to try to develop other strains. But, if the answer were 'Yes', he would realize he had asked the wrong question, for he would now want to know *which* of the varieties were superior ones in order to experiment further with these. He would therefore ask other questions of type (b), based on the same data: 'Is variety number 1 a superior one? Are varieties 2 and 3 superior?' And so on. He would then try to combine the answers (which could even turn out to be logically incompatible) to reach a conclusion as to which varieties merited further study. By using such an *ad hoc* combination of statistical procedures which were designed to answer questions of type (b) rather than the question of real interest to him, he could end up making a rather inefficient use of the data. Moreover, he almost never knew the actual efficacy of this conglomerate procedure, that is to say, the probability that it would actually select just those varieties with expected yields appreciably greater than that of the standard variety.

Wald's approach departed from the tradition of artificially restricting attention to questions of types (a) and (b) and instead tried to ask the question of real concern. In our example that might be 'Which varieties have expected yield at least 10 percent greater than that of the standard variety?' A procedure would then be designed to answer this question, with efficient use of the data and with a computation of the efficacy of the procedure in giving a correct answer.

Another interesting feature in the history of statistics is that, even in settings where a question of type (a) or (b) is appropriate, it has been less than forty years since statisticians worried much about making precise a reasonable measure of accuracy for statistical procedures and found how to construct procedures that use the data efficiently in terms of that measure. At the beginning of this century Karl Pearson and his school were the leading producers of statistical procedures, constructed largely on an intuitive basis, whose dangers we shall illustrate later in an estimation example. Beginning in 1912, R. A. Fisher showed how inefficient some of these intuitively appealing procedures could be. Over the next half century, Fisher contributed greatly to many mathematical developments in statistics and perhaps had more influence, especially on applied statisticians, than any other person. But some features of his work were unsatisfactory to those who felt that statistical inference should be based on a precise mathematical model with these two features: It makes explicit the penalties that can be incurred from reaching incorrect conclusions from the data, and it leads to the construction and use of a procedure that probably incurs small penalties by using the data efficiently. Jerzy Neyman, together with Karl Pearson's son Egon, was at the forefront of the resulting beginning of the modern mathematical theory of statistical inference.

The cost of being wrong

Ideally, the precise formulation of a statistical-inference model must include not only the list of possible answers to the question asked but also a statement of the relative harm of making various incorrect answers. For example, for the question (a) of estimating the value p characterizing our coin, the loss incurred from misestimating p will presumably be larger, the further the guess is from the actual p. For our grain example of type (c), you may want to try to formulate, at least qualitatively, the way the penalty for an incorrect guess might reasonably depend on the actual average yields and the guess made. We shall simplify our subsequent calculations in this article by neglecting the precise penalty values. In particular, all incorrect guesses are regarded as equally serious in the hypothesis-testing example treated later. However, these values do, in fact, play an important part in determining which statistical procedures are good ones.

From a practical point of view, these penalty values, to be expressed in units of money or utility, are very hard even to approximate in most problems. A manufacturer may be able to assess the cost he would incur from misclassifying a defective light bulb as 'good' in a quality control test. But what is the cost to an astronomer of misestimating by 20 percent the distance to a quasar he is studying? To go a step backward, in exploratory research one cannot always know in advance of the experiment the precise form of the question to be asked. Often a phenomenon shows up which has not occurred as a possibility to the experimenter. (The inviting practice of letting the data determine the question after the experiment and of answering it from these same data can lead to dangerous delusions about the accuracy of one's conclusions.) To go still another step backward in our pattern of formulation, it is often impossible to delimit precisely the class of *possible* probability mechanisms, one of which actually governs the experiment at hand.

What, then, is the worth of the theoretical developments of Neyman and Wald to the practical statistician? The answer is the same as in other areas of applied mathematics and science, where the careful study of a model that is not exactly correct can still lead to more useful decision-making or predictive procedures than will a formula or rule based on intuition or traditional rule of thumb. Moreover, the attempt to write down a precise model is often, in itself, remarkably helpful in clarifying the experimenter's thoughts about his problem. Incidentally, in view of the possible inaccuracy of the model, an important topic is the study of properties of a statistical procedure designed for use with one class of probability mechanisms when, because of incorrect formulation of the model, the actual mechanism is outside that class. This is beyond the realm of the present article.

Two simple examples illustrate some of the ideas arising earlier. In the first of these, we shall see that, even in the simplest models for a question of type (b), there are many procedures which use the data efficiently, and the choice among them is usually difficult.

Testing between simple hypotheses

We shall continue to use the language of independent tosses of a coin, although you can, of course, imagine H and T to stand for the outcomes of any dichotomous experiment. Suppose our coin is known in advance to be characterized by either the value $p = 1/3$ or by the value $p = 2/3$; no other values are possible. On the basis of three independent flips, we are to guess which value of p actually characterizes the coin. Each of the two possible guesses specifies a single probability mechanism, unlike the example 'The coin is unfair ($p \neq 1/2$)' discussed earlier, where the guess did not attempt to describe the exact value of p. When a 'hypothesis' consists of just one possible mechanism, it is said to be 'simple'. Thus, statisticians refer to our present problem as one of testing between two simple hypotheses. The characterization of procedures that use the data efficiently will turn out to be particularly simple for such problems and much simpler than that for other hypothesis-testing problems.

Here are five possible prescriptions for making a guess from the data; you can easily write down others.

Procedure 1: guess '$p = 2/3$' if there are at least two H's in the three tosses; guess '$p = 1/3$' otherwise.

Procedure 2: guess '$p = 2/3$' if the first H precedes the first T or if no T occurs; guess '$p = 1/3$' otherwise.

Procedure 3: guess '$p = 2/3$' if there is at least one H in the three tosses; guess '$p = 1/3$' otherwise.

Procedure 4: guess '$p = 2/3$' if there is at most one H in the three tosses; guess '$p = 1/3$' otherwise.

Procedure 5: ignore the data and always guess '$p = 2/3$'.

Which of these procedures would you use? Perhaps many people would feel intuitively that there is more tendency to obtain H's when $p = 2/3$ than when $p = 1/3$, so that Procedures 1, 2, and 3 do not seem too unreasonable and Procedures 4 and 5 seem unreasonable (the former because it works in the wrong direction; the latter, because it makes no use of the data). For a precise analysis, we first use a calculation like that of equation 1, near the beginning of this essay, to list the probabilities of the eight possible sequences of three tosses under each of the two possible probability mechanisms (see Table 1).

TABLE 1

Outcome	Probability of outcome when $p = 1/3$	Probability of outcome when $p = 2/3$
HHH	1/27	8/27
HHT	2/27	4/27
HTH	2/27	4/27
THH	2/27	4/27
HTT	4/27	2/27
THT	4/27	2/27
TTH	4/27	2/27
TTT	8/27	1/27

Then, for each procedure, we use Table 1 to compute the probability that the chance outcome will be such as to lead to a correct guess when $p = 1/3$. Next, we perform the same computation with $p = 2/3$. For example, when $p = 1/3$, Procedure 2 makes the correct guess '$p = 1/3$' if there is a T on the first toss, that is, if the outcome is THH or THT or TTH or TTT, with a total probability of 18/27, computed from the middle column. Similarly, summing the probabilities in the last column for the four other possible outcomes which lead to the guess

'$p = 2/3$' yields 18/27 for the probability of a correct guess when $p = 2/3$. A similar calculation for the other procedures gives Table 2.

TABLE 2

Procedure	Probability of correct guess when $p = 1/3$	Probability of correct guess when $p = 2/3$
1	20/27	20/27
2	18/27	18/27
3	8/27	26/27
4	7/27	7/27
5	0	1

What does this table show us about the appeal of various procedures? If a procedure were perfect, it would have probability one (certainty) of making a correct guess, whether the actual p is 1/3 or 2/3. It is not hard to see that no such procedure exists. Procedure 5 is always correct if $p = 2/3$, but never right if $p = 1/3$. The procedure which ignores the data and always guesses '$p = 1/3$' would have the opposite behavior. No other procedure has probability one of making a correct guess, whatever p may be.

Moreover, we see that no procedure maximizes the probability of making a correct guess when $p = 1/3$ and also when $p = 2/3$: Procedure 5, alone among all procedures, maximizes the probability of making a correct guess if $p = 2/3$, but it minimizes this probability when $p = 1/3$. Since a perfect maximizing procedure does not exist, how are we to select the procedure to be used?

To start with, Procedures 2 and 4 can be eliminated from contention since, whatever the actual value of p, each of these has a smaller probability of making a correct guess than does Procedure 1. Thus, while we have not yet decided on which procedure to use, we can tell anyone who proposes to use Procedure 2 that he should not do so, since we can give him another procedure which has a higher probability of yielding a correct guess, whether the actual p is 1/3 or 2/3. A subject of considerable study among theoretical statisticians is the characterization in different problems of those procedures, called 'admissible', which remain after the inferior procedures, such as Procedures 2 and 4, have been eliminated. The choice of the procedure to be used is thereafter restricted to the admissible procedures.

Procedures 1, 3, and 5 can be shown to be among the admissible procedures in our example; you can see for yourself that none of these three eliminates either of the others by having a higher probability in both columns.[2] The choice among

[2] Each admissible procedure is characterized, by the Neyman–Pearson Lemma, to guess '$p = 2/3$' for these outcomes in Table 1 for which the ratio of the last to the middle column is greater than some constant c. You will see that c can be taken to be 1 for Procedure 1, 1/4 for Procedure 3, and 1/10 for Procedure 5. For technical reasons, theoretical statisticians also include procedures which, when the outcome yields a ratio c, perform an auxiliary experiment to decide which guess to make. These need not concern us here.

Procedures 1, 3, 5, and the other admissible procedures must now involve some additional criterion which we have not yet mentioned. There are philosophical differences among statisticians as to the appropriateness of various criteria which have been suggested and used over the years. I shall next mention a few of these criteria.

Basic ways of choosing

The approach of Fisher in hypothesis testing, usually in more complex examples than ours, was to use a procedure that attained a previously specified probability of making a correct guess under one particular probability mechanism. This mechanism was often an older theory whose validity was to be tested by the experiment or 'no difference between new varieties and old' in the experiment we discussed in leading up to question (c). In our present example, let us suppose that $p = 1/3$ is the chosen mechanism. We must then specify the value a procedure is to yield in the middle column of Table 2. The Neyman–Pearson theory added to this formulation an aspect never treated by Fisher: Among all procedures with the specified value in the middle column, one should select that procedure with largest entry in the last column. Thus, among all tests with 8/27 in the second column, of which there are many in addition to Procedure 3 (for example, guess '$p = 1/3$' for outcomes THT and TTH, and '$p = 2/3$' otherwise), the Neyman–Pearson Lemma characterizes Procedure 3 as the one with maximum probability of making a correct guess when $p = 2/3$. If one accepts the figure 8/27, there is no doubt about using Procedure 3. The practical shortcoming is: Why 8/27 rather than some other figure (for example 20/27, which would have dictated using Procedure 1)? The value has often been chosen by tradition alone, in various ways in different fields of applications.

A second criterion is given by the 'minimax' principle. For each procedure, we compute the minimum of the two probabilities listed in Table 2: 20/27 for Procedure 1, 8/27 for Procedure 3, 0 for Procedure 5. This figure gives a measure of the worst possible performance of the procedure. The pessimist, worried about the possibility of encountering a coin whose p yields this worst performance, may want to choose a procedure for which this minimum probability is as large as possible. In the present example, that is Procedure 1. There are also some objections to this approach, but space does not permit us to discuss them here.

A third possibility is to compute for each procedure some weighted average of the two probabilities listed in Table 2 and to choose a procedure that maximizes this average. For example, if we compute $\frac{1}{2}$ the middle column plus $\frac{1}{2}$ the last column for each procedure, we find that this average is largest for Procedure 1. On the other hand, $\frac{1}{4}$ of the middle column plus $\frac{3}{4}$ of the last column is largest for Procedure 3. Thus, different weighted averages favor the choice of different procedures. The practical difficulty here is: How do you choose the weights? In a very few problems, we know so-called *prior* probabilities that the actual physical mechanism will be of one form or another. For example, in our coin problem, we might know that, over a long period, $\frac{3}{4}$ of the coins have come out of the mint with

$p = 2/3$, and $\frac{1}{4}$ with $p = 1/3$. In such a case, it is appropriate to use the prior probabilities $\frac{3}{4}$ and $\frac{1}{4}$ as weights, since the resulting average our procedure will maximize is then the total probability of a correct guess. This yields the choice of Procedure 3 in the example just given. Because the general scheme for computing such procedures uses the simple probabilistic formula known as Bayes' Theorem, the resulting procedure is called a Bayes procedure for the given prior probabilities. But there are few practical examples where such prior probabilities are known. In the absence of such knowledge, it is tempting to use equal prior probabilities to represent one's ignorance (which would lead to the use of Procedure 1 in our example) or to use the recently much-publicized subjectivist approach, which employs, in place of unknown physical prior probabilities, corresponding 'subjective probabilities', which are supposed to reflect a quantitative measure of the customer's feelings about the problem. Both these possibilities have also received considerable criticism, which we cannot discuss here.

A fourth possibility is to take note of the symmetry of the problem, as reflected in the fact that relabeling H as T and vice versa means interchanging the values $p = 1/3$ and $p = 2/3$, since a coin with probability 1/3 of yielding H has probability 2/3 of yielding T (which becomes H under the relabeling). The symmetry of the problem suggests the criterion of using a symmetric procedure, one that guesses '$p = 1/3$' for a given sequence (for example, TTH) if and only if it guesses '$p = 2/3$' for the relabeled sequence (for example, HHT). In our example, only Procedure 1, among all admissible procedures, has this symmetry; in fact, it can be shown to be better than all other symmetric procedures. Thus, under the criterion of symmetry (usually called 'invariance'), we would use Procedure 1. The difficulty with using this approach is that there are more complex problems in which no invariant procedure is admissible, so that the invariance criterion cannot lead to the choice of a satisfactory procedure; and there are other problems lacking symmetry, so that all procedures are equally 'symmetric', and this criterion does not choose among them. An example of the latter is our coin problem with the two possible values $p = 1/3$ and $p = 2/3$ replaced by $p = 1/5$ and $p = 2/3$. Under relabeling, these become $p = 4/5$ and $p = 1/3$, which are not merely the original pair of values in reversed order. There is no symmetry to the problem, so all procedures are equally symmetric.

You may well wonder at this point which procedure to use in our example. I cannot tell you without further knowledge of the background of the problem, the real meaning of the two events we labeled H and T (they might mean 'life' and 'death', etc.), the use you will make of your guess, and a re-examination of the relative losses (which we have tacitly assumed equal) for the two possible types of incorrect guess. Even with this knowledge, I might find it difficult to tell you that a particular procedure is the one and only obvious one to use, although there are many circumstances where I would stop this hedging and use Procedure 1.

I have included such a long description of some of the criteria people use to select a procedure and have indicated that all of them have shortcomings, in order to emphasize the view of many theoretical statisticians—that there is no simple recipe which will tell you how to choose a statistical procedure in all possible settings. This is a frustrating aspect of the subject and perhaps one without parallel in other mathematical areas. It presents a large target for future theoretical work. At the same time it points up the danger of using some simple, superficially appealing recipe to select a statistical procedure, and the need for expertise in that selection.

An even more fundamental danger occurs in the use of an intuitively appealing rule of thumb to construct a procedure, without any consideration of admissibility. The statistical literature is full of suggestions of procedures like Procedure 2 and is almost devoid of mention of Procedure 5 (which should be used if, for example, one knows that the prior probability that $p = 1/3$ is quite small, say $1/10$). In other settings where the mechanism is nothing as simple as coin-flipping, it is even easier to be led astray by intuition. We shall illustrate this now in an estimation problem that will exemplify the Pearson–Fisher controversy mentioned earlier.

An estimation example

A particle is moved along some fixed line by a force field around it. In each of three independent experiments, the particle, at rest, is placed at the same point P, and its position is measured one second later. Taking P as the origin, if the field were unaltered by outside influences and if the measurements were without error, the measured position of the particle after one second would be the same value, say θ millimeters, in each experiment. However, the experiments are not perfect, and we suppose that in each experiment the probability that the recorded measurement will be θ, $\theta-1$, or $\theta+1$ is $1/3$ for each. (This law of errors is being chosen for arithmetic simplicity rather than physical reasonableness, but it will illustrate a phenomenon that could also exist for more reasonable but more complicated models.) Here the force field, and thus θ, is unknown in advance of experiments, the object of which is to 'guess the value of θ'. Let us call the three measurements X_1, X_2, X_3. A statistical procedure is a real-valued function of these three quantities, that is, a rule for computing from them a real number that will be stated as our guess as to the actual value of θ for the particular field at hand.

A common intuitive line of reasoning for selecting a procedure begins by remarking that the value θ is the mean, or first moment, of the mass distribution corresponding to our law of errors, which assigns mass $\frac{1}{3}$ to each of the points $\theta-1$, θ and $\theta+1$. This reasoning then proceeds to note that we can summarize the three measurements in the form of an 'empiric mass distribution' that assigns to each real value the proportion of observations taking on that value. (For example, if $X_1 = 16{\cdot}2$, $X_2 = 17{\cdot}2$, $X_3 = 16{\cdot}2$, this empiric mass distribution would assign mass $\frac{2}{3}$ to the value $16{\cdot}2$ and $\frac{1}{3}$ to the value $17{\cdot}2$.) Since this chance empiric mass distribution has some tendency to resemble the underlying mass distribution corresponding to the probability law, the reasoning concludes that a guess as to the mean θ of the latter should be obtained from the mean of the former, which is easily seen to be the 'sample mean' $(X_1+X_2+X_3)/3$. This is a simple example of Pearson's 'method of moments'.

We can easily improve upon this procedure by altering the guessing rule whenever two of the three X_i take on a value 2 mm away from the third, guessing the midpoint between these two values rather than the sample mean in such cases. The sample mean will be in error by $\frac{1}{3}$ mm for such outcomes, while the new guess will be errorless. For example, if $X_1 = 14$, $X_2 = 16$, $X_3 = 14$, you can see that θ must be 15 because of the particular form of our law of errors, and this is what our new procedure guesses it to be. But the sample mean yields a guess of $\frac{44}{3}$, underestimating θ by $\frac{1}{3}$. Since the two procedures yield the same guess (and, hence, the same error) for all other types of outcomes, the new rule is certainly preferable to the sample mean, and it is not hard to see that the probability is $\frac{6}{27}$ that the outcomes will be of the type where this reduction of error occurs.

You may well think it an obvious improvement on the sample mean to replace it by the improved guess in the foregoing situations where the exact value of θ is obvious from the outcomes. But it is easy to alter the law of errors slightly, at the expense of arithmetic simplicity, to obtain a model where the sample mean can be improved upon without such an obvious motivation. The point is that the intuitively appealing guess, 'sample mean', used blindly in so many experimental settings, may be appropriate for some assumed laws of errors and terribly inefficient for others. Only a precise probabilistic analysis, and not any intuitive rule of thumb, can determine what procedures are reasonable ones for a given statistical model.

That is what the modern mathematical theory of statistical inference is all about.

Bibliography

Herman Chernoff and Lincoln E. Moses, *Elementary Decision Theory*, Wiley, New York, 1959.

J. L. Hodges, Jr., and E. L. Lehmann, *Basic Concepts of Probability and Statistics*, Holden-Day, San Francisco, 1964.

Irrational Rectangles

T. J. FLETCHER
University of Nottingham

What is a convenient shape for a piece of paper? Put this way the question is too vague to permit any mathematical answer, but if certain assumptions are made precise answers are implied. If paper is manufactured in large sheets it is convenient for a manufacturer to be able to cut it down and produce smaller standard sheets without wastage. In particular he might wish to arrange that if a standard sheet was cut in half two smaller standard sheets resulted.

The original question then becomes, what proportions must a rectangle have if the two smaller rectangles resulting when it is cut across the middle are similar to the original?

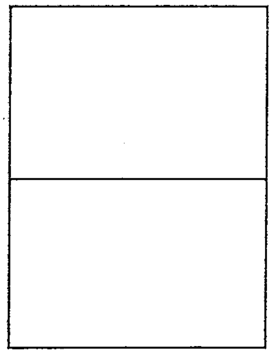


Figure 1

If the sides of the original are of lengths x and 1, then the sides of each half are $\frac{1}{2}x$ and 1. If these sides are in the same ratio as the sides of the original,

$$\tfrac{1}{2}x : 1 = 1 : x.$$

(Be careful to get the ratios expressed the right way round.)

This means that

$$x^2 = 2,$$

and so

$$x = \sqrt{2}.$$

The answer to the question is that the lengths of the sides of the rectangle must be in the ratio $\sqrt{2} : 1$.

What must the lengths of the sides be if the area of a sheet is to be 1 square metre? If the lengths are $s\sqrt{2}$ and s, then

$$s^2\sqrt{2} = 1$$

and so

$$s = 1/\sqrt[4]{2}$$

and the sheet has to be $\sqrt[4]{2}$ metres long and $1/\sqrt[4]{2}$ metres wide.

$\sqrt[4]{2} = 1{\cdot}189$ and so the dimensions required are 1189 mm and 841 mm. The half sheet will then measure 841 mm by 594 mm, and a half of this will once again have sides in the same ratio and will measure 594 mm by 420 mm.

In fact, determining the original size by the criteria stated, we are led to consider a range of paper sizes based on the following sequence:

1189, 841, 594, 420, 297, 210, 148, 105, 74, 52.

(These terms are rounded off in order to avoid fractions of a millimetre.)

This range of sizes is standard in Germany, where it is known as the DIN A series. DIN stands for Deutsche Industrie-Normen, meaning German Industrial Standards. As a move towards metrication these have also been adopted as standard sizes in British Drawing Offices. This range of paper sizes will not, of course, cover all requirements and there are other DIN series for other purposes.

It is well known that $\sqrt{2}$ is an irrational number, that is to say, it is not possible to find two integers a, b, such that $a^2 = 2b^2$. A simple algebraic proof by contradiction may be found in many books, but it is interesting to see a proof suggested by the paper sizes which have just been described. A carpenter uses rectangles of this shape, and he constructs them by first drawing a square $ABPQ$, then drawing an arc, centre A, radius AP to cut AQ produced at D. The rectangle $ABCD$ is then of the shape required.

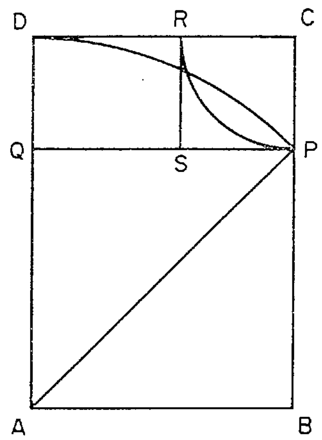


Figure 2

Continue this style of construction, and with centre C and radius CP draw an arc cutting CD at R. Complete the square $CPSR$. You should now be able to prove, as a simple exercise, that $SRDQ$ is a rectangle which is similar to $ABCD$. In slightly different words, removing two squares from our standard paper shape leaves another rectangle of the standard shape—although it is not a standard size.

Now let us assume that the sides of the original rectangle are commensurable, that is to say that they are in the ratio of two positive integers a and b. We will see later that this assumption is untenable, because it leads to a contradiction. The assumption would mean that there was a certain length h (say), such that AD was of length ah and AB was of length bh. What would be the dimensions of the rectangle $SRDQ$?

$$DQ = AD - AQ = AD - AB = (a-b)h,$$

and

$$DR = DC - CR = AB - CP = bh - (a-b)h = (2b-a)h.$$

This means that DR and DQ would be in the numerical ratio $(2b-a):(a-b)$. Now we have seen that this ratio is the same as the original ratio $a:b$, which we are assuming is $\sqrt{2}$. But the numbers $2b-a$ and $a-b$ are smaller than the numbers a and b, (why?), and so the geometrical construction has furnished two smaller numbers in the same ratio as the original.

Exercise

Those who prefer strictly algebraic arguments should prove, by algebra alone, that

$$a^2 = 2b^2 \Rightarrow (2b-a)^2 = 2(a-b)^2.$$

We may now repeat the geometrical argument and starting with $a' = 2b-a$ and $b' = a-b$ obtain two more positive integers, $2b'-a'$ and $a'-b'$, which are smaller still and which are in the same ratio $\sqrt{2}$ once again. The *geometrical* argument may be repeated indefinitely, but the *arithmetical* argument may only be repeated a finite number of times, since the numbers are positive integers to begin with and they get smaller each time.

Hence we have arrived at a contradiction, and so the assumption on which the argument was based, that two whole numbers could be found with the ratio $\sqrt{2}$, is untenable.

The numerical procedure which has just been described may be used in a slightly different way. We saw that, if $a^2/b^2 = 2$, then $a'^2/b'^2 = 2$ also; but we proved that such equalities cannot hold if a and b are integers. However, a^2/b^2 can be an approximation to 2, and it would be plausible to expect that if a^2/b^2 is approximately 2 then a'^2/b'^2 is approximately 2 as well. But which would one expect to be the better approximation? The numbers a', b' are smaller than a and b, and it would be surprising if the smaller numbers gave a better approximation, because if they did one could expect to repeat the process and derive two still smaller numbers which gave a still better approximation, and this is unlikely.

Now a' and b' are derived from a and b by the transformation

$$(a, b) \to (2b-a, a-b) = (a', b');$$

in other words,

$$a' = -a+2b,$$
$$b' = a-b.$$

What is the inverse transformation? Solving for a and b we have

$$a = a'+2b',$$
$$b = a'+b'.$$

Looked at this way it appears plausible that, if a'/b' is approximately $\sqrt{2}$, then $(a'+2b')/(a'+b')$ might be a better approximation. This is true, but not entirely obvious, and the proof is set as one of the problems at the end of the magazine. The following related exercise is rather easier. (It is convenient now to drop the primes as they have served their purpose.)

Exercise

If a and b are positive integers, prove that the difference between $(a+2b)^2/(a+b)^2$ and 2 is less than the difference between a^2/b^2 and 2.

The next rectangle to be considered has a shape which is perhaps even more famous. The Greeks asked what is the most beautiful shape for a rectangle? For no very clear reason they decided that it was the shape for which the rectangle had the following property:

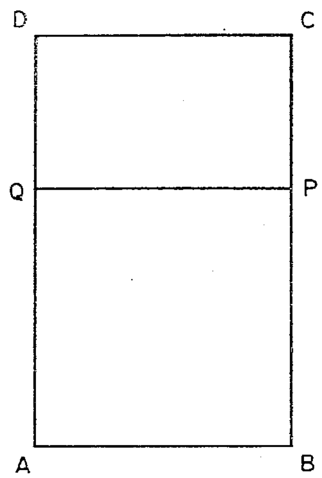


Figure 3

If a square $ABPQ$ is removed from the rectangle $ABCD$ then the remaining rectangle $PCDQ$ is similar in shape to the original. If AB is taken as a unit length and AD is equal to x, then

$$(x-1) : 1 = 1 : x.$$

This means that x satisfies the equation

$$x^2-x-1 = 0.$$

The roots of this equation are $\frac{1}{2}(1 \pm \sqrt{5})$, which are approximately 1·618 and $-0{\cdot}618$. It is traditional to call the positive root of the equation the *golden ratio*, and to denote it by the Greek letter gamma γ. It is easy to prove algebraically that $\sqrt{5}$ is irrational and so to conclude that γ is irrational; but algebraic notation had not been invented in Greek times, and it is in any case interesting to seek for a more geometrical line of proof. We will consider a method rather like the one employed for the previous rectangle.

Suppose it possible that the sides of the golden rectangle are in the numerical ratio a/b, where a and b are integers. As before, the lengths of the sides will then be ah and bh, where h is some unit of length.

In the figure $PCDQ$ is also a golden rectangle (this was part of the definition). But the sides of the rectangle will be bh and $(a-b)h$, with the positive integers $(b, a-b)$ in the same ratio as (a, b) and smaller (respectively).

Once again, we may repeat the geometrical construction indefinitely, but the corresponding arithmetical operation can only be done a finite number of times, and so there is a contradiction. Hence it is not possible for the sides of the rectangle to be in the ratio of two integers.

The transformation this time is $(a, b) \to (b, a-b) = (a', b')$; and successive applications of the inverse transformation yield, as may be expected, better and better rational approximations to γ. (See again the problem section.)

The golden ratio is to be found in another famous geometrical figure—the regular pentagon with its diagonals.

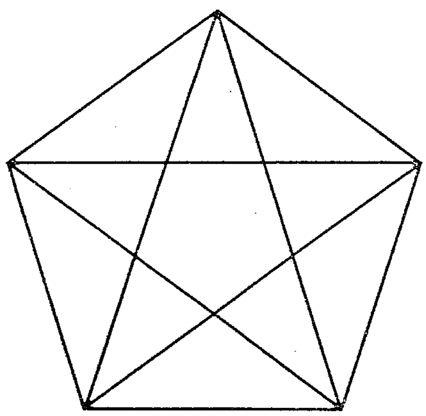


Figure 4

This figure is full of golden ratios. See how many you can find. The diagonal and the side of the pentagon are in golden ratio; this will be used shortly, and there are other occurrences of the ratio as well.

From this figure a further figure may be derived, consisting of an infinite sequence of interlaced pentagons.

How many golden ratios can you see now?

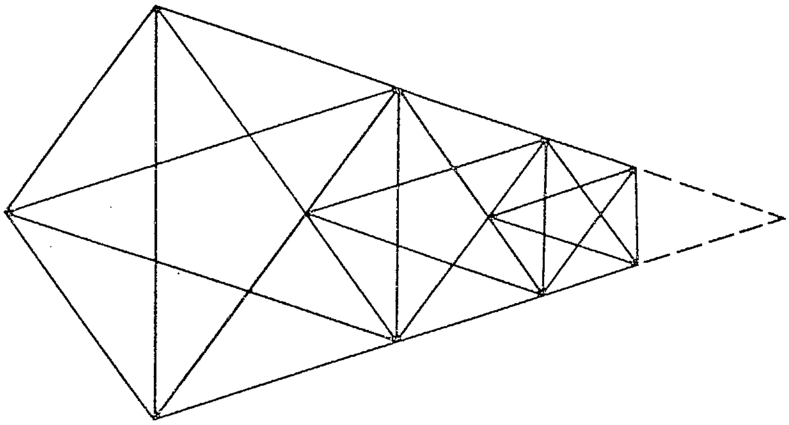


Figure 5

Exercise

As a final investigation, starting with the knowledge that the diagonal and the side of each pentagon are in golden ratio, find out how the previous argument may be applied to this figure, and the contradiction established once more.

Further information about the golden ratio may be found in Gardner (1) and in Northrop (2).

The discovery that there are geometrical ratios which cannot be measured by rational numbers was a major turning point in Greek mathematics. An interesting account of this may be read in Russell's *History of Western Philosophy* (3).

References

1. M. Gardner, *More Mathematical Puzzles and Diversions*, Pelican, Harmondsworth, Middlesex, 1966.
2. E. P. Northrop, *Mathematics*, Pelican, Harmondsworth, Middlesex, 1961.
3. B. Russell, *History of Western Philosophy*, Allen and Unwin, London, 1946.

An Original Solution of a Problem in Calculus

NICHOLAS YOUD

St. Paul's School, London

together with editorial comments

The following question has appeared in a Cambridge scholarship examination:

Show that, for positive integers m, n,

$$\int_0^{2\pi} \sin mx \sin nx \, dx = \begin{cases} 0 & (m \neq n), \\ \pi & (m = n). \end{cases}$$

Let f be an integrable function defined for $0 \leqslant x \leqslant 2\pi$, N a positive integer and let numbers α_n ($n = 1, 2, \ldots, N$) be defined by

$$\alpha_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \sin nx \, dx.$$

Show that, among all possible choices of numbers β_n ($n = 1, 2, \ldots, N$), the choice

$$\beta_n = \alpha_n \quad (n = 1, 2, \ldots, N)$$

makes the expression

$$\int_0^{2\pi} \{f(x) - \sum_{n=1}^{N} \beta_n \sin nx\}^2 \, dx$$

a minimum.

The first part is a simple exercise in integration. The proof of the second part can be expressed in the language of vector spaces.

We consider the vector space V consisting of all (real) functions which are (Riemann) integrable over the interval $(0, 2\pi)$. For any members f and g of V we put

$$f.g = \frac{1}{\pi}\int_0^{2\pi} f(x) g(x) \, dx. \tag{1}$$

It is easily checked that, whenever f, g, h belong to V and α, β are real numbers,

(i) $f.g = g.f$;

(ii) $(\alpha f + \beta g).h = \alpha(f.h) + \beta(g.h)$;

(iii) $f.f \geqslant 0$.

Thus $f.g$ is an *inner product* for V; and all members f of V have associated with them the *norm* $\|f\|$ given by

$$\|f\| = \sqrt{(f.f)},$$

i.e.,

$$\|f\|^2 = \frac{1}{\pi}\int_0^{2\pi} f^2(x) \, dx.$$

If $u_1, u_2, u_3, \ldots$ are the elements of V defined by

$$u_n(x) = \sin nx \quad (n = 1, 2, \ldots),$$

then, by the first part of the question,

(iv) $u_m . u_n = 0$ when $m \neq n$; and

(v) $\|u_n\|^2 = u_n . u_n = 1$ for $n = 1, 2, \ldots$.

Also, the definition of the real numbers $\alpha_1, \alpha_2, \ldots$ in the second part is simply

$$\alpha_n = f.u_n \quad (n = 1, 2, \ldots),$$

and we wish to show that, for all possible choices of the real numbers $\beta_1, \ldots, \beta_N$,

$$\left\| f - \sum_{n=1}^{N} \beta_n u_n \right\|$$

is a minimum when $\beta_n = \alpha_n$ $(n = 1, 2, \ldots, N)$.

At this stage Nicholas Youd argued as follows. Denote by V_N the subspace of V consisting of functions of the form

$$\sum_{n=1}^{N} \lambda_n u_n.$$

If, now,

$$f_N = \sum_{n=1}^{N} \alpha_n u_n,$$

then (i)–(v) show that, for $k = 1, \ldots, N$,

$$(f - f_N) . u_k = \left(f - \sum_{n=1}^{N} \alpha_n u_n \right) . u_k = f . u_k - \alpha_k = 0.$$

Thus, if $f_N^{\perp} = f - f_N$,

$$f = f_N + f_N^{\perp},$$

where f_N is orthogonal to V_N (i.e., $f_N^{\perp} . v = 0$ for all v in V_N). It follows that

$$\left\| f - \sum_{n=1}^{N} \beta_n u_n \right\|^2 = \left(f_N + f_N^{\perp} - \sum_{n=1}^{N} \beta_n u_n \right) . \left(f_N + f_N^{\perp} - \sum_{n=1}^{N} \beta_n u_n \right)$$

$$= \left(f_N - \sum_{n=1}^{N} \beta_n u_n \right) . \left(f_N - \sum_{n=1}^{N} \beta_n u_n \right) + 2 \left(f_N - \sum_{n=1}^{N} \beta_n u_n \right) . f_N^{\perp} + f_N^{\perp} . f_N^{\perp}$$

$$= \left\| f_N - \sum_{n=1}^{N} \beta_n u_n \right\|^2 + \| f_N^{\perp} \|^2$$

$$= \left\| \sum_{n=1}^{N} (\alpha_n - \beta_n) u_n \right\|^2 + \| f_N^{\perp} \|^2; \qquad (2)$$

and the choice $\beta_n = \alpha_n$ $(n = 1, 2, \ldots)$ clearly minimizes (2).

We note that the identity (2) is of the form

$$\| g \|^2 = \| g_N \|^2 + \| g_N^{\perp} \|^2,$$

where $g = g_N + g_N^{\perp}$ and g_N belongs to V_N, while $g_N^{\perp}$ is orthogonal to V_N; it is, therefore, the analogue of Pythagoras' theorem. The standard abstract proof of the approximation theorem in question lacks this geometrical background, but is shorter. It depends on the identity

$$\left\|f-\sum_{n=1}^{N}\beta_n u_n\right\|^2=\left(f-\sum_{n=1}^{N}\beta_n u_n\right).\left(f-\sum_{n=1}^{N}\beta_n u_n\right)$$

$$=f.f-2f.\sum_{n=1}^{N}\beta_n u_n+\left(\sum_{n=1}^{N}\beta_n u_n\right).\left(\sum_{n=1}^{N}\beta_n u_n\right)$$

$$=f.f-2\sum_{n=1}^{N}\beta_n(f.u_n)+\sum_{n=1}^{\infty}\beta_n^2(u_n.u_n)$$

$$=\|f\|^2-2\sum_{n=1}^{N}\beta_n\alpha_n+\sum_{n=1}^{N}\beta_n^2$$

$$=\|f\|^2-\sum_{n=1}^{N}\alpha_n^2+\sum_{n=1}^{N}(\alpha_n-\beta_n)^2. \tag{3}$$

Since $(\alpha_n-\beta_n)^2 \geqslant 0$ for all β_n, the right-hand side is least when $\beta_1=\alpha_1,\ldots,\beta_N=\alpha_N$.

Nicholas Youd has (re)discovered a portion of the modern theory of Fourier series.

Given an arbitrary vector space V in which an inner product is defined, a sequence $(u_1, u_2, \ldots)$ of elements of V for which (iv) and (v) hold is called *orthonormal*; and if f is any member of V, the real numbers $\alpha_1, \alpha_2, \ldots$ are called the *Fourier coefficients* of f (with respect to the system $(\alpha_1, \alpha_2, \ldots)$). The identities (3) hold in this general situation and show, as before, that

$$\left\|f-\sum_{n=1}^{N}\beta_n u_n\right\|$$

is least when $\beta_n=\alpha_n$ $(n=1,2,\ldots,N)$. The advantage of framing the theory in such general terms is that it applies to a large number of particular cases which would otherwise have to be treated separately. Two further examples will suffice.

First we again take V to be the space of functions integrable over $(0, 2\pi)$ and with inner product specified by (1). But we now let our orthonormal sequence $(u_1, u_2, \ldots)$ be given by

$$u_1(x)=1/\sqrt{2},\quad u_2(x)=\cos x,\quad u_3(x)=\sin x,\quad u_4(x)=\cos 2x,\quad u_5(x)=\sin 2x, \ldots.$$

(It is easy to prove that this sequence *is* orthonormal.)

Secondly we might take V to be the space of functions integrable over $(-1, 1)$ and with inner product

$$f.g=\int_{-1}^{1}f(x)\,g(x)\,dx.$$

The sequence of (normalized) *Legendre polynomials* p_n defined by

$$p_n(x)=\sqrt{\left(\frac{2}{2n+1}\right)}\,\frac{1}{2^n n!}\,\frac{d^n}{dx^n}(x^2-1)^n$$

may then be shown to be orthonormal (but this is not easy).

The reader should have no difficulty interpreting the general result proved above for these two cases.

The abstract approach to mathematics, and its justification, are well illustrated by Nicholas Youd's solution of his scholarship problem.

Friends and Acquaintances

HAZEL PERFECT

University of Sheffield

1. Shaking hands at a party

At a certain party, some (but not necessarily all) of the guests shake hands with each other. We shall prove that the number of guests who shake hands an odd number of times must itself be even. (The proof is very simple, and you may like to try to supply it yourself before reading further.)

Whenever A shakes hands with B, then B also shakes hands with A. Therefore, if we count the number of handshakes made by each guest and add all these together, the total will be an even number, $2n$ say. Now suppose m people shake hands an even number of times and m' people shake hands an odd number of times. The number $2n$ is made up of a sum of m even numbers and a sum of m' odd numbers. The first sum is always even, regardless of whether m is itself even or odd. Therefore the second sum must also be even, and this can only happen if m' is even. Thus our proposition is proved.

2. Three acquaintances or three strangers

The problem just described belongs to 'Combinatorics'. This is a part of mathematics with many different aspects, which we might briefly describe as the study of arrangements of objects according to given rules, and the enumeration and classification of such arrangements. All the problems discussed in this article are of combinatorial type. The theory of permutations and combinations belongs to Combinatorics. Another important branch of the subject which has been studied intensively in recent years is Graph Theory.[1] Graphs are just sets of points, certain pairs of which are joined by lines which we call 'edges'. Let us represent the guests $A, B, C, \ldots$ at the party mentioned in Section 1 by points (with the same labels $A, B, C, \ldots$). We shall join A and B by an edge whenever A and B shake hands. The proposition that we proved above asserts, in graph-theoretic terms, that in a graph

[1] If you read on, you will see that the 'graphs' referred to here are quite different from those drawn on squared paper to represent numerical data, or the sketch-graphs which illustrate the behaviour of elementary functions.

the number of points through which there pass an odd number of edges is even. As we saw, this can be proved by a very simple counting argument.

Let us consider another party now, at which six people are present. Can you devise an argument to show that, before the party, either three (or more) of the guests were mutual acquaintances or three (or more) of them were mutual strangers?

A graph provides an appropriate visual aid to help us to solve this problem. We shall represent the six people by six points $A_1, A_2, \ldots, A_6$ in the plane, no three of them being collinear, and we shall join $A_i A_j$ by a red line or a blue line according as A_i and A_j were acquaintances or strangers before the party. Then each pair of points is joined by either a red line or a blue line, and we wish to show that the figure contains at least one 'monochromatic' triangle, i.e., at least one triangle in which all three sides have the same colour. Consider the five edges through A_1. At least three of them must be the same colour. For definiteness, suppose $A_1 A_2$, $A_1 A_3$, $A_1 A_4$ are red, and look at the triangle $A_2 A_3 A_4$. If it is monochromatic, we have finished. If not, then it has both blue and red edges; in particular, it has at least one red edge, say $A_2 A_3$, and $A_1 A_2 A_3$ is red. The proof is therefore complete. A little experimenting will convince you that, if there are only five points instead of six, then there may be no monochromatic triangle in the figure.[2]

A slightly stronger form of the result is that the figure of six points always contains at least *two* monochromatic triangles. You will find it easy to show that the number 2 cannot be improved upon and, further, that there exists a figure which contains exactly two red triangles and no blue triangle and another figure which contains exactly one triangle of each colour.

Much more general results than these are to be found in the mathematical literature. For instance, the result that a graph with enough points (in fact, with at least six points), and in which every two points are joined by a red or a blue edge, contains a monochromatic triangle is a very special case of an important theorem of the logician and mathematician F. P. Ramsey.

3. Dancing partners

Yet another party, at which both boys and girls are present, serves to introduce our next problem. Suppose there are n boys at the party and that, *for every positive integer $k \leqslant n$, every k boys between them are acquainted with at least k girls at the party*. A celebrated result of Philip Hall asserts that, in these circumstances, each boy is able to choose a dancing partner from among the girls he knows so that no two boys choose the same partner.

This is a more difficult problem than the two earlier ones, and we shall give an argument only in bare outline. Let the n boys be labelled $1, 2, \ldots, n$ and, for $1 \leqslant i \leqslant n$, denote by A_i the set of girls with whom the ith boy is acquainted. The italicized conditions above are evidently conditions on the sets $A_1, \ldots, A_n$ and we shall refer to them as 'Hall's conditions'. Our objective is to show that, when Hall's conditions

[2] For the vertices of the triangles you must, of course, only use the five given points, and you must disregard any other points where the edges happen to intersect.

are assumed, we can find n distinct elements, one from each of the sets $A_1, \dots, A_n$. Now if each A_i consists of a single element (i.e., if each boy is acquainted with just one of the girls) then Hall's conditions certainly ensure this. The idea of the proof is to reduce the general situation to this simple special case. Suppose, then, that one of the sets A_i contains at least two distinct elements x, y. It can be shown (and this is the hard core of the proof, the details of which we omit) that one of x, y can be removed from A_i without destroying Hall's conditions. Thus, step by step, we may proceed until each A_i has been reduced to a single element, in such a way that Hall's conditions are preserved at each stage in the reduction. In this way the theorem is proved, and the elements of the resulting sets constitute a possible choice of dancing partners for the n boys.[3]

Hall's theorem lies at the centre of an important branch of Combinatorics which deals with 'representatives' of families of sets. A more practical illustration of Hall's theorem than the one we have just given is to the following 'assignment problem'. A certain firm advertises some vacant jobs of various kinds. In general each applicant is qualified for certain of the jobs and not qualified for others, and the question arises—under what conditions is it possible to assign each man to a job for which he is qualified? Hall's theorem provides the answer.

4. The 'Friendship Theorem'

Our final problem concerns a peculiar community of people. In this community, every two members of the population have just one common acquaintance.[4] The Friendship Theorem affirms that there is then some member of the community who knows everyone else. The only relatively simple proof of this theorem known to me depends on ideas from linear algebra (specifically, properties of matrices), rather than on purely combinatorial reasoning. An entirely combinatorial proof has, in fact, been devised, but it is far from simple and has not been published. So here is a problem awaiting an elementary solution. You are warned, however, that perhaps there is none. You may find it more rewarding to look instead for different interpretations of the Friendship Theorem (in terms of graphs for instance) and from these perhaps to discover some interesting applications.

[3] On another occasion I have described a different proof of Hall's theorem (*Mathematical Gazette*, **53** (1969), pp. 13–19).

[4] We have to make a convention here, namely that no one is acquainted with himself.

Careers for Women Graduates in Mathematics

R. S. H. G. THOMPSON

Imperial College, London

The telephone rang. It was the wife of a professor of medicine, and she was in great distress: her daughter had just been awarded a scholarship at Girton and the girl actually intended to read Mathematics. 'How can she possibly earn a living from mathematics?' the mother asked, 'What sort of job will she be able to get if she does mathematics at the university?' It was very easy to comfort her. But her enquiry made me wonder how one would answer similar questions from a girl who has decided to go to a University (Technical College, or similar institution), who has the ability to follow a mathematics course, with at least moderate success, and yet does not know what opportunities would be open to her after graduation. Of course one would hope such a girl had enough scientific background to combine ancillary science with her degree course, but even if she were the most 'Arts' of arts-women much of what follows would apply to her directly or with only little modification. (I know of one executive in a large modern industrial complex who began his career as a teacher of art.)

My remarks will apply most directly to those who enter on a course in a British University leading to a Special Honours Degree in mathematics. What applies to those who enter on a course leading to a General Degree, or a mathematics course in universities outside Britain, will depend on the level of the mathematics covered.

Part of my responsibility in the mathematics department in which I work is to advise young mathematicians on their immediate future after graduating. Naturally some of them continue, either in their original college or in a different university, to work for a higher degree. In the past, for those who hoped to make a career in the academic world, this has been the Ph.D. Because of government-imposed financial restrictions, academic opportunities have become strictly limited; hence many students now expect their doctorate to lead them to a career in Industry. In contrast to those studying for a Ph.D., there are now many who follow a first degree in mathematics with a year on an M.Sc. course in such subjects as statistics for economics, operational research, systems analysis or even in a school of Business Studies. Many of these courses are directed to the immediate needs of Industry or Commerce; for such people the employment opportunities are many and diverse.

To give an indication of the variety of mathematical jobs available, one need only scan a few issues of the *New Scientist* and the *Times Educational Supplement*. In these, there will be advertised lectureships in British, Australian and other Commonwealth Universities (as well as Technical Colleges or similar institutions). Their specialities will range from Pure Mathematics, through Applied Mathematics and Mathematical Statistics, to Numerical Analysis and Computing Science. There are always a large number of posts available for mathematics teachers in schools. Sometimes there are a few editorial positions advertised with publishers specialising in mathematics or science texts, and many posts in computer programming, market and operational research with industrial and commercial firms, or insurance companies. There are also from time to time positions advertised in the Civil Service either for research workers on special economic projects, or for statisticians in the Central Statistical Office. The experience of one mathematical department in Britain is that, roughly speaking, one third of mathematics graduates go into teaching, either at schools, or after further training and research, in universities, one third go into commerce and industry, and one third into the various branches of the Civil Service.

At this point I am conscious that a girl might be asking 'But what is there special here for girls?' 'What is the difference in opportunities for girls?' The answer is that, broadly, there is little difference between the job prospects for women and men. And here Industry is in no different position from Business or Commerce, although the occasional confusing difficulty may arise. One Research and Development man from a large, multifarious, national industry was asked by a girl at the beginning of his day's interviews with undergraduates if his organisation was prejudiced against the employment of women graduates. He stressed that they were not, and illustrated this by instances of women who had reached high managerial positions in the organisation. He was himself horrified to discover that the specific field of interest of this girl was the sole area restricted to 'males only'. He told me, later, how extremely embarrassed he had been.

There are, naturally, on some of the production sites of heavy industry, places where one would not expect a woman to be employed. In this connection it is amusing (!) to note that a large photographic company, part of whose processes need to be carried out in dark-rooms, does not employ women on such processes. There is also an international computer firm which employs many women graduates, but not on its sales side, because (so they say) the customers have more confidence in a man! This can be contrasted with a manager in the oil industry who offered a position to a Scottish girl student before her degree result was known. She was awarded a 'First' and her professor tried hard to persuade her to stay in the university for research. But the oil firm asked her to come and talk things over with them; which they did so successfully that she is now employed by them. One of their directors explained to me that they had a particular job for her to do, a job that men had tried and in which they had failed to get the desired results, but one which they expected a woman would be specially able to cope with. It is, incidentally, indicative of the far-sighted attitude of industry that she was not to be put on that work until she had been with the firm for two years.

It is sometimes thought that firms are loath to employ women graduates because of the likelihood of their getting married and leaving after what is initially a training period. In fact, many firms have said that their loss-rate from first employment is worse with men than with women.

Earlier in this article we considered the question, 'What for women specifically?' Now we come to the question 'What for mathematicians specifically?' The answer depends somewhat on the excellence (or class) of degree obtained, but in many ways surprisingly little. For those who hope to become actuaries or to enter research laboratories where specialised technical work is carried out, a highly developed mathematical experience is necessary. But those of our students who leave us with weak (third-class) Honours Degrees usually have little difficulty in finding rewarding and interesting employment. Perhaps this can be summarised in the dictum of an aircraft corporation interviewer. He had outlined the skills, attitudes and aptitudes he was looking for, and then added finally, 'What does all that add up to? A mathematician!'

The computer and programming side of industry is an obvious goal for mathematicians. Here the girl who wishes to have some continuation of her employment after she is married has an advantage. There is, for instance, one American research department which farms out considerable programming work to married women in this country.

Sometimes students nearing the end of their time in a mathematics department realise that the bread-and-butter work they will soon be doing may not involve much of the technical expertise in mathematics they have acquired. So they ask, 'Have we wasted our time learning mathematics?' The answer is emphatically, 'No' When commerce, insurance companies, firms of accountants, industry, or government departments seek mathematics graduates, they are for the most part looking for men and women who will, after a few years' work, rise to managerial and executive positions. And for such positions they will have been particularly well fitted by submission to the discipline of mathematics.

It is also important to realise that for those who begin an undergraduate course in an honours school of mathematics, there is almost always the possibility of changing over to an undergraduate course in (for instance) one of the engineering or science disciplines. These, nowadays, carry a very strong mathematical content. The need to insert 'almost always' in the last sentence reminds me that the girl contemplating going to a university should take some care in deciding which one to make her first choice. This is so, not only because of the wide diversity of emphases and methods of presentation in the different schools of mathematics, but also because of the ancillary possibilities or compulsory topics, and the general conditions of work.

A final word beyond the brief of this article: to understand the complexities of the modern, computerised world, and to be fully aware of the universe around us, one needs to be well trained in the wisdom as well as the learning of mathematics. Our present picture of the world and its setting in the universe depends heavily on mathematical understanding. Any intelligent person can be shown what mathematics is; only the wise one can appreciate what it is not.

Mathematics of Today—A Sixth Form Conference

F. BOSTOCK

University of Southampton

A three day conference designed for sixth form pupils, planned jointly by the University of Southampton and *Mathematical Spectrum* was held on 6, 7 and 8 April 1970 at the University. School staff were also invited. It was found necessary to restrict the number of participants, and a total of 214 pupils and 12 teachers attended the conference from 79 schools as far apart as South Devon, Kent and Manchester. There were 125 men and 101 women, and of these a large proportion were in residence at one of the University Halls.

The conference programme was designed to give an appreciation of some fundamental aspects and uses of mathematics. There were two common lectures each day which were expected to be of interest to all, and two specialized lectures from each of five branches of mathematics: applied mathematics, computation, operational research, pure mathematics and statistics. Most of the lectures were complete in themselves so that participants were not restricted to any one of the five branches. The lectures were given by 26 members of staff of the University of Southampton and Professor J. Gani, of the University of Sheffield. Careers in mathematics was the subject of an evening discussion, and on another evening staff and students of the University provided music and refreshments at a dance.

It is evident from several comments and letters that the conference was well received. A similar conference is now planned for Easter 1971; further information is available from the Sixth Form Conference Secretary, Department of Mathematics, The University of Southampton, Highfield, Southampton SO9 5NH.

Problems and Solutions

Readers who have not yet reached the age of 20 on 1 October 1970 are invited to submit solutions to some or all of the problems below: the most attractive solutions will be published in subsequent issues. When writing to the Editorial Office, please state your full name and the postal address of your school, college or university.

Problems

3.1. (i) Prove that, if a, b are positive integers, the difference between $(a+2b)/(a+b)$ and $\sqrt{2}$ is less than the difference between a/b and $\sqrt{2}$, so that

$$\tfrac{1}{1},\ \tfrac{3}{2},\ \tfrac{7}{5},\ \tfrac{17}{12},\ \tfrac{41}{29},\ \ldots$$

is a sequence of better and better rational approximations to $\sqrt{2}$. (This method was described by Theon of Smyrna in the second century A.D.)

(ii) Obtain the first seven terms of the sequence, beginning with 1/1, in which successive terms are generated by the transformation $a/b \to (a+b)/a$. Show that this sequence provides steadily improving rational approximations to $\frac{1}{2}(1+\sqrt{5})$.

3.2. Let n be a positive integer. Show that among any $n+1$ different integers chosen from $1, 2, 3, \ldots, 2n-1, 2n$ there are always two such that one divides the other.

3.3. Prove that $\log x$ cannot be expressed in the form $f(x)/g(x)$, where $f(x)$, $g(x)$ are polynomials in x.

Solutions to Problems in Volume 2, Number 2

16. Let m and n be natural numbers such that $m \leqslant n$. (a) Show that n can be written as a sum of m natural numbers in $\binom{n-1}{m-1}$ different ways. (b) Determine the number of solutions of the equation

$$x_1+x_2+\ldots+x_m = n$$

when $x_1, x_2, \ldots, x_m$ are restricted to non-negative integers.

Solution. (a) Put $n = 1+1+\ldots+1$. The number of ways of expressing n as a sum of natural numbers is the number of ways of choosing $m-1$ '+' signs from $n-1$, i.e., it is $\binom{n-1}{m-1}$.

(b) The equation

$$x_1+x_2+\ldots+x_m = n$$

is equivalent to

$$(x_1+1)+(x_2+1)+\ldots+(x_m+1) = m+n,$$

so the number of non-negative integral solutions of the former is the number of positive integral solutions of

$$y_1+y_2+\ldots+y_m = m+n.$$

By (a), this is $\binom{m+n-1}{m-1}$.

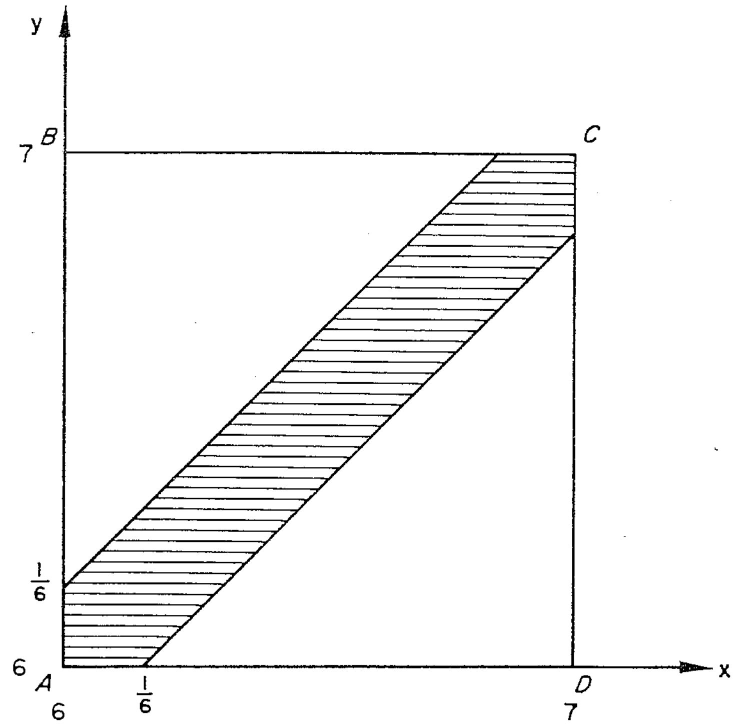
Also solved by Y. Y. Ma (Ipswich School), D. R. Simmons (Enfield Grammar School).

17. Peter and Paul agree to meet at their favourite restaurant. Both choose their arrival times at random between 6 p.m. and 7 p.m., and both wait 10 minutes after arriving. What is the probability that they will meet?

Solution by D. R. Simmons (Enfield Grammar School)

Let x, y be the respective arrival times of Peter and Paul. Then $6 \leqslant x, y \leqslant 7$. For Peter and Paul to meet, we must have $y \leqslant x+\frac{1}{6}$, $x \leqslant y+\frac{1}{6}$. The pairs of values of (x, y) for which Peter and Paul will meet are given by the points in the shaded region of the diagram. Hence the probability that they will meet is

$$\frac{\text{Area of shaded region}}{\text{Area of } ABCD} = \frac{11}{36}.$$



Also solved by A. P. Beverley (King George V School, Southport), Y. Y. Ma (Ipswich School), A. Peil (King George V School, Southport), M. J. Mellish (The King's School, Chester), D. Whitgift (Dulwich College), M. R. Pooley (Gresham's School, Holt), S. R. Blake (Rugby School).

18. Show that, if m is a positive rational number, then $m+(1/m)$ is an integer only if $m = 1$.

Solution by Y. Y. Ma (Ipswich School)

Let $m = p/q$, where p and q are positive integers having no common prime factors. Then

$$m+\frac{1}{m} = \frac{p^2+q^2}{pq}$$

so that, if this is an integer, then p and q both divide p^2+q^2. But then p divides q^2 and q divides p^2. Since p, q have no common prime factors, this means that $p = q = 1$ and $m = 1$.

Also solved by A. Jones and S. A. E. Briggs (The Grammar School, Ebbw Vale), Diane J. Gleek (Queen Mary School, Lytham), M. J. Mellish (The King's School, Chester), D. R. Simmons and David A. McEwan (Enfield Grammar School), M. R. Pooley (Gresham's School, Holt), S. R. Blake (Rugby School).

19. Let m, n be integers such that $m \geqslant 2$, $n \geqslant 3$. Show that there exist positive integers a, b such that

$$m^n = a^2-b^2.$$

Give an example to show that the result is false when $n = 2$.

Solution by J. D. Hoddy (Hulme Grammar School for Boys, Oldham)

Let r be a positive integer. Then

$$(r+1)^2-r^2 = 2r+1 \quad \text{and} \quad (r+2)^2-r^2 = 4(r+1).$$

Thus every odd integer greater than 1 and every multiple of 4 greater than 4 can be expressed as the difference of two squares of positive integers. If m is odd, then m^n is odd (and $m^n > 1$), so that m^n can be expressed in the required form. If m is even, say $m = 2p$, then $m^n = 2^n p^n$ and, since $n \geqslant 3$, m^n is a multiple of 4 and greater than 4. Hence, when m is even, then m^n can be expressed in the required form. The result is seen to be false when $n = 2$ if we take $m = 2$; in fact, $m = 2$ is the only case when the result is false.

Also solved by Y. Y. Ma (Ipswich School), M. J. Mellish (The King's School, Chester), D. R. Simmons and David A. McEwan (Enfield Grammar School), M. R. Pooley (Gresham's School, Holt), S. R. Blake (Rugby School).

A Survey Review of Books on Vectors and Applications used in Sixth Forms in Schools

HUGH NEILL

St Paul's School, London S.W.13

The books which I have considered in this context are the following:

Advanced Level Vectors. By A. P. ARMIT. Heinemann Educational Books Ltd, London, 1968. Pp. 152. 18*s*.

An Introduction to Vectors. By A. E. COULSON. Longmans, London, 1967. Pp. 170. 15*s*.

Vectors and their Applications to Geometry and Mechanics. By A. J. FRANCIS. Bell, London, 1968. Pp. 140. 12*s*. 6*d*.

Vectors. By E. H. LEATON. George Allen and Unwin Ltd, London, 1968. Pp. 176. 21*s*.

Elementary Vector Algebra. By A. M. MACBEATH. Oxford University Press, 1966. Pp. 136. 14*s*.

Vector Algebra. By E. M. PATTERSON. Oliver and Boyd, London, 1968. Pp. 144. 17*s*. 6*d*.

Theory and Problems of Vector Analysis. By M. R. SPIEGEL. Schaum Outline Series, McGraw-Hill, New York, 1959. Pp. 225. 34*s*.

Introduction to Elementary Vector Analysis. By J. C. TALLACK. Cambridge University Press, 1966. Pp. 139. 17*s*. 6*d*.

S.M.P. Advanced Mathematics. Books 1, 2, 3, 4. Cambridge University Press, 1966, 1968, 1968, 1969. Pp. 1382. Total price 95*s*.

I apologize to any author whose book I may inadvertently have omitted from this list. There are number of books which give a full discussion of vectors as a preliminary to a course on linear algebra, but I have deliberately not considered them because their style and approach is bound to be different from these books which largely concentrate on using vectors as a tool in geometry or mechanics.

1. General remarks

I have fairly strong views about vectors and it is only fair that I should come into the open and reveal them. I think a book on vectors should do several things, but in particular these two:

(i) It should lay a reasonable foundation for a more advanced study of linear algebra. At this level most work is done in a maximum of three dimensions where it is not unreasonable to assume properties of bases and the unique representation of any vector in terms of the basis, but the future specialist should surely have some idea of how these ideas are generalised to higher dimensions as well as the problems involved in this generalisation. If the problem is tackled, well and good. If it is not, then I would like to see the nature of the assumptions involved explicitly stated rather than glossed over.

(ii) It should encourage (implicitly or explicitly) the use of vectors without co-ordinate systems before using the representations

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \text{ or } \quad a = a_1 i + a_2 j + a_3 k.$$

Solutions achieved without the co-ordinate systems are more elegant, more satisfying and demand greater undersatnding of the ideas involved than do the same ideas when they rely on a particular co-ordinate system which is later imposed. For the same reasons I prefer to see, in this type of book on vectors, $\boldsymbol{a}.\boldsymbol{b}$ defined as $ab \cos \theta$ and used in this form before the derivation of $a_1 b_1 + a_2 b_2 + a_3 b_3$. I know that pupils find it easier at the start to work with co-ordinate systems, and less able pupils may have to work almost entirely using them, but the specialist should certainly be encouraged to manage without them if he can.

For ease of reference I have analysed the contents of the books in the following table. All the books mentioned discuss scalar product, so in that part of the table I have indicated whether the definition $\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$ or $\boldsymbol{a}.\boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$ has been used. In the other parts of the table a cross indicates that the book deals with that topic.

	Discussion of basis	Attack on scalar product	Vector product	Derivation
Armit	—	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	×	×
Coulson	See below	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	×	×
Francis	—	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	—	×
Leaton	×	$\boldsymbol{a}.\boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$	×	×
Macbeath	See below	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	×	—
Patterson	×	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	×	×
Spiegel	×	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	×	×
Tallack	—	$\boldsymbol{a}.\boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$	—	×
S.M.P. Books	×	$\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$	×	×

2. Comments on individual books

A. P. Armit, Advanced Level Vectors

Armit represents vectors by directed line segments but never tells us what a vector is. He assumes properties of bases in two and three dimensions and his whole approach centres around the orthonormal basis $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ and the notation $\boldsymbol{a} = a_1 \boldsymbol{i} + a_2 \boldsymbol{j} + a_3 \boldsymbol{k}$. There is a chapter involving co-ordinate geometry in which vector equations of curves are given; the line is in the form $\boldsymbol{r} = \boldsymbol{a} + \lambda \boldsymbol{b}$ which seems to me to be a proper use of vectors, but I feel that it is unnecessary and unhelpful to say that the parabola is $\boldsymbol{r} = at^2 \boldsymbol{i} + 2at\boldsymbol{j}$. Perhaps the author has been forced into this by the London 'A' level syllabus. The introduction to scalar product and the worked examples are done independently of co-ordinate systems but *no* exercises of this type are given to the student. In fact, the vector is being used only for its concise notation, all the problems being worked in co-ordinates in the way they always have been. I would hesitate to use this book with a pupil who wanted to be a mathematician, but its wealth of exercises and down-to-earth style make it useful for the 'single subject' student.

A. E. Coulson, **An Introduction to Vectors**

My immediate reaction was to like this book as all the right things (for me that is) are done, but further acquaintance has turned me away, perhaps because everything is done piecemeal. Vectors in two dimensions and three dimensions are separated by scalar product and there are two discussions of linear dependence, again well separated. Bases are mentioned and the reader will probably form an intuitive but imprecise idea of what constitutes a basis. I did not think there were enough exercises; this was particularly true on the use of vectors without co-ordinate systems. The snippets of historical introduction to the various topics were for me the most interesting part of the book, which I do not feel I can recommend as a standard text book.

A. J. Francis, **Vectors and their Application to Geometry and Mechanics**

Like Armit's, this book has been inspired by the new London 'A' level syllabus. I liked it—it is small, compact and well laid out—but I found the lack of an index an irritant (and surely unnecessary?). Like Armit also, the book has plenty of exercises, but, unfortunately, nearly all of them use a co-ordinate system. This, together with the fact that vector products are not mentioned make it a suitable book for the single subject 'A' level student, but not for a serious double subject student,

E. H. Leaton, **Vectors**

This book virtually rules itself out as a suitable school text book by its lack of exercises. The author's approach differs from most of the other authors in that he seems to me to be an algebraist seeking application for his algebra. It is not surprising then to find vectors defined as n-tuples of scalars, scalar product defined as

$$\boldsymbol{a}.\boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

and a discussion of the vector space structure. Because of this shift in approach I think it is a book worth having in a school library so that pupils can supplement their usual classroom approach.

A. M. Macbeath, **Elementary Vector Analysis**

Macbeath thinks of vectors entirely in geometrical terms, there being no mention of applications to other branches of mathematics. He uses projections to discuss components and mentions the word 'base' more to acknowledge that a problem exists rather than to do anything about it. The scalar product is introduced, with worked examples and exercises on the $ab \cos \theta$ form before meeting a co-ordinate system. My pupils have always liked this book because of its clear layout, because topics are easy to find and because of its worked examples which show clearly that there are a variety of methods which may be used to solve a particular problem with great satisfaction to be gained in choosing an elegant and economic method. This is most certainly a book to have available for all pupils whether they be mathematics specialists or not.

E. M. Patterson, **Vector Algebra**

The author defines geometrical vectors (presumably to keep the usual algebraic definition separate) as translations in Euclidean three-dimensional space and then discusses linear dependence and basis in some detail before moving on to scalar product, vector product and derivatives. The style is to give a bare minimum of bookwork and to instruct by worked examples; this tends to make unusual (but not unwelcome) demands on school readers. One of the things I like about this book is that there is a good selection of exercises of a non-co-ordinate system type, especially in the chapter on scalar product. This is not an easy book for schoolboys, requiring as it does a certain maturity of approach, but I have been very pleased with it as a text book for second year sixth formers who are specialist (double subject) mathematicians.

M. R. SPIEGEL, **Theory and Problems of Vector Analysis**

This is really not a school book at all and certainly not a practical proposition as a text book. It is one of the Schaum's Outline Series which instruct by worked examples and then give supplementary exercises. The examples and exercises are well thought out and the diagrams are particularly helpful. My older and most able pupils have found it very useful as a reference library book and have liked working from it. The contents table earlier in this article is not appropriate for this book; it covers all the vector calculus that is normally found in a University course.

J. C. TALLACK, **Introduction to Elementary Vector Analysis**

Vectors are defined as representations of displacements. No discussion of linear dependence is given, the idea of components being handled by projections and the uniqueness of the decomposition into components is established geometrically. The scalar product is tackled two ways, first by using $\boldsymbol{a}.\boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$ and then by turning round and starting with $\boldsymbol{a}.\boldsymbol{b} = ab \cos \theta$. There are plenty of worked examples and probably just enough exercises, but the book does not go far enough for the double subject student. However, for the single subject student it is an attractive and well produced book.

S.M.P. Advanced Mathematics, Books 1, 2, 3, 4

The S.M.P. 'A' level books are not, of course, confined to vectors, but I have included them in this list because there is, to my mind, a good elementary discussion of bases and linear dependence to be found in them. There are plenty of exercises at all levels with varieties of applications, but I was sorry to find that nearly all the geometry was done using a co-ordinate system without any non-co-ordinate introductory exercises. These books are, of course, almost essential reading for all teachers and pupils whether they are doing the S.M.P. Course or not; the vector treatment merely reinforces this conclusion.

* * *

Statistics for Technology. By CHRISTOPHER CHATFIELD. Penguin Books, Harmondsworth, Middlesex, 1970. Pp. 359. 30*s*., paperback.

This is an attractively presented run-through of elementary statistical methods for technologists. The key ideas are simply and clearly introduced within the limitations of the intended audience which are evidently formidable. No knowledge of matrices or vectors is assumed; nor of complex variables, nor indeed any capacity for mathematical generality or sustained logical argument. Only simple formula manipulation is required. The level of understanding generated is correspondingly superficial and so much is stated without proof that it is difficult to comprehend what sort of picture of Statistics the technologist reader will have at the end of it all. The relatively few numerical examples are well placed to illustrate the mathematical manipulations but serve a purely formal, almost ritualistic, arithmetical purpose; they give very little idea of the complexity of circumstantial data. So many are moreover cast in terms of rocket and missile technology that a rather dismal impression is given; quite at variance with the actual richness and subtlety of the range of statistical applications.

There is the usual formal reference to the Computer but it is allowed no real influence; the orthogonal polynomials and orthogonal linear designs of the hand-machine era are given (rather than non-linear least squares, for instance) and there is hardly any discussion of computational methods. Random number generation is discussed in pre-computer terms and simulation does not seem to be mentioned in the text. Moreover, the computational facility provided by the computer has nowadays made multivariate analysis the usual form in Applied Statistics. The justification for discussing univariate problems of elementary statistics is consequently more the pedagogic virtue of introducing new ideas in simple cases rather than on their practical merits. The status of univariate manipulation is thus reduced; it has now to be judged by whether it leads to an insight into multivariate manipulation. Engineers and technologists are practical men and learn by using techniques as much as by formal logic, and this book, supplemented by practical statistical work, should give them the feel of univariate Statistics. But without a more advanced mathematical level (e.g., SMP 'A' level) it is difficult to believe that a basis can be had for any real idea of what multivariate methods achieve.

There are one or two curiosities like 'Cusum' techniques in quality control (quite inappropriate for picking up *changes* in a process) but the author wisely abjures transient window-dressing ('Information Theory', 'Games Theory', 'Bayesian Theory', 'Decision Theory') which more formal texts often feel obliged to mention. The book is based on lectures given to students at Bath University of Technology and they may well have reached their mathematical ceiling at 'O' level (SMP-style). Such students will get all they can master here without irrelevant matter.

Clearly it is not a suitable text for sixth formers taking Statistics at 'A' level but it is well worth buying for the school library. Pupils, and teachers, will find the very absence of mathematical proofs generates a corresponding virtue. That is: the utility of each proposition, the idea of its proof and the point of introducing it at a given juncture; all these are expressed more clearly and overtly than in many texts offering more formal logical continuity and more mathematically well-founded theory.

Queen Mary College, London D. E. BARTON

Mathematics. By DAVID BERGAMINI and the Editors of TIME–LIFE BOOKS. TIME–LIFE Books Pocket Edition, 1969. Pp. 187. 12*s*. 6*d*.

This book contrives not only to provide information concerning basic techniques in the major mathematical fields, but also to stimulate interest in mathematics by introducing the historical background and practical uses of these techniques. Considering the size of the volume and the copiousness of the photographs and diagrams accompanying the text, an astonishing amount of information is provided. This is arranged as a series of general topics which are considered in historical order. Biographical notes on the principal founders and theorists of each topic, and mention of the historical and modern consequences of mathematical theory, accompany the descriptions of advances in knowledge and ideas.

The book is more than amply supplied with photographs and diagrams, which tend to dominate the rest of the material. Each picture is accompanied by a paragraph of description, which is occasionally an unnecessary restatement of the text, or else introduces a topic which, though relevant, is not included in the main body of the book. The latter feature, although certainly adding to the reader's enjoyment, tends to divide his interest between the text and the copious notes on the illustrations. A happier solution would have been to incorporate these notes within the text with suitable references to the appropriate pictures. As it stands, the book has little cohesion in this respect. The pictures and diagrams, although all very attractive, are occasionally too elementary, and, more frequently, have only the most tenuous relevance to the subject. However, the book would lose much of its unusual character and appeal by their removal.

The text itself, by attempting both to give a brief historical summary and also to pursue the extensions of mathematical theories, is of necessity somewhat erratic. Unhappily it also tends to become somewhat repetitive, particularly over biographical details in the later sections. The treatment of mathematical ideas, however, is both attractive and concise, and the digressions arising from the development of these ideas are fascinating.

If there is lack of cohesion, there is no lack of interesting material in this book: examples of the part that mathematics plays in such diverse subjects as natural history, cartography, astronomy and art serve to involve the reader in the mathematical topics no matter what his interests may be.

The book is to be recommended as an interesting source of background material for those who are studying mathematics in secondary school.

Science VI Form, D. T. MILROY
Nottingham High School

Elementary Number Theory. By U. DUDLEY. W. H. Freeman and Company Ltd, Folkestone, 1970. Pp. ix+262. 70*s*.

Number theory, or at least the elementary part associated with divisibility of the integers and its applications, is rapidly becoming part of sixth form school mathematics. This is largely because the basic ideas involved are essential at an early stage in the build-up of mathematical knowledge. However, the subject is also useful in motivating later mathematical generalisations, is interesting and challenging in itself, and has the added advantage that it is easy to describe problems which are still unsolved. A book on the subject at this level must be judged against these facts.

This book, by Underwood Dudley, is undoubtedly very readable, interesting for its mathematics and its historical remarks, has a large and excellent collection of problems with some hints for solutions, and does not forget to mention unsolved problems. The author has chosen to omit connections with algebra and some people will consider this a mistake in the present approach to unification in mathematics.

A few points that occurred to me in reading the book are:

(a) The term 'well ordering principle' is introduced without explanation on page 3 when the term 'least (and greatest)-integer principle' is used for the same idea earlier.

(b) The book studiously avoids any mention of equivalence relations, the partition of the set of integers into congruence classes (mod m) and the sets of representatives (mod m).

(c) The term 'least residues' would, as used, be more clearly described as 'least non-negative residues'.

The chapter called 'Formulas for Primes' contains an interesting result of W. H. Mills, but has appeared too early to include the recent interesting work on a prime generating polynomial in a large number of real variables whose existence has been established as a result of the solution of the tenth Hilbert Problem by the Russian mathematician Matijasevich.

The mention of this fact may excite readers to find out more about what is implied. I can thoroughly recommend as a first stage in this process a careful and active reading of Underwood Dudley's excellent and well-printed book.

University of Glasgow J. HUNTER

Complex Numbers. By D. G. TAHTA (2, Pennsylvania Crescent, Exeter). Printed by Wm Pollard & Co. Ltd, Exeter. 20 pp.

In such a short booklet (20 pp.) there is obviously room for no more than a very brief introduction to complex numbers; but by aiming the book at very mathematically inclined readers, the author has managed to cover quite a wide range of subjects. His technique of giving references (or just saying 'find out about . . .') means that, if the book were conscientiously worked through, the reader, by following up all the lines of inquiry suggested, would end up with a fairly extensive and wide-ranging knowledge of the concepts involved—and of various side-issues, like matrices or finite arithmetic. These are brought in without comment: the author is assuming knowledge which the reader must then acquire in order to continue.

But the reader must be fairly interested in the subject beforehand, and quite practised at abstract visualisations also (especially in the section on matrices and quaternions). Such suggestions as 'investigate the relation $y = x^x$' or 'investigate the logarithm of i' will encourage a few but repel many.

The presentation is generally clear and pleasing, although the absence of page numbers and the elusiveness of the author's name annoy somewhat at first. In general, this book would be useful but hardly well used in a school's mathematical library—it might provoke some mathematics teachers to deeper-than-usual thought.

Second Year VI Form, K. R. Moore
Nottingham High School

Notes on Contributors

Jack Kiefer has spent most of his professional career at Cornell University, where he is now a Professor of Mathematics. As an undergraduate at Massachusetts Institute of Technology, he concentrated on engineering and economics until he became interested in mathematical statistics. After taking his S.B. and S.M. degrees there, he did graduate work and received his Ph.D. at Columbia University in 1952. His research has been primarily in statistics and probability theory. He has received a Guggenheim Fellowship.

T. J. Fletcher taught for some years in schools and a technical college. He is now an H.M. Inspector of Schools and regularly spends short periods of secondment at the Shell Centre for Mathematical Education in the University of Nottingham.

Nicholas Youd contributed his article while a member of the sixth form at St Paul's School, London. He gained an Exhibition in mathematics at Queens' College, Cambridge.

Hazel Perfect has been both a school teacher and a university teacher and is at present a Senior Lecturer in Pure Mathematics in the University of Sheffield. Her main interests are in matrix theory and in combinatorics. She is the author of two books, *Topics in Geometry* and *Topics in Algebra* (both published by Pergamon Press), which seek to narrow the gap between sixth form and university mathematics.

R. S. H. G. Thompson is a lecturer on the staff of the Mathematics Department at Imperial College, London University. He is especially concerned with the two-fold task of advising students on career prospects and of acting as liaison officer between the department and the many employers of mathematics graduates. It is perhaps interesting to note that when asked what he would like said about himself in this section of *Spectrum* he replied, 'Say that I have a great interest in students and in what happens to them when they come into our Mathematics Department'. His knowledge of universities and university life is wide, and it embraces a considerable number of years as a senior member of a small Chinese university.

MATHEMATICA PRIMA

No. 21 1970

This magazine is the work of senior boys in the St. Dunstan's College Mathematical Society. Copies are available at 5/– each (postage included).

All enquiries and orders should be sent to

A. J. Sherlock, Esq., M.A.,
Head of the Mathematics
Department,
St. Dunstan's College,
Catford,
LONDON S.E.6.

Approximation Theory

An international symposium held at the University of Lancaster in July 1969

Edited by A. Talbot
Mathematics Department
University of Lancaster, England

July 1970, viii+356 pp., 75s. (£3.75)

Here in one volume are some of the major statements of recent years on approximation theory which in this age of the computer has been developing at an unprecedented rate.

The book contains mainly recent unpublished work by eminent mathematicians from many countries—among them the first account in English of the powerful "method of functionals" developed by the Soviet mathematician E. V. Voronovskaya.

Graduate students or mathematicians using the approximations of function for practical applications such as computers, or those interested in certain branches of numerical analysis or functional analysis will find this a useful and stimulating study.

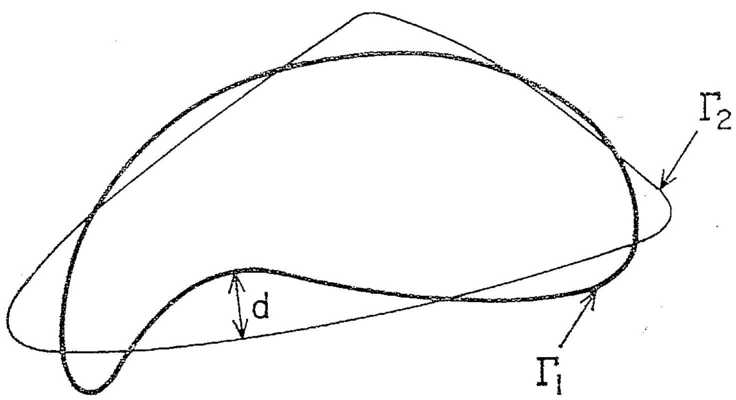


Illustration of general purpose curve fitting.

Academic Press
London and New York

Berkeley Square House
Berkeley Square
London W1X 6BA
England

111 Fifth Avenue
New York
N.Y. 10003
USA

Contents

PRICES (*postage included*)

	Individual annual subscription (1 volume of 2 issues)	Bulk annual subscriptions (5 or more copies of 1 volume of 2 issues)	Single issues
U.K. only	6/–	5/– each	3/6 each
Australia	$A1.10	$A0.90	$A0.60
U.S.A.	$1.00	$1.00	$0.70

Enquiries about rates, subscriptions, and advertisements should be directed to:

Oxford University Press, Education Department,
Walton Street, Oxford,

or to Oxford University Press Branches Overseas.

Printed in England by John Wright & Sons Ltd, Bristol