CONTENTS

POLYNOMIALS FOR BEST APPROXIMATION OVER SEMI-INFINITE AND INFINITE INTERVALS

Herbert E. Salzer

FOREWORD

It is well known that Chebyshev polynomials play a useful key role in giving close polynomial approximations to functions defined over finite intervals. By employing Chebyshev polynomials, we can approximate a polynomial of very high degree (which is practically the same as any continuous function) by another polynomial of much lower degree, with great accuracy. This approximation is a consequence of the property that, for any finite interval $[a,b]$, for fixed degree n, the Chebyshev polynomial (normalized to $[a,b]$) has a maximum deviation from zero which is less than that of any other polynomial of degree n, and having the same leading coefficient of x^n. But Chebyshev polynomials do not serve as well in approximations over infinite intervals, where an approximating polynomial will usually require a "damping" or "weight" factor. The simplest analogue of Chebyshev polynomials over infinite intervals would be polynomials having the property that, for fixed degree n, when multiplied by e^{-x} or e^{-x^2} for $[0,\infty]$ or $[-\infty,\infty]$ respectively, they have a maximum deviation from zero which is less than that of any other polynomial (damped by the factor e^{-x} or e^{-x^2}) of degree n, and having the same leading coefficient of x_n. The coefficients of these hitherto unknown polynomials are shown to satisfy a system of simultaneous transcendental equations whose general solution appears extremely difficult, especially since those equations involve also the abscissae of the extrema of the damped polynomials. Those equations are solved here only for the first few values of n. These polynomials are also shown to posses simple and instructive geometric pictures.

It is well known that the Chebyshev polynomials

$$C_n(x) \equiv \frac{1}{2^{n-1}} \cos(n \arccos x)$$

are highly useful in problems of best approximation over a finite region $[a,b]$, by virtue of the property that $C_n(x)$ is the polynomial of degree n, with leading coefficient 1, which has the smallest maximum deviation from 0, in the interval $[-1,1]$. The applications of this property are too numerous to cite here (one may consult the work of C. Lanczos and others). Now for the semi-infinite (range $[0,\infty]$) or infinite (range $[-\infty,\infty]$) interval, where there is a weight factor of e^{-x} or e^{-x^2} respectively (most common cases arising in practice, especially when dealing with a series of Laguerre or Hermite polynomials), one can ask a similar question for the purpose of "economization" of functions of the form $e^{-x}P(x)$ or $e^{-x^2}P(x)$ respectively, where $P(x)$ is a polynomial of high degree. Specifically, just as in the finite case, it is of interest to know:

I. The values of the coefficients of the polynomial $P_n(x)$, leading coefficient 1, such that $e^{-x}P_n(x)$ deviates the least from 0 in the interval $[0,\infty]$, *i.e.* with the smallest maximum absolute value of its extrema and initial value at $x = 0$.

II. The values of the coefficients of the polynomial $Q_n(x)$, leading coefficient 1, such that $e^{-x^2}Q_n(x)$ deviates the least from 0 in the interval $[-\infty,\infty]$, *i.e.* with the smallest maximum absolute value of its extrema.

The answers to I. and II. will provide useful sets of polynomials to be employed in a wide variety of approximation problems, including differential and integral equations over the ranges $[0,\infty]$ or $[-\infty,\infty]$. Of course, an important tool in such work will be the expression of x^n in terms of $P_n(x)$ or $Q_n(x)$, in order to obtain economical expressions for high degree polynomials by employing terms of lower degree. Also, Laguerre and Hermite polynomials in terms of $P_n(x)$ and $Q_n(x)$ respectively, should prove to be quite useful.

In a paper by J. Shohat, 'Sur les fonctions s'écartant le moins possible de zero dans un intervalle infini', *Annals of Ural (Ekaterinburg) University*, Vol. 1, 1922, pp. 1-33, there is an existence theorem which is related to the above problem, namely: If $\lim_{x=\mp\infty} f(x) = 0$, $\lim_{x=\mp\infty} p(x)x^i = 0$ $(i = 0, 1, 2, \ldots)$, $p(x) \geq p_0 > 0$ for $a \leq x \leq b$, then for every n there is a polynomial of the nth degree, $A_n(x)$, such that $f(x) - p(x)A_n(x)$ deviates the least from 0 in the interval $[-\infty,\infty]$. In Shohat's theorem, since the $f(x)$ is arbitrary for any given $p(x)$, choose $f(x) \equiv p(x)x^n$ and consider the best approximation polynomial of the $(n-1)$th degree, say $A_{n-1}(x)$. From the fact that $p(x)x^n - p(x)A_{n-1}(x)$ has the smallest maximum deviation from 0, it is apparent that for the weight factor $p(x)$, the polynomial $B_n(x)$ with leading coefficient 1, and such that $p(x)B_n(x)$ deviates least from 0, exists and is precisely $x^n - A_{n-1}(x)$. In particular, for $p(x) = e^{-x^2}$, the polynomials $Q_n(x)$ must exist. That the polynomials $P_n(x)$ exist seems intuitively obvious, but formally their existence follows readily from the property, established below by an independent constructive demonstration, that $Q_n(x)$ is an even function when n is even.*

*Strictly speaking, that is still on the assumption that there always is a solution to the system (B') given below.

Equations which, if solvable, are sufficient for determining $P_n(x)$ and $Q_n(x)$, are found by a device similar to that used in obtaining the Chebyshev polynomials, *i.e.* a certain condition of equality in height (or depth) of extrema. But in the Chebyshev case we are fortunate in having at our disposal the knowledge that we can actually exhibit such a polynomial, since $\cos n\phi$ which happens to be a polynomial in $x = \cos \phi$, has $n + 1$ equal extrema and end values for x in the interval $[-1,1]$. But there is no such ready mathematical information at our disposal for finding $P_n(x)$ and $Q_n(x)$.

Considering the semi-infinite case, suppose that there is a polynomial $P_n(x)$, with leading coefficient 1, such that $e^{-x}P_n(x)$ has the absolute value of each of its extrema equal to $|a_0|$, the absolute value at $x = 0$. Notice that $e^{-x}P_n(x)$ cannot have more than n extrema, since its derivative is $e^{-x}[P_n'(x) - P_n(x)]$, and the factor in brackets is an nth degree polynomial, while e^{-x} never vanishes. Thus for $n = 5$, the picture would appear as follows:

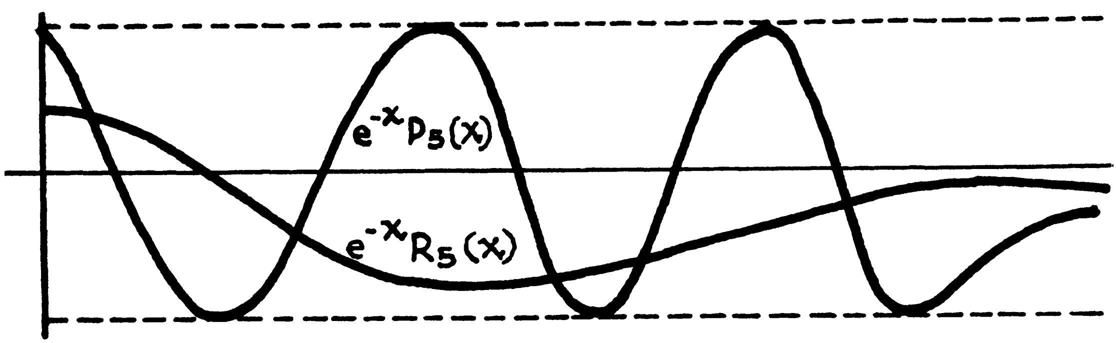


Figure 1

($P_n(x)$ may, without affecting the generality of the argument, begin with a negative value.) It follows that there cannot exist another $R_n(x)$, leading coefficient 1, such that $e^{-x}R_n(x)$ is nowhere as great, because from Figure 1, drawing in any such $e^{-x}R_n(x)$, it is apparent that there would be at least n intersections. Then $e^{-x}[P_n(x) - R_n(x)]$, with a polynomial factor of degree $n - 1$, vanishes in at least n points, which implies that $R_n(x) \equiv P_n(x)$. At the abscissae of the n maxima or minima, x_i $(i = 1,\ldots,n)$,

$$(A)\quad \begin{cases} P_n'(x_i) = P_n(x_i), & n \text{ equations,} \\ e^{-x_i}P_n(x_i) = (-1)^i a_0, & n \text{ equations.} \end{cases}$$

The $2n$ unknowns are x_i $(i = 1,\dots,n)$ and a_i $(i = n-1,\dots,0)$, where $P_n(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$. It would suffice to solve for the a_i's only, but from the appearance of the system (A), that does not seem possible.

Instead of specifying that $P_n(x)$ begins with x^n, we can normalize the $P_n(x)$ by multiplication by $\frac{1}{a_0}$, so that (keeping the same notation $P_n(x)$):

$$(A')\quad \begin{cases} P_n'(x_i) = P_n(x_i), \\ P_n(x_i) = (-1)^i e^{x_i} \end{cases}$$

System (A') lends itself somewhat more readily toward simple geometric picturization. The normalized polynomial $P_n(x)$ is tangent to e^x or $-e^x$, alternately, at the points x_i (Figure 2a). For, at x_i, $P_n(x_i) = (-1)^i e^{x_i}$ and also, $P_n(x_i)$ = its derivative at that point. But the derivative of $(-1)^i e^x$ is also $(-1)^i e^x$. Also, $P_n(x)$ can never exceed $(-1)^i e^x$ at any point, for then $e^{-x}P_n(x)$ would exceed 1 in absolute value. Since $P_n(x)$ is of the nth degree, the nth point of tangency cannot be followed by a maximum or minimum (Figure $2b$ or $2c$).

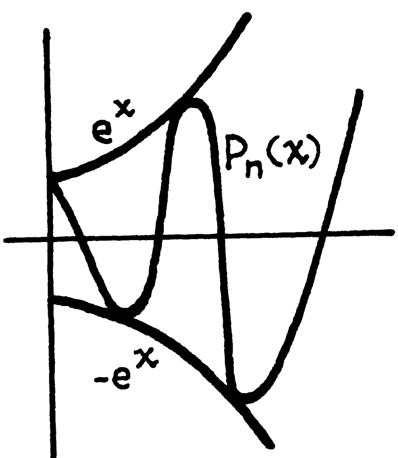


Figure $2a$

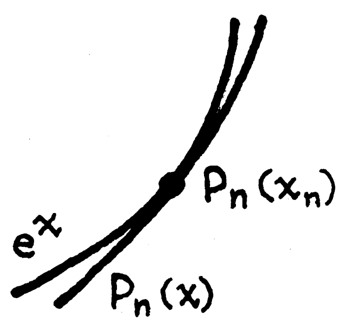


Figure $2b$

The equations (A) or (A') are difficult to solve. Even in the linear case, (where $P_1(x)$ is normalized) for $e^{-x}(a_1x + 1)$ to have a minimum absolute value of the maximum, which is equal to the initial value, it is necessary to solve

(1) $a_1x_1 + 1 = -e^{x_1} = a_1$, the solution of which is pictured in Figure $2d$

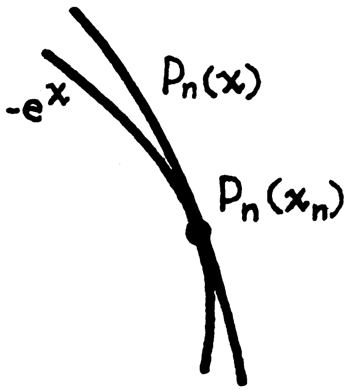


Figure $2c$

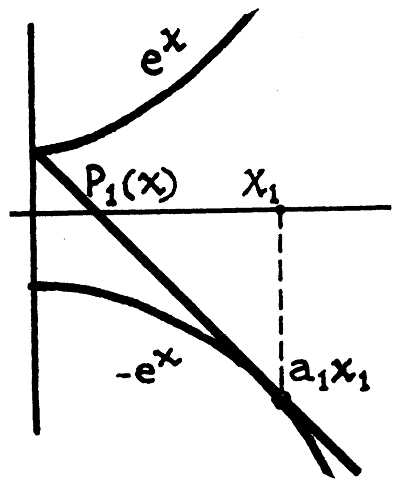


Figure $2d$

Eliminating x_1 as $\frac{(a_1 - 1)}{a_1}$, the transcendental equation to solve for a_1, is

(2) $$-a_1 = e^{1 - \frac{1}{a_1}}.$$

The solution of (2) yields $P_1(x) = -3.59112\ 148\ x + 1$, (or in the original form, $P_1(x) = x - 0.27846\ 4543$, and $x_1 = 1.27846\ 4543$).

The calculation of the coefficients of $P_2(x)$ (original form), involves the solution of the system

(3) $$x_1^2 + a_1x_1 + a_0 = 2x_1 + a_1 = -e^{x_1}a_0,$$

(4) $$x_2^2 + a_1x_2 + a_0 = 2x_2 + a_1 = e^{x_2}a_0.$$

From the two equations on the left side of (3) and (4),

(5) $$x_1 = \frac{1}{2}(2 - a_1 - \sqrt{a_1^2 - 4a_0 + 4});$$

$$x_2 = \tfrac{1}{2}(2 - a_1 + \sqrt{a_1^2 - 4a_o + 4}).$$

Substituting into the right hand side of (3) and (4), and taking the sum and product of the resulting equations, one has the following two equations in two unknowns, a_o and a_1:

$$(6) \qquad 4 = 2a_o e^{1-\frac{a_1}{2}} \sinh \tfrac{1}{2}\sqrt{a_1^2 - 4a_o + 4}$$

$$(7) \qquad -4a_o + a_1^2 = a_o^2 e^{2-a_1}.$$

The problem can be reduced to the solution of equations with one unknown, by setting

$$(8) \qquad x = -4a_o + a_1^2.$$

Then (6) becomes

$$(9) \qquad 2 = \pm\sqrt{x} \sinh \tfrac{1}{2}\sqrt{x + 4} .$$

Since x must be positive, $\sqrt{x+4}$ must be real; thus only the + sign holds in (9). Then, employing the numerical value of x, which equals 1.75691 536, since $a_o^2 = \frac{1}{16}(x - a_1^2)^2$, the quantity a_1 is found from the equation

$$(10) \qquad x = \frac{1}{16}(x - a_1^2)^2 e^{2-a_1}.$$

From (10), either

$$(11) \qquad 4\sqrt{x}\, e^{\frac{a_1}{2} - 1} = a_1^2 - x,$$

or

$$(11') \qquad 4\sqrt{x}\, e^{\frac{a_1}{2} - 1} = x - a_1^2.$$

Equation (11′) is rejected because it yields inadmissible values for a_1 and a_o. From (11), a_1 is obtained, and then a_o is had from (8). The result is $P_2(x) = x^2 - 1.62005\,5\,x + 0.21691\,54$. The abscissae of the points of tangency of $P_2(x)$ to $\pm e^x$ (here $P_2(x)$ is considered as normalized so that $a_o = 1$) are $x_1 = 0.61034\,87$ and $x_2 = 3.00970\,6$. (In a_o, a_1, x_o and x_1, the 7th significant figure is not guaranteed.) The function $e^{-x}P_2(x)$ was tested by comparison with $e^{-x}x^2$ and $e^{-x}L_2(x)$, where $L_2(x)$ is the Laguerre polynomial $x^2 - 4x + 2$. It was shown that the absolute value of the extrema of $e^{-x}P_2(x)$ are at the most only about $\frac{2}{5}$ as large as the greatest absolute value of $e^{-x}x^2$, and very much smaller than the greatest value of $e^{-x}L_2(x)$, (about $\frac{1}{10}$ the size in the interval from 0 to 5).

The infinite case has a similar solution. The derivative of $e^{-x^2}Q_n(x)$ cannot vanish at more than $n + 1$ points. Suppose $Q_n(x)$ were found such that $e^{-x^2}Q_n(x)$ had all its $n + 1$ extrema equal in absolute value. Thus for $n = 6$, the picture would appear as follows:

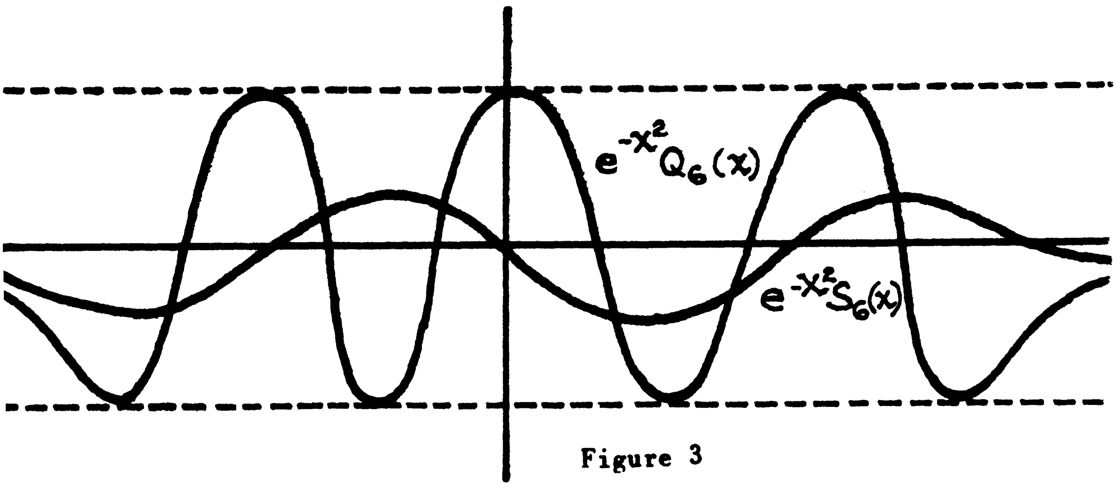


Figure 3

(There is no loss in generality in taking either $Q_n(x)$ or its negative). Again, there cannot be any polynomial $S_n(x)$, leading coefficient 1, such that $e^{-x^2}S_n(x)$ is everywhere less than these equal extrema, because from Figure 3, drawing in any such $e^{-x^2}S_n(x)$, there would be at least n intersections. Then $e^{-x^2}[Q_n(x) - S_n(x)]$, which has a polynomial factor only of the $(n-1)$th degree, would have n roots, implying that $S_n(x) \equiv Q_n(x)$.

The equations satisfied at the abscissae of the extrema, x_i ($i = 1, \ldots, n + 1$), are

$$(B)\begin{cases} Q_n'(x_i) = 2x_iQ_n(x_i), & n + 1 \text{ equations,} \\ e^{-x_i^2}Q_n(x_i) = (-1)^{i-1}e^{-x_1^2}Q_n(x_1), & n \text{ equations.} \end{cases}$$

Thus (B) consists of $2n + 1$ equations for the $2n + 1$ quantities

$a_0, a_1, \ldots, a_{n-1}, x_i, i = 1, \ldots, n, n + 1$. If $Q_n(x)$ is normalized by letting the absolute value of the extrema of $e^{-x^2}Q_n(x)$ equal 1 (also letting the first extremum be positive) then (keeping the same notation $Q_n(x)$):

$$(B') \begin{cases} Q_n'(x_i) = 2x_iQ_n(x_i), \\ Q_n(x_i) = (-1)^{i-1}e^{x_i^2}, \text{ with } e^{-x_1^2}Q_n(x_1) = 1. \end{cases}$$

Now (B') has $2n + 2$ equations, since it involves an extra unknown a_n.

The normalized polynomial $Q_n(x)$ is tangent to the curves $\pm e^{x^2}$, alternately, because at x_i, $Q_n(x) = (-1)^{i-1}e^{x^2}$, and also the slope of $Q_n(x)$ equals the slope of $(-1)^{i-1}e^{x^2}$

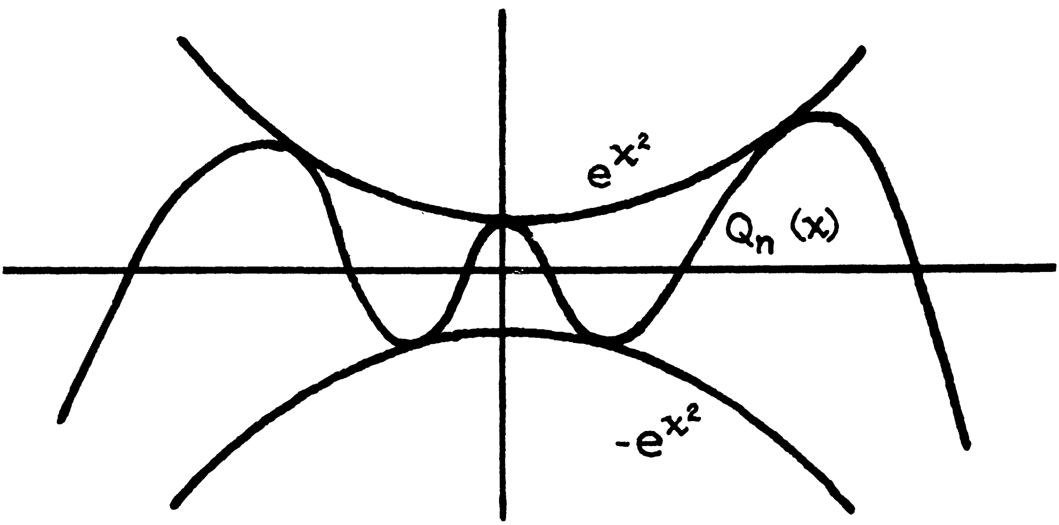


Figure 4*a*

The normalized straight line $Q_1(x)$ has the following picture:**

**That $a_0 = 0$, will be shown below.

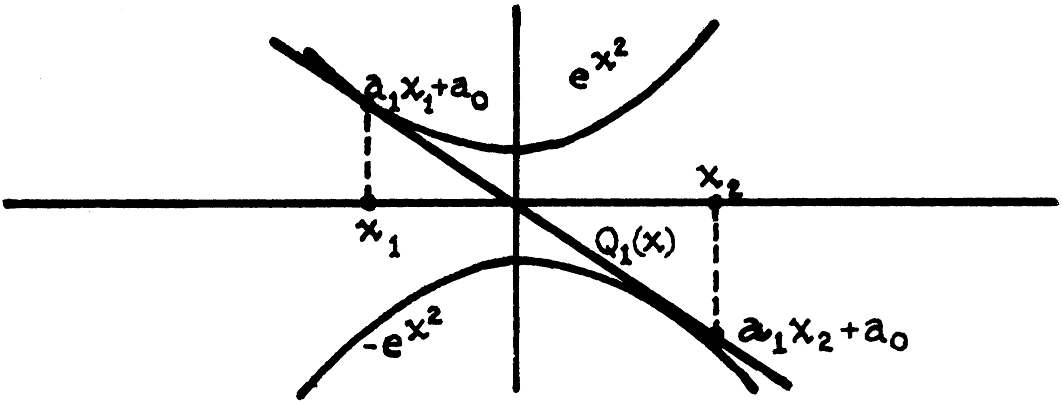


Figure 4b

To find $Q_1(x)$ (no normalization) formally from (B):

$$(12)\qquad 1 - 2x_1(x_1 + a_0) = 0,$$

$$(13)\qquad 1 - 2x_2(x_2 + a_0) = 0,$$

$$(14)\qquad e^{-x_1^2}(x_1 + a_0) = -e^{-x_2^2}(x_2 + a_0).$$

From (12) and (13), since $x_2 > x_1$,

$$(15)\ x_1 = \tfrac{1}{2}(-a_0 - \sqrt{a_0^2 + 2});\quad x_2 = \tfrac{1}{2}(-a_0 + \sqrt{a_0^2 + 2}).$$

Substituting into (14), and after obvious cancellations and rationalizations,

$$(16)\qquad e^{a_0\sqrt{a_0^2+2}} = a_0^2 - a_0\sqrt{a_0^2 + 2} + 1.$$

It is easily proven that $a_0 = 0$ is the only solution of (16). Of course, that $a_0 = 0$ is also obvious from the fact that, since $e^{-x^2} > 0$, $e^{-x^2}(x + a_0)$ will always have either a higher maximum or lower minimum than $e^{-x^2}x$, depending upon whether $a_0 > 0$ or $a_c < 0$. Still another way of seeing that $a_0 = 0$ is by drawing a straight line from the origin, tangent to e^{x^2} on the right, and another straight line tangent to $-e^{x^2}$ on the left. By symmetry, those two lines make equal angles with the y-axis, and hence must constitute a single straight line which is $Q_1(x)$ (normalized).

Before considering $Q_n(x)$ of higher order, it can be readily shown that $Q_n(x)$ is an odd polynomial for n odd, and an even polynomial for n even; in fact, $Q_{2m}(x) \equiv P_m(x^2)$. For $n = 2m + 1$, consider a curve of degree $m + 1$ passing through the origin and alternately tangent to $\pm e^{x^2}$ at $m + 1$ points on the positive side of the x-axis. Together with its reflection with respect to the origin, it constitutes a curve of degree $2m + 1$ which must be $Q_{2m+1}(x)$ (normalized). A simpler geometric picture of $Q_{2m+1}(x)$ is had by setting $u = x^2$, so that the resulting curve through the origin $\sqrt{u}\, T_m(u)$ is alternately tangent to $\pm e^u$. For $n = 2m$, consider a curve of degree m in the variable x^2, passing through the y-axis at $y = +1$ and alternately tangent to $\pm e^{x^2}$ at m points on the positive side of the x-axis. Together with its reflection with respect to the y-axis, it constitutes a complete curve of degree $2m$, which must be $Q_{2m}(x)$, (normalized). Considering $u = x^2$, from the preceding construction of $Q_{2m}(x)$, it is seen that $Q_{2m}(x) = P_m(u) = P_m(x^2)$.

To find $Q_3(x) = x^3 + a_1 x$ (no normalization) it is sufficient to solve

(17) $$e^{-x_1^2}(x_1^3 + a_1x_1) = -e^{-x_2^2}(x_2^3 + a_1x_2),$$

(18) $$3x_1^2 + a_1 = 2x.(x_1^3 + a_1x_1),$$

(19) $3x_2^2 + a_1 = 2x_2(x_2^3 + a_1x_2)$, where x_1 and x_2 denote the positive abscissae of the points of tangency Setting $x^2 \equiv X$, so that $x_i^2 \equiv X_i$,

(20) $$X_1 = \tfrac{1}{4}(3 - 2a_1 - \sqrt{4a_1^2 - 4a_1 + 9});$$

$$X_2 = \tfrac{1}{4}(3 - 2a_1 + \sqrt{4a_1^2 - 4a_1 + 9}).$$

Substituting into (17),

(21) $$\sqrt{3 - 2a_1 - \sqrt{4a_1^2 - 4a_1 + 9}}\; e^{\frac{1}{2}\sqrt{4a_1^2 - 4a_1 + 9}}(3 + 2a_1 - \sqrt{4a_1^2 - 4a_1 + 9}) =$$
$$- \sqrt{3 - 2a_1 + \sqrt{4a_1^2 - 4a_1 + 9}}\;(3 + 2a_1 + \sqrt{4a_1^2 - 4a_1 + 9}).$$

Equation (21) can be further simplified by multiplying first by $\sqrt{3 - 2a_1 - \sqrt{4a_1^2 - 4a_1 + 9}}$, which introduces an extraneous root

$a_1 = 0$,* and once more by $3 + 2a_1 - \sqrt{4a_1^2 - 4a_1 + 9}$, which again has $a_1 = 0$ as the only extraneous root. When all simplifications of radicals have been made, equation (21) for a_1 becomes

$$(22)\quad -4a_1\sqrt{-8a_1} = e^{\frac{1}{2}\sqrt{4a_1^2-4a_1+9}}\left(27 + 4a_1^2 - (9+2a_1)\sqrt{4a_1^2 - 4a_1 + 9}\right).$$

From (22), $a_1 = -0.82018\ 29$, so that $Q_3(x) = x^3 - 0.82018\ 29x$. The positive abscissae of the points of tangency to $\pm e^{x^2}$ (here referring to the normalized $Q_3(x)$), are $x_1 = 0.4390500$, $x_2 = 1.45856\ 71$. (In a_1, x_1 and x_2, the last significant figure is not guaranteed.) The function $e^{-x^2}Q_3(x)$ was tested by comparison with $e^{-x^2}x^3$ and $e^{-x^2}H_3(x)$, where $H_3(x)$ is the Hermite polynomial $x^3 - 1.5x$. It was shown that the absolute value of the extrema of $e^{-x^2}Q_3(x)$ are no more than about $\frac{1}{2}$ as large as the greatest absolute value of $e^{-x^2}x^3$, and only about $\frac{2}{5}$ as large as the greatest absolute value of $e^{-x^2}H_3(x)$.

In order to be of substantial use in approximation problems, $P_n(x)$ and $Q_n(x)$ must be known up to a much higher order than $n = 2$. However, the calculation of the coefficients of $P_n(x)$ and $Q_n(x)$ for $n > 2$ seems to be a task, the difficulty of which increases very rapidly with n. Indeed, finding a suitable general method for any n, seems almost like a challenge problem.

*$a_1 = 0$ cannot be an admissible a_1, since then (18) and (19) become $2x^4 - 3x^2 = 0$, and the roots $x_1 = 0$, $x_2 = \sqrt{1.5}$ contradict (17).

Computation Laboratory
National Bureau of Standards

THEOREMS ON QUADRATIC RESIDUES*

Albert Leon Whiteman

1. *Introduction.* As a by-product of his investigations on the class number of quadratic forms, Dirichlet,[1] in 1839, established the following remarkable theorem.

Theorem 1. If p is a prime of the form $4j + 3$, then among the integers

$$1,\ 2,\ 3,\ \ldots,\ \frac{(p-1)}{2},$$

there are more quadratic residues of p than non-residues.

Dirichlet's proof of this proposition involves the use of infinite series. Despite its simple arithmetic nature no independent proof has ever been discovered.[2]

The object of this paper is to derive two theorems which are equivalent to Theorem 1 and which appear to be of interest in themselves. These are

Theorem 2. If p is an odd prime, then

$$\sum_{n=1}^{p-1} \cot \frac{\pi n^2}{p} \gtrless 0,$$

according as $p \equiv 1$ or $3 \pmod 4$.

Theorem 3. If p is an odd prime, then

$$\sum_{n=1}^{\frac{(p-1)}{2}} \left[\frac{n^2}{p}\right] \gtrless \frac{(p-1)(p-5)}{24},$$

according as $p \equiv 1$ or $3 \pmod 4$.

Here, as is usual, $[x]$ denotes the greatest integer not exceeding x.

2. *Lemmas.* We shall need four lemmas.

Lemma 1. Let p be an odd prime. Then

$$\sum_{n=1}^{\frac{p-1}{2}} \left(\frac{n}{p}\right) = -\frac{2-\left(\frac{2}{p}\right)}{p} \sum_{n=1}^{p-1} n\left(\frac{n}{p}\right),$$

*Presented under a different title to the Maryland-District of Columbia-Virginia Section of the Mathematical Association of America at College Park, Maryland, December 6, 1947.

[1] L. Dirichlet, Recherches sur diverses applications de l'analyse infinitésimale à la théorie des nombres, Journal für die reine und angewandte Mathematik, Vol. 19 (1839), p. 324 and Vol. 21 (1841), pp. 1-134. An excellent exposition of Dirichlet's methods is given in the first volume of Landau's Vorlesungen über Zahlentheorie.

[2] A form of Dirichlet's proof which is not expressed in terms of class number theory has been given by Kai-Lai Chung, Note on a theorem on quadratic residues, Bulletin of the American Mathematical Society, Vol. 47 (1941), pp. 514-516.

where $\left(\frac{n}{p}\right)$ denotes the Legendre symbol.

When p is of the form $4j + 3$, we have, on the one hand,

$$(2) \qquad \sum_{n=1}^{p-1} n\left(\frac{n}{p}\right) = \sum_{n=1}^{\frac{p-1}{2}} n\left(\frac{n}{p}\right) + \sum_{n=1}^{\frac{p-1}{2}} (p - n)\left(\frac{p-n}{p}\right)$$

$$= 2\sum_{n=1}^{\frac{p-1}{2}} n\left(\frac{n}{p}\right) - p\sum_{n=1}^{\frac{p-1}{2}} \left(\frac{n}{p}\right)$$

On the other hand,

$$(3) \qquad \sum_{n=1}^{p-1} n\left(\frac{n}{p}\right) = \sum_{n=1}^{\frac{p-1}{2}} 2n\left(\frac{2n}{p}\right) + \sum_{n=1}^{\frac{p-1}{2}} (p - 2n)\left(\frac{p-2n}{p}\right)$$

$$= 4\left(\frac{2}{p}\right)\sum_{n=1}^{\frac{p-1}{2}} n\left(\frac{n}{p}\right) - p\left(\frac{2}{p}\right)\sum_{n=1}^{\frac{p-1}{2}} \left(\frac{n}{p}\right).$$

When p is of the form $4j + 1$, both members of equation (1) equal zero. Combining (2) and (3), we get Lemma 1.

Lemma 2. If k is a positive integer which does not divide the integer m, then

$$\sum_{n=1}^{k-1} n \sin \frac{2\pi mn}{k} = -\frac{k}{2} \cot \frac{\pi m}{k}.$$

From the identity

$$\sum_{n=1}^{k-1} nx^n = \frac{kx^k}{x-1} - \frac{x(x^k - 1)}{(x-1)^2}, \quad x \neq 1,$$

it follows that

$$\sum_{n=1}^{k-1} n \sin \frac{2\pi mn}{k} = I \sum_{n=1}^{k-1} ne^{\frac{2\pi imn}{k}} = I \frac{k}{e^{\frac{2\pi im}{k}} - 1}$$

$$= Ik \frac{e^{\frac{-2\pi im}{k}} - 1}{\left|e^{\frac{2\pi im}{k}} - 1\right|^2} = \frac{-k \sin \frac{2\pi m}{k}}{4 \sin^2 \frac{\pi m}{k}}$$

$$= -\frac{k}{2} \cot \frac{\pi m}{k},$$

where $I(u)$ denotes the imaginary part of u.

The next lemma is a classical result about Gauss sums.

Lemma 3. If p is an odd prime and n is an integer, then

$$\sum_{m=1}^{p-1} \sin \frac{2\pi m^2 n}{p} = 0 \text{ if } p \equiv 1 \pmod 4,$$
$$= \left(\frac{n}{p}\right)\sqrt{p} \text{ if } p \equiv 3 \pmod 4.$$

A simple proof of this lemma has recently been given by Estermann[3].

Lemma 4. If $((x))$ is defined for non-integral x by

$$((x)) = x - [x] - \frac{1}{2},$$

then

$$((x)) = -\sum_{n=1}^{\infty} \frac{\sin 2\pi nx}{n\pi}.$$

This is a well known Fourier series expansion.

3. *Proof of Theorems* 1, 2 *and* 3. We shall first establish the identity

$$(4) \qquad \frac{1}{4\sqrt{p}} \sum_{m=1}^{p-1} \cot \frac{\pi m^2}{p} = \sum_{n=1}^{\frac{p-1}{2}} \left(\frac{n^2}{p}\right) - \frac{(p-1)(p-5)}{24},$$

where p denotes an odd prime. From Lemma 2, we have

$$\cot \frac{\pi m}{p} = -\frac{2}{p} \sum_{n=1}^{p-1} n \sin \frac{2\pi mn}{p}.$$

Hence, using Lemma 3, we obtain when p is of the form $4j + 3$,

$$\sum_{m=1}^{p-1} \cot \frac{\pi m^2}{p} = -\frac{2}{p} \sum_{m=1}^{p-1} \sum_{n=1}^{p-1} n \sin \frac{2\pi m^2 n}{p}$$

$$(5) \qquad = -\frac{2}{p} \sum_{n=1}^{p-1} n \sum_{m=1}^{p-1} \sin \frac{2\pi m^2 n}{p}$$

$$= -\frac{2}{\sqrt{p}} \sum_{n=1}^{p-1} n \left(\frac{n}{p}\right)$$

$$= -2\sqrt{p} \sum_{n=1}^{p-1} \left(\left(\frac{n}{p}\right)\right) \left(\frac{n}{p}\right).$$

Since $\left(\left(\frac{n}{p}\right)\right)$ is periodic in n modulo p and $((-x)) = -((x))$, it follows that

$$\sum_{m=1}^{p-1} \cot \frac{\pi m^2}{p} = -2\sqrt{p}\left[\sum_{n=1}^{\frac{p-1}{2}} \left(\left(\frac{n^2}{p}\right)\right) - \sum_{n=1}^{\frac{p-1}{2}} \left(\left(\frac{-n^2}{p}\right)\right)\right]$$

[3] T. Estermann, On the sign of the Gaussian sum, Journal of the London Mathematical Society, Vol. 20 (1945), pp. 66-67.

$$= -4\sqrt{p}\sum_{n=1}^{\frac{p-1}{2}}\left(\left(\frac{n^2}{p}\right)\right)$$

$$= -4\sqrt{p}\left[\sum_{n=1}^{\frac{p-1}{2}}\frac{n^2}{p} - \sum_{n=1}^{\frac{p-1}{2}}\left[\frac{n^2}{p}\right] - \sum_{n=1}^{\frac{p-1}{2}}\frac{1}{2}\right]$$

$$= 4\sqrt{p}\left[-\frac{(p-1)(p-5)}{24} + \sum_{n=1}^{\frac{p-1}{2}}\left[\frac{n^2}{p}\right]\right],$$

from which (4) follows at once in the case in which p is of the form $4j + 3$. When p is of the form $4j + 1$, both members of equation (4) are equal to zero. The proof of (4) is thus complete.

Combining the results in Theorem 1, Lemma 1, (4) and (5), we deduce Theorems 2 and 3.

For the sake of completeness we include the proof of Theorem 1. Using Lemmas 3 and 4, we have,

$$\sum_{m=1}^{p-1}{}' m\left(\frac{m}{p}\right) = p\sum_{m=1}^{p-1}{}'\left(\left(\frac{m}{p}\right)\right)\left(\frac{m}{p}\right)$$

$$= -\frac{p}{\pi}\sum_{m=1}^{p-1}{}'\left(\frac{m}{p}\right)\sum_{n=1}^{\infty}\frac{\sin\frac{2\pi mn}{p}}{n}$$

(6) $$= -\frac{p}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sum_{m=1}^{p-1}{}'\left(\frac{m}{p}\right)\sin\frac{2\pi mn}{p}$$

$$= -\frac{p}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sum_{m=1}^{p-1}{}'\sin\frac{2\pi m^2 n}{p}$$

$$= -\frac{p\sqrt{p}}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{n}{p}\right).$$

It remains to prove that if p is of the form $4j + 3$, then

(7) $$\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{n}{p}\right) > 0.$$

The sum of the series in (7) is different from zero since the sum in the left side of (6) is different from zero. We have for $s > 1$, the Euler factorization,

(8) $$\left[\sum_{n=1}^{\infty}\frac{1}{n^s}\left(\frac{n}{p}\right)\right]\prod_q\left[1 - \frac{1}{q^s}\left(\frac{q}{p}\right)\right] = 1,$$

where q runs through the sequence of primes. The series in (8) is uniformly convergent for $s \geq 1$ and its sum is therefore continuous at $s = 1$. The product in (8) is convergent for $s > 1$ and each factor is positive. Hence, the result stated in (7) follows. Combining Lemma 1, (6) and (7), we obtain Theorem 1.

University of Southern California

TEACHING OF MATHEMATICS

Edited by

Joseph Seidlin and C. N. Shuster

This department is devoted to the teaching of mathematics. Thus articles on methodology, exposition, curriculum, tests and measurements, and any other topic related to teaching, are invited. Papers on any subject in which you, as *a teacher,* are interested, or questions which you would like others to discuss, should be sent to Joseph Seidlin, Alfred University, Alfred, New York.

AMERICA'S GREATEST

Duane Studley

Since the use of atomic bombs and radar in the last war physics has attracted rather wide not to say notorius attention. The general public is not very well informed about what physics is and this ignorance does not help in fulfilling the new duties of citizenship in the modern world. As a step in the right direction it might be well to become acquainted with the history of physics and with some of the famous physicists. To start let's meet the greatest physicist that America has produced. His name is not widely known; perhaps you have never heard it before. Josiah Willard Gibbs, professor of mathematical physics at Yale University 1871 - 1903, is recognized as preeminent among all the physicists who have lived and worked in the western hemisphere.

Born February 11, 1839 in New Haven, Connecticut, he grew and lived in an environment dominated by Yale. His father was professor of theology and sacred literature in the Yale Divinity School; his mother the daughter of a Princeton professor. At the age of two he had scarlet fever and as a result was frail during boyhood. Hopkins Grammar School supplied his first formal instruction. He had four sisters, two of whom died young. The New Haven of his first years was the New Haven of Noah Webster, Benjamin Silliman, John Trumbull and J. G. Percival.

In 1854 aged fifteen he entered Yale. It didn't change life much because home was so like the College. He had many friends and spent much time hiking and camping out. As he grew older his friends found him "a universal bugbear". In his senior year he was elected Phi Beta Kappa, won a prize and a scholarship for postgraduate study. Also in the class of 1858 was Addison Van Name who later became his brother-in-law.

Gibbs continued his graduate work and at the outbreak of the civil war was working towards his Master's degree. Military service did not attract him. Another prize assured continuation of graduate work so he continued toward the doctor's degree. This he received in 1863 presenting "On the Form of the Teeth of Wheels in Spur Gearing" for his dissertation. His mother had died during his early college years and his father died in 1861. He and his sisters continued to live in the old High Street house. During these years he worked on several inventions, a railroad brake for which he received a patent, an engine governor was another. Yankee inventiveness was at its peak at the time. The Connecticut Academy of Sciences accepted him as a member.

As soon as the estate was settled he and his sisters, Anna and Julia, went abroad. First to Paris where Gibbs entered the Sorbonne. Enrolled, he took courses given by Liouville, Serret, Chasles, Duhamel, Darboux, Briot, Delaunay and Bertin. Next the South of France, then on to Berlin. Van Name arrived to marry Julia and shortly after their marriage in Berlin they left for New Haven. Gibbs himself never married. Anna and Willard stayed on, Willard entering the University of Berlin. The teachers included Dove, Du Bois-Reymond, Erman, Kronecker, Kummer, Kundt, Magnus, Paalzow, Quincke, Weierstrass and Forster among whom were three of the greatest mathematicians of the day.

Heidelberg completed the tour of universities. There he had the chance to know Helmholtz, Kirchhoff, Bunsen, Cantor, Hesse and Weber. After return to Paris and a visit to the Riviera they sailed home in June 1869. Gibbs was now thirty years old. His brother-in-law had been made Librarian of Yale.

In England the need of physics laboratories had been felt by the Universities and in 1871 James Clerk Maxwell was appointed the first professor of experimental physics at Cambridge. At the same time Yale was studying its own needs and decided to establish a chair of physics. Gibbs was appointed professor of mathematical physics. His first paper was published in the Transactions of the Connecticut Academy in 1873. It was titled Graphical Methods in the Thermodynamics of Fluids. Six months later the second paper "A Method of Geometrical Representation of the Thermodynamic Properties of Substances by Means of Surfaces." appeared.

Maxwell read the second paper and recognized its value. He actually constructed three surfaces representing the thermodynamical properties of water and sent one of these "statues of water" to Gibbs. The recognition by Maxwell was about all the recognition Gibbs had at this time. At Yale there was a tendency to deprecate his chair and in fact he didn't receive a salary during the first ten years of his professorship. Maxwell went so far as to write a chapter in his Theory of Heat on Gibbs's work.

In June 1874 Gibbs presented a paper "The Principles of Thermodynamics as Determining Chemical Equilibrium". No funds were available to publish it. A subscription was taken up and finally the 300 page paper, consisting of about 160,000 words, was published in three parts. The Great Paper as it has come to be called was in spite of its length highly condensed. Many have claimed that it is easier to rediscover Gibbs than to read him. Often called "The Equilibrium of Heterogeneous Substances" this paper founded and practically exhausted the new science of chemical thermodynamics. Ignorance of this work has caused vast waste of time and effort on the part of many investigators. Gibbs made little effort to popularize his work and what effort has been made by others has not had a very wide effect. Ostwald translated this paper into German and Le Chatelier translated it into French. Through the years this paper has proved very important especially in chemical industry and in metallurgy. It has even been applied in physiology and economics. One section of five pages called the phase rule has been of outstanding importance. The whole theoretical structure was built on two hypotheses: the energy of the world is constant and the entropy of the world tends to a maximum.

Van der Waals read the paper and started studies in the new chemical physics. One of his students Bakhuis Roozeboom applied the phase rule to the study of steel. Holland and later Germany nurtured the new science. During the first world war chemical thermodynamics proved vital in the production of explosives.

Gibbs's favorite maxim was the whole is simpler than its parts. Contemporary science depends, not on single points of knowledge, but on clusters and combinations. Gibbs's methods were most effective in organizing and ordering knowledge. He insisted that mathematics is a language.

After the founding of Johns Hopkins University in 1876 Gibbs was invited to deliver a series of lectures on Theoretical Mechanics there. He accepted and lectured using a new notation on which he was working. This notation had its origin in Grassmann's Ausdehnungslehre and Sir William Rowan Hamilton's quaternions. Gibbs's formulation however was superior as a tool of physics associating as it did physical notions with mathematical notions. It has come to be called vector analysis and it has outlasted all rival methods.

In 1881 Gibbs was offered a professorship at Johns Hopkins. He was going to accept but his colleagues didn't want him to leave Yale so he stayed. After ten years he at last received a salary. His vector analysis appeared in a privately printed volume *Elements of Vector Analysis* for his students but he never wrote and published a book about it. That he left to one of his students, E. B. Wilson, who wrote a text which was published in 1901.

A controversy developed with P. G. Tait. Tait championed quaternions and Gibbs defended his notation. It was quite unlike the Newton-Leibniz controversy which was over priority. As far as modern theoretical physics is concerned Gibbs won.

Among Gibbs's students there were Lee De Forest, the inventor, and Irving Fisher, the economist. His phase rule has been called the Rosetta Stone of science. Gradually he gained recognition; the peak of his reputation will be posthumous and has not yet arrived. His vector analysis has been an indispensable tool in radar.

In 1879 he had been elected to the National Academy of Sciences and from time to time he read papers before it. His last work was in statistical mechanics and its quality is measured by the fact that it has withstood the impact of the modern quantum theory. Many nineteenth century theories have been banished by quantum theory. His book *Elementary Principles of Statistical Mechanics* was published by Scribners in 1902.

Gibbs died on April 28, 1903 of an intestinal obstruction but he will live forever in his work. Millikan has said "Gibbs lives because, profound scholar, matchless analyst that he was, he did for statistical mechanics and for thermodynamics what Laplace did for celestial mechanics and Maxwell did for electrodynamics, namely, made his field a well-nigh finished theoretical structure ...."

PROJECTIVE GEOMETRY

H. S. M. Coxeter

1. *Introduction.* The plane geometry of Euclid may be described as the geometry of the straight-edge and compasses (see Chapter 3). It was proved by the Danish geometer Georg Mohr (1672) that nothing is lost by discarding the straight-edge and using the compasses alone. For instance, we can still construct a point mid-way between two given points, though the procedure is quite complicated. We naturally ask how much can still be done if we discard the compasses instead, and use the straight-edge alone. At first sight it looks as if nothing at all will be left: we cannot even carry out the construction in Euclid's first proposition. But in fact a very beautiful and intricate collection of propositions emerges: propositions of which Euclid never dreamed, because his interest in circles led him in a different direction. Some of these propositions were discovered by Pappus of Alexandria in the fourth century A.D., and some by two Frenchmen of more recent times: the philosopher Pascal and the architect Desargues, who both died in 1662, though Pascal was the younger by thirty years. Meanwhile, the closely related subject of perspective had been studied by such artists as Leonardo da Vinci and Dürer. But the systematic development of projective geometry is not much more than a hundred years old: Poncelet's *Traité des proporiétés projectives des figures* appeared in 1822, and von Staudt's *Geometrie der Lage* in 1847.

2. *The axioms of incidence.* As Miss Mazziotta remarked, 'A geometric proof is ... based on argument from stated specific assumptions, through a series of logical steps, to an inevitable conclusion.' The specific assumptions, called *axioms*, are concerned with a few primitive concepts which we do not attempt to define. All other concepts used are defined in terms of these. Mario Pieri (1860-1913) made the astonishing discovery that the only primitive concepts needed for the real projective plane (including the theory of cyclic order) are *point*, *line*, and the relation of *incidence*.

It is important to remember that, unlike Euclid's line (which is finite but capable of extension), Pieri's line is a single complete entity. We need not be worried by the idea of its being 'infinitely long,' for in this kind of geometry length is not defined.

Actually, Pieri left the number of dimensions unrestricted; but if we restrict it to two, we find that ten axioms suffice. These are conveniently interspersed with definitions which have the desirable effect of rendering the axioms reasonably concise. The following neat way of expressing the first six is due to Karl Menger*:

*Independent self-dual postulates in projective geometry, *Reports of a Mathematical Colloquium* (2), 8 (1949), pp. 81-87.

AXIOM I *Two distinct* $\genfrac{}{}{0pt}{}{\textit{points}}{\textit{lines}}$ *are incident with at least one* $\genfrac{}{}{0pt}{}{\textit{line}}{\textit{point.}}$

II· *Two distinct points cannot both be incident with two distinct lines.*

III *There exist two points and two lines such that each of the points is incident with just one of the lines.*

IV *There exist two points and two lines (the points not incident with the lines) such that the join of the points is incident with the intersection of the lines.*

Definitions. The unique $\genfrac{}{}{0pt}{}{\text{line}}{\text{point}}$ incident with two $\genfrac{}{}{0pt}{}{\text{points } P,\ Q}{\text{lines } p,\ q}$ is called their $\genfrac{}{}{0pt}{}{\text{join}}{\text{intersection}}$ and is denoted by $\genfrac{}{}{0pt}{}{PQ}{p\cdot q.}$ Any $\genfrac{}{}{0pt}{}{\text{points}}{\text{lines}}$ incident with one $\genfrac{}{}{0pt}{}{\text{line}}{\text{point}}$ are said to $\genfrac{}{}{0pt}{}{\text{lie on}}{\text{pass through}}$ the $\genfrac{}{}{0pt}{}{\text{line}}{\text{point}}$ or to be $\genfrac{}{}{0pt}{}{\text{collinear}}{\text{concurrent}}$ or to belong to a $\genfrac{}{}{0pt}{}{\text{range}}{\text{pencil.}}$ The symbols are easily combined; *e.g.*, the join of $p\cdot q$ and $r\cdot s$ is $(p\cdot q)(r\cdot s)$. The figure of three points and three lines, where each line joins two of the points, is called a *triangle*. More specifically, the triangle PQR consists of three points P, Q, R and three lines QR, RP, PQ. Similarly, the figure of four points P, Q, R, S and four lines PQ, QR, RS, SP is called the *simple quadrangle** $PQRS$. The two remaining joins PR and QS are called the *diagonal lines*, and the two remaining intersections $PQ\cdot RQ$ and $PS\cdot QR$ the *diagonal points*. One quadrangle is said to be *inscribed* in another if its four vertices lie on respective sides of the other in their proper cyclic order.

AXIOM V *If one simple quadrangle is inscribed in another in such a way that one ofits diagonal points lies on a diagonal line of the other, then its second diagonal point lies on the second diagonal line of the other.*

(See Fig. 1, where the quadrangle $pqrs$ is inscribed in $PQRS$, and the diagonal points $p\cdot r$, $q\cdot s$ lie on the diagonal lines QS, PR.)

AXIOM VI *The intersection of the two diagonal lines of a simple quadrangle cannot lie on the join of the two diagonal points.*

3. *The principle of duality.* This is not the whole story, but we can go a long way without appealing to the remaining four axioms, so let us pause and take stock of the situation. We immediately observe an elegant symmetry: each axiom retains its meaning when we interchange

points with *lines* and *joins* with *intersections*.

*The word *simple* is used to distinguish this from the *complete* quadrangle used in most treatments of projective geometry, such as Coxeter, *The Real Projective Plane* (New York, 1949), p. 14.

This symmetry is the basis for the so-called *principle of duality*. Since these substitutions can be made in the axioms, they can still be made in any theorem deduced from the axioms. In fact, after proving any theorem we can at once assert the dual theorem, because we know that a valid proof of the dual theorem could be written down mechanically by dualizing every step in the proof of the original theorem.

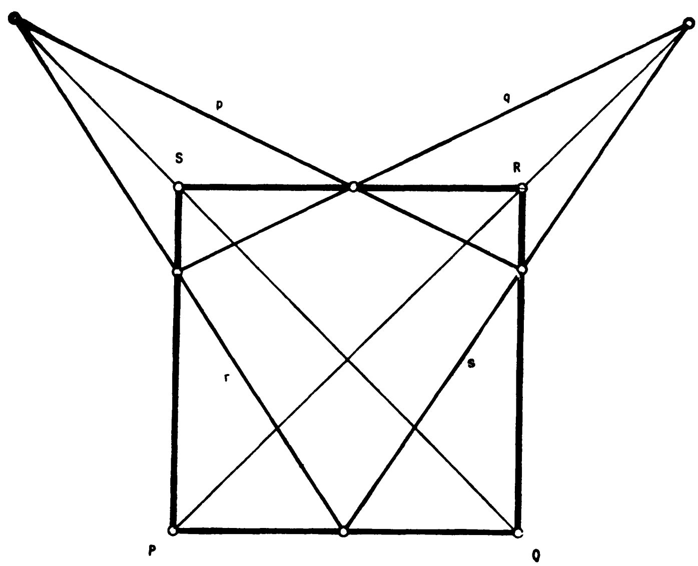


Figure 1

This principle was understood to a limited extent by Maurolycus (1494-1575), Snellius (1581-1626), Brianchon((1760-1854) and Poncelet (1788-1867); but it was first plainly expounded by Gergonne (1771-1859), who gave it the name *duality*.

4. *The projective plane and the affine plane.* To what extent do the above axioms agree with our usual conception of plane geometry? Anyone who plays about with a straight-edge will find Axioms II, III, IV, and VI intuitively obvious. Axiom V can be deduced from simple properties of planes in three-dimensional space. The upper half of Axiom I is merely Euclid's Postulate I, which we can accept without question. But the lower half, necessitated by the principle of duality, seems to deny the possibility of two lines being parallel. To solve this difficulty we resort to a trick discovered by the German astronomer Kepler at the beginning of the 17th century and rediscovered by Desargues within a few years.

The trick is to regard the plane of ordinary points and lines as being embedded in a new kind of plane which contains also *ideal* points on an *ideal* line, enabling us to say that two parallel lines meet in an ideal point and that all the lines parallel to a given line pass through the same ideal point. We postulate an ideal point for every direction, two opposite directions determining the same ideal point. Thus an ordinary point and an ideal point determine an ordinary line; but two ideal points determine the ideal line.

The ordinary plane is called the *affine* plane, the extended plane the *projective* plane, the ideal line the *line at infinity*, and the ideal points the *points at infinity*. Thus the projective plane may be described as the affine plane plus the line at infinity and all its points.*

In this spirit we define a parallelogram as a simple quadrangle whose diagonal points are both at infinity. Axiom VI asserts that the two diagonal lines cannot be parallel!

The problem of extending the affine plane is of the same nature as that of extending a limited region such as a rectangular sheet of paper. In carrying out a complicated construction we often find in practice that two lines whose intersection we seek are so nearly parallel that their point of intersection is inaccessible. It is possible to draw another line that would pass through this inaccessible point, by means of some auxiliary points and lines within the limited region. This kind of construction, applied to two parallel lines, is found to yield another line parallel to both.

However, this aspect of the projective plane obscures its essential symmetry: the fact that every point behaves just like every other point, and every line like every other line. It is therefore preferable to regard the points and lines of the projective plane as representing the lines and planes through a fixed point in ordinary three-dimensional space. With this interpretation, Axiom I simply states that two such lines determine a plane, and two such planes a line. Both statements evidently hold without exception.

5. *Desargues' Theorem and harmonic conjugacy.* From Axiom V we readily deduce

DESARGUES' THEOREM: *If two triangles have corresponding vertices joined by three concurrent lines, then the three intersections of corresponding sides are collinear.*
(See Fig. 1, where the joins of corresponding vertices of the two triangles *PQR* and *pqr* all pass through *S*, and consequently the intersections of corresponding sides all lie on *s*.)

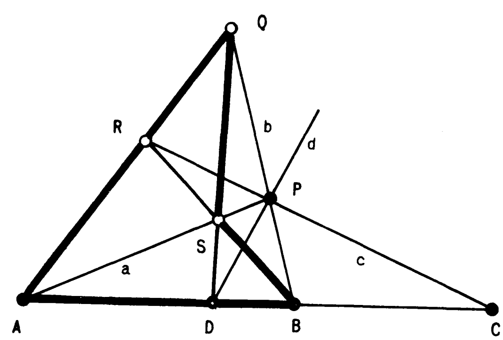


Figure 2

*For further details see Coxeter, *Non-Euclidean Geometry* (Toronto, 1947), pp. 159-178.

Four collinear points A, B, C, D are said to form a *harmonic set* if there is a simple quadrangle whose diagonal points are A and B while its diagonal lines pass through C and D respectively. We call C and D *harmonic conjugates* (of each other) wo* A and B, writing

$$\mathrm{H}(AB,\ CD)$$

as an abbreviated statement of the relation (see Fig. 2). It is a consequence of Desargues' Theorem that the harmonic conjugate of C wo A and B is independent of the choice of the construction-lines QA, QB, RC.

Dually, four concurrent lines a, b, c, d are said to form a harmonic set if there is a simple quadrangle whose diagonal lines are a and b while its diagonal points lie on c and d respectively, and we write

$$\mathrm{H}(ab,\ cd).$$

By considering the quadrangle $ABSQ$, we can identify a, b, c, d with the lines PA, PB, PC, PD of Fig. 2. Thus a harmonic set of points is projected from any point P by a harmonic set of lines, and (dually) any section of a harmonic set of lines is a harmonic set of points.

In defining the harmonic conjugate of C wo A and B, it is convenient to admit the special case when C coincides with B, so that P, S and D also coincide with B. This means that B (and likewise A) is its own harmonic conjugate wo A and B. But Axiom VI shows that in every other case the harmonic conjugates C and D are distinct.

6. *Perspectivity*. When we speak of the points on a line as forming a *range*, we think of them as the possible positions of a variable point X that 'runs along' the line. Dually, the lines of a *pencil* are the possible positions of a variable line x that 'rotates about' the center of the pencil. We often have occasion to consider a one-to-one *correspondence* between two ranges, or two pencils, or a range and a pencil. One trivial case, which must not be ignored, is when the two corresponding points (or lines) continually coincide; this correspondence is called the *identity*. The next simplest case is when the range is a section of the pencil, so that we have a correspondence $x \to X$ where X is the point in which x meets a fixed line.

The correspondence between two ranges that are different sections of one pencil is called a *perspectivity*. In such a case we write

$$X \;\overline{\wedge}\; X' \quad \text{or} \quad X \overset{O}{\overline{\wedge}} X',$$

meaning that the join XX' continually passes through a fixed point O.

As an instance of the use of this notation, consider the following sequence of three perspectivities (Fig. 3) which has the effect of interchanging pairs among any four collinear points $(AA')(BB')$:

$$AA'BB' \overset{Q}{\overline{\wedge}} RR'PB' \overset{A}{\overline{\wedge}} QQ'PB \overset{R}{\overline{\wedge}} A'AB'B.$$

*The preposition *wo* (rhyming with *so*) has been coined by some British mathematicians as a convenient abbreviation for 'with respect to' or 'with regard to.'

Since any perspectivity transforms a harmonic set into another harmonic set, it follows that a harmonic set remains harmonic after we have interchanged the four points in pairs. Interchanging the pairs (*AC*)(*BD*), we deduce that the relation H(*AB*, *CD*) implies H(*CD*, *AB*).

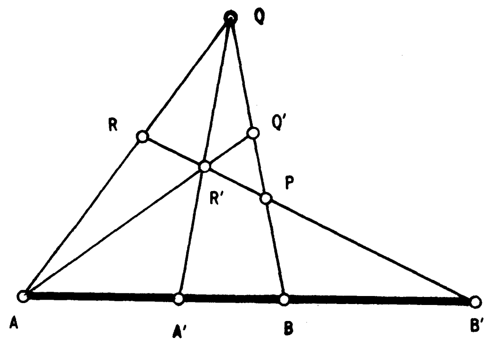


Figure 3

7. *Segment and interval.* We are now ready for Pieri's extraordinarily subtle definition of a segment, which is the basis for his theory of order. We are accustomed to thinking of a segment *AB* as the locus of points that lie 'between' *A* and *B*. For the purposes of projective geometry, however, this notion is inadequate: besides being connected by this 'minor' segment, *A* and *B* are also connected by a supplementary 'major' segment passing through the single point at infinity on the line *AB*. Since we have agreed to treat the point at infinity just like any other point on the line, the distinction between major and minor has no place in the strictly projective theory. To pick out one of the two segments *AB* we have to specify that a certain point does or does not lie in the chosen segment. Pieri uses the notation (*ACB*) for the segment that contains a point *C*. Of course this *C* can be replaced by any other point belonging to the segment. But without making any appeal to such intuituve ideas, he defines the segment as follows:

DEFINITION. *The segment (ACB) is the locus of the harmonic conjugate of C wo a variable pair of distinct points which are harmonic conjugates wo A and B.*

In other words, (*ACB*) is the locus of the point *X* given by

$$H(AB,\ MN), \qquad H(MN,\ CX).$$

Each position for *M* on the line *AB*, except *A* and *B* themselves, yields a distinct point *N* and a corresponding point *X*. If we omitted the word distinct' in the definition, we would have to allow *M*, *N* and *X* to coincide with *A* or with *B*: then instead of the segment (*ACB*) we would have the *interval* (*ACB*), which consists of the segment plus its end points *A* and *B*.

8. *The axioms of order.* A few properties of segments are immediate consequences of the above definition: *e.g.*, the segment (*ACB*)

is the same as (BCA); and if D is in (ACB), C is in (ADB). But we cannot go far without using some further assumptions, such as Pieri's three axioms of order:

AXIOM VII *If C, on AB, does not belong to the interval* $(\bar{A}C\bar{B})$, *it belongs to the segment (ABC).*

VIII *If D belongs to both (ABC) and (BAC), it cannot belong to (ACB).*

IX *If D belongs to (ACB), and E to (ADB), then E belongs to (ACB).*

Axiom IX shows that the segment (ACB), containing D, could equally well be called (ADB). Axiom VIII shows that an interval cannot cover the whole line. A point D that lies on the line without belonging to the interval $(\bar{A}C\bar{B})$ is said to be *separated* from C by A and B, and we write

$$AB \,//\, CD.$$

(Thus the sign // can be read 'separate'.) The fundamental properties of separation were enunciated by another Italian, Vailati, in 1895. For their deduction from Pieri's axioms (1899), and for the consequent theory of sense, the reader is referred to Chapter 3 of my book, *The Real Projective Plane* (1949),* pp. 33, 25. In the present outline we must be content with the assertion that all the ordinary properties of cyclic order can be rigorously deduced.

9. *The axiom of continuity.* In particular, a sequence of points $A_0, A_1, A_2, \ldots$ is said to be *monotonic* if $A_0A_n // A_1A_{n+1}$ for every integer $n > 1$; and a point M is called a *limit* of this sequence if it satisfies the following two conditions:

(1) for every integer $n > 2$, $A_1A_n // A_2M$;

(2) for every point P with $A_1P // A_2M$, there exists an integer n such that $A_1A_n // PM$.

In ordinary language, this means that the sequence is monotonic if each of the A's precedes the next as we go along the line without getting beyond A_0; and the requirements for a limit M are:

(1) The points $A_1, A_2, \ldots$ all precede M.

(2) Every point that precedes M precedes some A_n.

We are now ready for the final axiom:

AXIOM X *Every monotonic sequence of points has a limit.*

It follows easily that this limit is unique.

10. *Ordered correspondence.* A perspectivity transforms a harmonic set into another harmonic set, and consequently preserves any relation that is definable in terms of the harmonic relation. In particular,

*In subsequent footnotes this work will be cited as RPP. The McGraw-Hill Book Company have kindly granted me permission to reproduce some of the figures.

it preserves cyclic order: If $ABCD \barwedge A'B'C'D'$ and $AB // CD$, then $A'B' // C'D'$. The general correspondence having this property is called an *ordered* correspondence.

Any point M that coincides with its corresponding point M' is called an *invariant* point of the correspondence. (Some authors prefer the name *double* point.) For instance, a perspectivity between ranges on two lines has just one invariant point, where the two lines intersect. But the most interesting ordered correspondences arise when the two ranges are superposed: both on the same line. Such a correspondence is said to be *direct* or *opposite* according as it preserves or reverses *sense* (*i.e.*, the distinction between left-to-right and right-to-left along the line —— one of the concepts that can be defined in terms of separation*). A trivial but important instance of a direct correspondence is the identity, where every point on the line is invariant.

To get an intuitive picture of what is happening in a direct or opposite correspondence $X \to X'$, imagine two runners running all round a circular race track, starting at the same time and finishing at the same time, never stopping or turning back but otherwise free to go as fast or slow as they please. Then X and X' represent the respective positions of the two runners at any instant. The correspondence is direct or opposite according as the runners are going in the same direction or in opposite directions. An invariant point occurs where the runners meet or where one overtakes the other. In the direct case this may happen any number of times, even infinitely often, for the runners might remain side by side. Thus there may be any number of invariant points, from none at all to infinitely many. But in the opposite case a little thought reveals that the runners will meet exactly twice before each returns to his own starting point (or if they started from the same point they will meet once more elsewhere). This means that every opposite correspondence should have exactly two invariant points; but the rigorous proof of this fact requires considerable labor.**

11. *Projectivity.* Any correspondence that perserves the harmonic relation (and consequently preserves order) is called a *projectivity* and is denoted by

$$X \barwedge X'.$$

This means that, whenever the relation H(AB, CD) holds for four positions of X, the relation H($A'B'$, $C'D'$) holds for the four corresponding positions of X'. Thus a projectivity is a special case of an ordered correspondence, and a perspectivity is a special case of a projectivity.

Given three collinear points A, B, C, we can take the harmonic conjugate of one wo the other two, then the harmonic conjugate of this wo two of the previous points, and so on for ever. The resulting

*Coxeter, *Non-Euclidean Geometry*, p. 32.
**RPP, pp. 144 and 31.

harmonic net is an infinite set of points which has the appearance of covering the line completely (although actually it does not). To be more precise, we can prove the Lüroth-Zeuthen Theorem:*

Every segment contains a point of any given harmonic net.

If a projectivity has as many as three invariant points, it leaves invariant the whole harmonic net determined by those three, and consequently leaves every point invariant, *i.e.*, it must be the identity. (For, if not, there must be at least one point X having a distinct corresponding point X'. Let M and N be two points of the harmonic net, separating X and X'. Then the projectivity, although it is obviously direct, transforms the triad MXN into the triad $MX'N$ having the opposite sense, which is absurd.) This is the crucial step in the proof of

THE FUNDAMENTAL THEOREM OF PROJECTIVE GEOMETRY:** *A projectivity is determined when three points of one range and the corresponding three points of the other are given.*

This means that, if we are given three points A, B, C of one range and the corresponding three points A', B', C' of another, then every point X of the first determines a unique point X' of the second, such that

$$ABCX \barwedge A'B'C'X'.$$

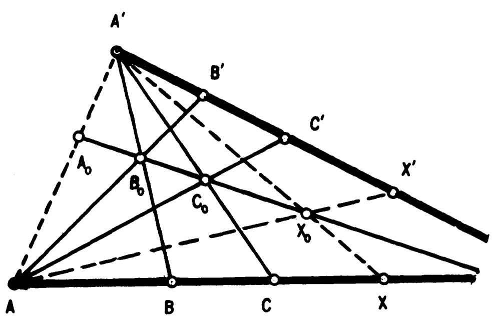


Figure 4

If the two ranges are on distinct lines, as in Fig. 4, the construction of X' is achieved by two perspectivities:

$$ABCX \overset{A'}{\barwedge} A_0B_0C_0X_0 \overset{A}{\barwedge} A'B'C'X'.$$

If instead they are both on the same line, we need a third perspectivity to obtain a corresponding range on another line before applying that simple construction. But we never need more than three perspectivities. Of course, the product of any number of perspectivities is still a projectivity; the remarkable fact is that every such

*RPP, pp. 140-142.

**RPP, p. 38.

product has the same effect as a suitable set of three, or possibly fewer. In particular, there is one important case where a single perspectivity suffices:

If a projectivity between ranges on two distinct lines has an invariant point, it is a perspectivity.

12. *The number of invariant points.* We have seen that a projectivity between ranges on one line cannot have more than two invariant points without being merely the identity. The projectivity is said to be *elliptic*, *parabolic*, or *hyperbolic* according as it has no invariant point, one invariant point, or two invariant points.

The actual construction for an elliptic projectivity requires the full allowance of three perspectivities. (An example of such a projectivety is

$$ABC \;\bar{\wedge}\; BCA,$$

where A, B, C are any three distinct collinear points.*) On the other hand, for a projectivity $MAB \;\bar{\wedge}\; MA'B'$ having a known invariant point M, we can use the simpler construction

$$MABX \overset{Q}{\;\bar{\wedge}\;} MA_0B_0X_0 \overset{R}{\;\bar{\wedge}\;} MA'B'X'$$

of Fig. 5, where A_0 and B_0 are any two points collinear with M. If there is a second invariant point, this must be collinear with $Q = AA_0 \cdot BB_0$ and $R = A'A_0 \cdot B'B_0$, so it can only be the point $QR \cdot AB$. Hence the projectivity $MAB \;\bar{\wedge}\; MA'B'$ is parabolic or hyperbolic according as the line QR does or does not pass through M.

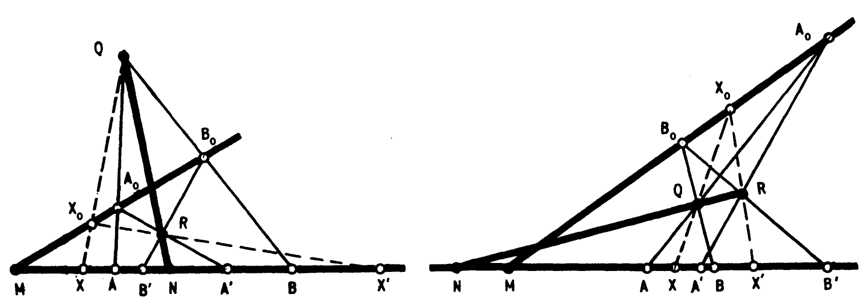


Figure 5

A hyperbolic projectivity may be either direct or opposite. For, if M and N are its two invariant points, the projectivity may be expressed as

$$MNA \;\bar{\wedge}\; MNA',$$

*RPP, p. 43.

and the sense determined by the three points MNA may or may not agree with the sense determined by MNA'. In the latter case $MN // AA'$. Hence,

The hyperbolic projectivity $MNA \barwedge MNA'$ *is opposite if* $MN // AA'$ *and direct otherwise.*

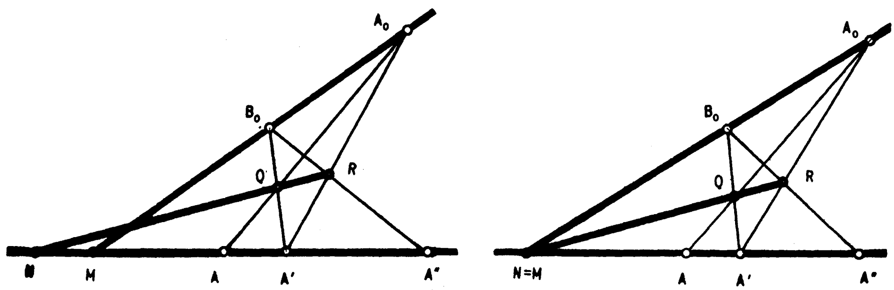


Figure 6

When the B of Fig. 5 coincides with A', we have the situation illustrated in Fig. 6, where

$$MNAA' \barwedge MNA'A''.$$

When M coincides with N, so that the projectivity is parabolic, we have H(MA', AA'') from the quadrangle A_0B_0QR. In this case $MA' // AA''$, so that the sense determined by MAA' agrees with the sense determined by $MA'A''$. Hence,

Every parabolic projectivity is direct.

13. *Involution.* We have considered a projectivity relating A to A', and A' to A''. It may happen that A'' coincides with A, so that the projectivity *interchanges* the two points A and A'. It is easily proved* that if this happens for one non-invariant point A it happens for every point: every X is interchanged with its corresponding X'. Such a projectivity is called an *involution.*

An **involution** may be elliptic or hyperbolic, but it cannot be parabolic; for, if it interchanges A, A' and has an invariant point M, it must have a second invariant point N, given by H(MN, AA'). Thus a hyperbolic involution is simply the correspondence between harmonic conjugates wo its two invariant points M and N.

The involution $AA'B \barwedge A'AB'$, being determined by the two pairs AA' and BB', is conveniently denoted by

$$(AA')(BB')$$

It follows from §7 that the segment (ABA') contains B' (as well as B) if and only if there exist two points M and N such that H(MN, AA') and H(MN, BB'). In other words, the involution $(AA')(BB')$ is hyper-

*RPP, p. 46.

bolic if and only if the pairs AA' and BB' do not separate each other. Therefore

$$(AA')(BB') \text{ is elliptic if and only if } AA' \,//\, BB'.$$

By expressing $(AA')(BB')$ in the form $AA'B \,\overline{\wedge}\, A'AB'$, we see that the involution is direct or opposite according as it is elliptic or hyperbolic.

Pappus discovered the following simple construction for the companion of any point X in a given involution $(AA')(BB')$. Draw a simple quadrangle in such a way that two opposite sides pass through A and A', and the other two through B and B', while one diagonal line passes through X; then the other diagonal line passes through X'.

When A coincides with A', and B with B', this reduces to the ordinary construction for the harmonic conjugate of X wo A and B. In this case the involution is naturally denoted by $(AA)(BB)$.

14. *Direct and opposite projectivities.* In such a brief account of this great subject, we naturally have to omit the proofs of most of the statements. But let us make an exception in favor of the following proof, because it illustrates the rich content of one-dimensional geometry; moreover, although the theorem is classical, the present method for dealing with it seems to be new.

. THEOREM. *Every elliptic projectivity is direct.*

We prove this by showing that every opposite projectivity is hyperbolic. We already know that every opposite *involution* is hyperbolic, so let us restrict consideration to an opposite projectivity that is *not* an involution. Let A be any point that is not invariant. Suppose the projectivity takes A to A', A' to A'', and A'' to A'''. Being opposite, it reverses sense; therefore the sense determined by $AA'A''$ disagrees with that determined by $A'A''A'''$, and $A'A'' \,//\, AA'''$. Changing the notation, we have a sequence of points

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots,$$

each transformed into the next (see Fig. 7); and the above separation holds for any consecutive four: points n and $n+1$ separate $n-1$ and $n+2$. Thus the points form four monotonic sequences (see §9): positive even numbers, positive odd numbers, negative even numbers, negative odd numbers. By our axiom of continuity, these sequences have limits, say M, M', N, N'. Since even points are transformed into odd points, and vice versa, the projectivity interchanges the limits M and M'. Since it is not an involution, this is impossible unless M and M' coincide. Similarly N and N' must coincide. Thus the projectivity has two invariant points M and N, and is hyperbolic as we wished to prove.

We have now classified the general projectivity into four categories: elliptic, parabolic, direct hyperbolic, and opposite. Involutions occur in the first and fourth categories.

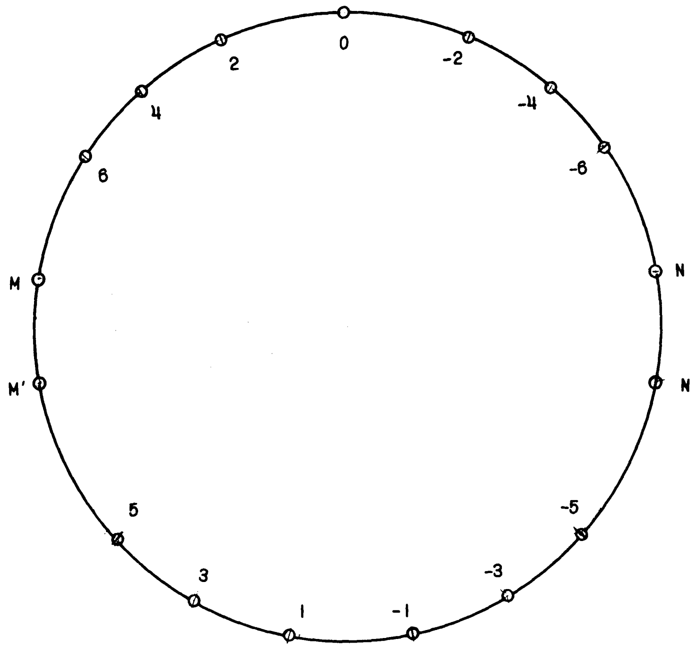


Figure 7

15. *Collineation and correlation.* The notion of correspondence extends easily from the line to the plane. By a two-dimensional correspondence $X \to X'$ we mean a rule for associating every point X with every point X' so that there is exactly one X' for each X and exactly one X for each X'. A correspondence $x \to x'$ for lines is defined similarly.

A *collineation* is the special case where a point and line are incident if and only if the corresponding point and line are incident, so that ranges correspond to ranges, and pencils to pencils. More precisely, the range of points X on a given line y corresponds to a range of points X' on the corresponding line y'. Moreover, four positions of X forming a harmonic set correspond to four positions of X' forming a harmonic set; for, any quadrangle used in constructing the first set corresponds to a quadrangle having the same relation to the second set. Thus a collineation induces a projectivity between ranges on corresponding lines, and (dually) a projectivity between pencils through corresponding points. It is easily proved* that, if a quadrangle is left invariant, the collineation can only be the identity (which leaves every point and every line invariant).

Instead of associating points with points and lines with lines, we can just as well associate points with lines and lines with points. A *correlation* is the special case where a point and line are incident

*RPP, p. 52.

if and only if the corresponding line and point are incident, so that ranges correspond to pencils, and pencils to ranges. Incidences are dualized: the range of points X on a given line y corresponds to a pencil of lines x' through the corresponding point Y'. Moreover, four positions of X forming a harmonic set of points correspond to four positions of x' forming a harmonic set of lines; for, any quadrangle used in constructing the first set corresponds to a quadrangle having the dual relation to the second set. Thus a correlation induces a projectivity between any range and the corresponding pencil.

The Fundamental Theorem (§11), belonging to one-dimensional geometry, has the following two-dimensional analogue:

A collineation or correlation is determined when a pair of corresponding quadrangles are given.

16. *Polarity.* In general, a correlation relates a point X to a line x', and relates this line to a new point X''. The correlation is 'involutory' if X'' always coincides with X, in which case we may omit the prime (') without causing any confusion. An involutory correlation is called a *polarity*. Thus a polarity relates X to x, and vice versa. Following Servois and Gergonne, we call X the *pole* of x, and x the *polar* of X. In virtue of the general properties of a correlation, the polars of all the points on a line a form a projectively related pencil of lines through the pole A.

Since a polarity dualizes incidences, if A lies on b, a passes through B. In this case we say that A and B are *conjugate points*, a and b are *conjugate lines*. If A and a are incident, A is a self-conjugate point and a a self-conjugate line. If two lines a and b are not conjugate, each point on a is conjugate to a definite point on b, and these points are related by a projectivity. In particular, the points on any non-self-conjugate line occur in conjugate pairs which belong to an involution. This remark enables us to prove*

HESSE'S THEOREM. *If the pairs of opposite vertices of a simple quadrangle are pairs of conjugate points, then the two diagonal points are likewise conjugate.*

Dually, if the pairs of opposite sides of a simple quadrangle are pairs of conjugate lines, then the two diagonal lines are likewise conjugate.

The figure of five points P, Q, R, S, T and five lines

$$p = RS, \quad q = ST, \quad r = TP, \quad s = PQ, \quad t = QR$$

is called a *pentagon*. The points P, Q, ... are called *vertices*, and the lines p, q, ... respectively *opposite sides*. It can be proved** that the correlation which relates four vertices of a pentagon to the respectively opposite sides relates the fifth vertex to the fifth side, and is a polarity. Thus any pentagon determines a unique polarity.

*RPP, p. 61.
**RPP, p. 64.

(But infinitely many other pentagons would serve equally well to determine the same polarity.)

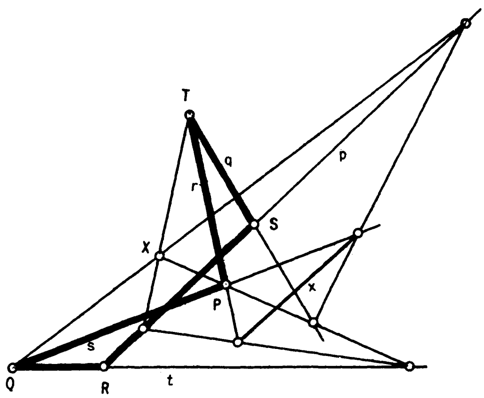


Figure 8

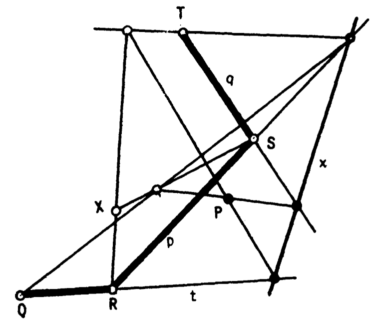


Figure 9

Let X be any point on neither of the sides r, s of a given self-polar pentagon $PQRST$. Then its polar is the line

$$[r\cdot(t\cdot PX)(p\cdot TX)][s\cdot(q\cdot PX)(p\cdot QX)]$$

(see Fig. 8). Dually, the pole of a line x through neither R nor S is the point

$$R[T(p\cdot x)\cdot P(t\cdot x)]\cdot S[Q(p\cdot x)\cdot P(q\cdot x)]$$

(see Fig. 9).

The pentagon $PQRST$ or $pqrst$ has five *diagonal points*

$$A = t\cdot q,\quad B = p\cdot r,\quad C = q\cdot s,\quad D = r\cdot t,\quad E = s\cdot p.$$

The involutions of conjugate points on the sides are evidently

$$(RB)(SE),\quad (SC)(TA),\quad (TD)(PB),\quad (PE)(QC),\quad (QA)(RD).$$

By examining the way in which these pairs of points separate one another, we see that, in the four cases shown in Fig. 10 (which are just the four projectively distinct types of pentagon), the involutions on the five sides fall into the following categories:

(1) all elliptic,
(2) two elliptic and three hyperbolic,
(3) one elliptic and four hyperbolic,
(4) all hyperbolic.

Each hyperbolic involution has two invariant points, which are self-conjugate. Hence the polarities determined by pentagons (2), (3), (4) all admit self-conjugate points. On the other hand, it is not surprising to find that the polarity determined by pentagon (1) admits *no* self-conjugate point. Such a polarity is said to be *elliptic*.* It plays an important role in elliptic geometry (see Chapter 9), where each point has a unique 'absolute polar.'

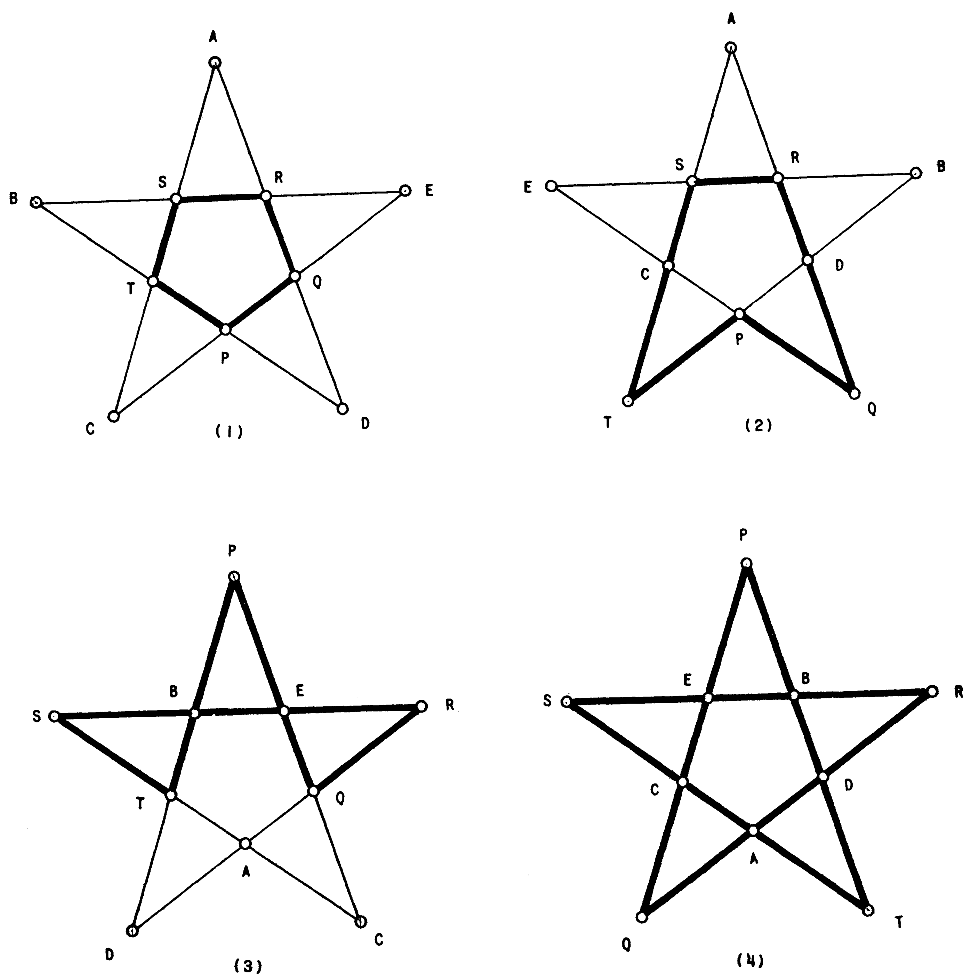


Figure 10

17. *The conic.* If a polarity admits one self-conjugate point M, it also admits a self-conjugate line m, the polar of M. Any other line through M contains an involution of conjugate points having M for one invariant point. The other invariant point of the involution provides another self-conjugate point on this arbitrary line through

*RPP, p. 70.

M. Conversely, any self-conjugate point N, distinct from M, is joined to M by a line which contains an involution $(MM)(NN)$ of conjugate points. Thus the presence of one self-conjugate point implies the presence of infinitely many. Their locus is called a conic, and their polars are called *tangents*. This definition exhibits the conic as a self-dual figure: the locus of self-conjugate points and also the envelope of self-conjugate lines.

A tangent justifies its name by meeting the conic only at its pole, the *point of contact*. Other lines are called *secants* or *exterior lines* according as they meet the conic twice or not at all, *i.e.*, according as the involutions of conjugate points on them are hyperbolic or elliptic. Any two conjugate points on a secant MN are harmonic conjugates wo M and N, since they are paired in the involution $(MM)(NN)$. It follows easily that, if a simple quadrangle is inscribed in a conic, the join of its two diagonal points is the polar of the intersection of its two diagonal lines.* Hence we can construct the polar of a given point X, not on the conic, by drawing through X any two secants, PR and QS; then the polar is

$$(PQ \cdot RS)(QR \cdot SP).$$

The pole of a given line AB can be constructed as the intersection of the polars of A and B; and the tangent at a given point M on the conic can be constructed by joining M to the pole of an arbitrary line through M.

The pole of an exterior line is an *interior point*; every line through such a point is a secant.** But the pole of a secant is an *exterior point*, from which two tangents can be drawn. Every point on an exterior line is an exterior point. It follows that two conjugate lines (each passing through the pole of the other) cannot both be exterior. Hence a self-polar pentagon cannot have two alternate sides that are exterior lines. This remark limits the possible number of exterior sides of a self-polar pentagon to 2, 1, or 0, in agreement with Fig. 10, cases (2), (3), (4).

To draw a self-polar pentagon for a given conic, take any two conjugate points T, Q, and let a line p meet the polars t, q in points R, S. The polars r, s of these new points intersect in P, the pole of RS, and we find that the pentagon $PQRST$ has the desired properties. Since each of the lines t, p, q may be either a secant or an exterior line (except that t and q must not both be exterior), we see that the same conic admits self-polar pentagons of all three types (2), (3), (4).

Considering the way the pentagons are actually drawn in Fig. 10, we see that the conic is a hyperbola in cases (2), (3), and a circle in case (4). But this distinction belongs to Euclidean geometry, not to projective geometry.

*RPP, p. 74.
**RPP, p. 72.

The following construction* for a conic through five given points was discovered independently by Braikenridge and Maclaurin about 1733. Let A, B, C, A', B' be the five given points, as in Fig. 11. Then the conic is the locus of the point

$$C' = A(z \cdot CA') \cdot B(z \cdot CB'),$$

where z is a variable line through the point $AB' \cdot BA'$.

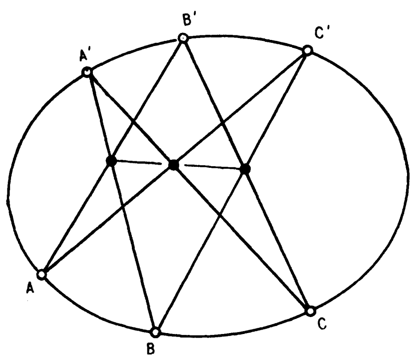


Figure 11

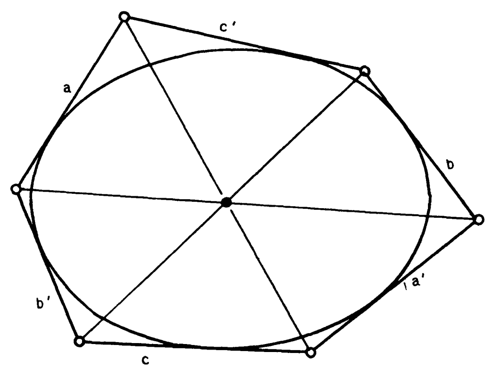


Figure 12

We may speak of $AB'CA'BC'$ as a hexagon inscribed in the conic. We observe that its three 'diagonal points'

$$BC' \cdot CB', \quad CA' \cdot AC', \quad AB' \cdot BA'$$

are collinear. This is Pascal's Theorem, discovered in 1639, when Pascal was only sixteen years old. It was dualized by Brianchon, more than 150 years later (see Fig. 12):

If a hexagon is circumscribed about a conic, the three diagonal lines are concurrent.

Five lines of general position have a natural cyclic order, namely the order in which they occur as tangents of the unique conic that touches them all. Taken in this order, they form a *convex* pentagon.** The other types of pentagon considered above are easily enumerated by examining the possible ways of upsetting this natural order.

18. *Various extensions of the theory.* For simplicity we have concentrated our attention on the real projective plane. But the

*RPP, p. 79.

**Cf. H. S. White, The plane figure of seven real lines, *Bulletin of the American Mathematical Society*, 38 (1932), p. 60, where it is proved that five lines decompose the plane into eleven regions: five triangles, five quadrangles, and one pentagon.

theory extends readily to the *complex* projective plane, which is obtained by modifying Axiom VII. This provides a greater degree of uniformity: every projectivity has one or two invariant points, every polarity determines a conic, and every line meets a conic.

Another important extension is to three-dimensional space. Instead of the lower half of Axiom I we have a more complicated statement, due to Pasch (1843-1913), which enables us to define a plane in terms of points and lines. Then we need the existence of a point outside a given plane, and another axiom saying that any two planes intersect in a line. (This has the effect of limiting the number of dimensions to three.) We now have a different principle of duality: points, lines and planes are related respectively to planes, lines and points. Two intersecting lines determine a plane ab and a point $a \cdot b$; these are dual concepts. The theory of polarities is analogous to the two-dimensional case, except that three-space admits the so-called *null* polarities, where *every* point lies on its polar plane. When we make the geometry complex instead of real, every ordinary polarity determines a quadric surface, but the null polarity still behaves quite differently.

Finally we must mention yet another line of development, considered by some to be the most significant of all. This is where *coordinates* are introduced, so that geometry is exposed to the powerful machinery of modern algebra.

Here are some suggestions for further reading. Of the following ten books, the first four (like the present outline) stress the synthetic aspect, the next three the analytic or algebraic, while the remaining three strike a balance between the two aspects and explain the relation between them:

L. CREMONA, *Elements of Projective Geometry*, Oxford, 1913.

G. B. MATHEWS, *Projective Geometry*, London, 1914.

C. W. O'HARA and D. R. WARD, *An Introduction to Projective Geometry*, Oxford, 1937.

A. N. WHITEHEAD, *The Axioms of Projective Geometry*, Cambridge (England), 1906.

W. C. GRAUSTEIN, *Introduction to Higher Geometry*, New York, 1930.

W. V. D. HODGE and D. PEDOE, *Methods of Algebraic Geometry (vol. 1)*, Cambridge (England), 1947.

J. A. TODD, *Projective and Analytical Geometry*, New York, 1947.

H. F. BAKER, *Principles of Geometry* (6 vols.), Cambridge (England), 1929-1933.

G. de B. ROBINSON, *The Foundations of Geometry*, Toronto, 1940.

O. VEBLEN and J. W. YOUNG, *Projective Geometry* (2 vols.), Boston,

University of Toronto

CURRENT PAPERS AND BOOKS

Edited by

H. V. Craig

This department will present comments on papers previously published in the MATHEMATICS MAGAZINE, lists of new books, and book reviews.

In order that errors may be corrected, results extended, and interesting aspects further illuminated, comments on published papers in all departments are invited.

Communications intended for this department should be sent in duplicate to H. V. Craig, Department of Applied Mathematics, University of Texas, Austin 12, Texas.

Cybernetics. By Norbert Wiener. New York, 1948. 194 pp. (Wiley)

Briefly stated the new science of cybernetics seeks to develop a theory which will embrace the common elements in the behavior of automatic machines and the human nervous system with particular emphasis on the aspects of communication and control. In this book the author presents the current status of this effort. It is a book your local bookseller knows well. It has been in the best seller class through its happy combination of provocative content, peculiarly attractive style, wide publicity, and message of interest to the entire scientific fraternity. It is in no sense a textbook or a purely scholarly treatise; witness the innumerable reviews that have appeared in non-scientific publications. It rather defies classification. The publishers were astonished by its success; at least five printings so far. Paradoxically the only comparable surprise seems to have been the parallel success of Al Capp's "Life and Times of the Schmoo".

Right in the introduction the author in his finest style sets about captivating his readers with a delightfully informal survey of the development of cybernetics from its beginnings over a decade ago up to the year 1947. Then comes a mildly philosophical chapter on "Newtonian and Bergsonian Time" where he traces the evolution of physical science from the reversible dynamics of classical astronomy with its emphasis on mechanism, through the statistical unidirectional mechanics of thermodynamics with its emphasis on energy, up to the present day emphasis on communication and control as, for example, in the design of servomechanisms. A chapter on "Groups and Statistical Mechanics" is for the most part a discussion of the modern form of Gibbs' ergodic theorem (Koopman, von Neumann, G. D. Birkhoff). This requires a preliminary discussion of groups, Lebesgue measure, limit in mean, etc. It is so brief and sophisticated that it may be meaningless to those not already familiar with the subject. However it should be said that it really puts the finger on key principles and can serve to orient an interested student or rusty-expert. There is a coda to this chapter on the Maxwell demon which is required reading for scientific philosophers. A long involved chapter on "Time Series, Information and Communication" is the most unblushingly mathematical and hence the most difficult in the book. It suffers from insufficient introductory material and references, an overpowering welter of Wieneresque formulae, and numerous inexcusable typographical errors. Incidentally these errors run through the entire book almost as if there had been no galleys but only in this and the two adjacent chapters are they really confusing. Of particular interest is a rather vaguely motivated mathematical definition of "amount of information", a quantity exhibiting the properties of negative entropy, and a proof of the second law of thermodynamics for communication engineering: "No operation on a message can gain information on the average". As a paper for a technical journal this chapter qualifies but it seems overly obscure and elaborate to serve well the purposes of the book as a whole. Perhaps it can be considered a foretaste of the author's very recent treatise on time series. "Feedback and Oscillation" is a short essay on servomechanisms which assumes the reader to be a fairly competent communications engineer. It concludes with a development of one of the main themes of the book, the servo-mechanical nature of neuro-physical reactions. The next three chapters, "Computing Machines and the Nervous System", "Gestalt and Universals", "Cybernetics and Psychopathology", reassume the descriptive character of the initial chapters. Here the relationship between computing machines (ENIAC type) and the physiology and even psychology of the nervous system is carried far. For example it is conjectured that under certain conditions the former might well exhibit conditioned reflexes. And more generally, mental and nervous disorders both functional and organic have their counterpart in possible computing machine behavior even to the extent of similar remedial treatment, shock treatment for example. Another interesting idea: how a process of scanning similar to that of television may be fundamental in the sensorial recognition of similarity and might be used to effect sensory prosthesis in disabled persons. Some somber thoughts develop. One is that the human brain has likely reached in size and complication a point of diminishing returns. Another is the forecast that the plebeian type of human brain power, perhaps up to and including minor executives, may in the future be replaced by generalized ENIAC's just as muscle power has already been superceded in great measure by power machinery. One ray of hope for mathematicians: the author thinks that such fine effective tools as these new machines will still need those with a high level of understanding and technical skill to take full advantage of their potentialities. The final chapter, "Information, Language, and Society" is short and sociologically meaty. The author puts into clear language the dilemma of social science, how it stands in a central position between macro- and micro-physics where the coupling between observer and observable is hardest to minimize and where statistics stands almost helpless because of the shortness of the statistical runs available as data. There is an appendix on the possibility of constructing a fairly bright chess automaton that is authentically automatic, not just a large container with a noted chess expert inside.

Certain mild shortcomings of the book have been duly noted above. As for concluding laudatory remarks, they are hardly necessary. In the months since publication the book has spoken pretty well for itself.

College Park, Maryland John L. Vanderslice

The Real Projective Plane. By H. S. M. Coxeter, McGraw-Hill Book Co., 10 plus 196 pages. $3.00.

At the present time the subject of Synthetic Projective Geometry is an area in which few mathematicians in this country are vitally interested. As a result, it is a pleasure to see a new work in this field which presents the material in a clear, rigorous, interesting manner. The book is clearly written, the printing is excellent, and the numerous diagrams are both clearly drawn and distinctly labeled.

The first seven chapters of this work deal with material customarily covered in a first course in Synthetic Projective Geometry. Enough of a postulational basis is given to show the need for and use of, such a treatment. Conics are introduced via self conjugate points of a hyperbolic polarity. Emphasis is placed on the concept of correspondence and its relation to the theory of transformations.

The latter chapters deal with the specialization of a projective geometry into an affine and a metric geometry, the concept of continuity, and an introduction into the study of coordinate systems, with especial emphasis on homogeneous projective coordinates.

There is enough material in this text for a three hour course in Projective Geometry over a year. For shorter courses, certain portions would have to be omitted, but the salient portions of the first seven chapters could be used for a first course (3 hours for one semester) in this subject.

There are numerous exercises in the text. For classes in Geometry which are mainly composed of students who are seeking a degree and are not too interested in mathematics, it is probable that additional exercises of an easier, more formal nature would have to be supplied by the teachers.

In terms of the mathematical background of students in the United States, this book would be suitable for a course available to Juniors, Seniors, and First Year Graduate Students in the average college. Technically, this book does not demand calculus as a prerequisite, but in general it would be desirable for the student to have mathematical maturity equivalent to that obtained by completing a course in the calculus before attempting to read the work. Mathematical maturity is especially demanded in the last three chapters, where the idea of continuity is discussed in detail (including some theorems on point sets and Dedekind's Axiom) and where geometric interpretations of addition and multiplication are introduced.

R. G. Sanger

Introduction to Applied Mathematics. By F. D. Murnaghan. New York, 1948. ix + 389 pp. (Wiley)

In all probability this book is the swan song of Professor Murnaghan as a resident of the northern hemisphere. But let us hope that many more such contributions will come from him below the equator even if in Portuguese. While his books have been mostly in the field of applied mathematics one is constantly impressed by the sure and uncompromising foundation of pure mathematics which he brings to them all. He never resembles either a purist out slumming or an engineer paying lip service to mathematical rigor. Continuing this tradition the present book is not just another primer for junior engineers. The general subject matter is that of a course for graduate students of mathematics and the sciences given for the last twenty years at Johns Hopkins. A bright senior with a good course in advanced calculus behind him might also study it with profit although the word "introduction" in the title is an understatement. Not too much emphasis is placed on actual physical problems; the main effort is to give the student a thorough grounding in some graduate mathematics which he will find useful in advanced scientific work. The reviewer, for one, thoroughly enjoyed the book and thinks he learned a lot of new things both mathematical and pedagogical.

To give a general idea of the content, here are the chapter headings: Vectors and Matrices, Linear Vector Functions, Function Vectors-Fourier Series, Curvilinear Coordinates, Laplace's Equation, Spherical Harmonics and Bessel Functions, Boundary Value Problems, Integral Equations, The Calculus of Variations, The Operational Calculus. It is not a catch-all list by any means and, except possibly for the last two, forms a close-knit structural whole. Apparently the author's purpose was to cover a limited number of related topics fully and with satisfying rigor. As is usual with Professor Murnaghan strikingly original methods of presentation appear in unexpected places. And his style seems somehow to inveigle the reader into taking his time and gaining a maximum of understanding.

Vectors we all know are a special pet of the author's which he believes everyone should use from cradle days. Their presentation here is of course excellent. From the beginning he uses components and matric methods and, having arrived at n-dimensions (via two and three), uses the complex field in order to prepare for the logical development of orthogonal functions and Fourier series as vector analysis in function space. Later, when he comes to introduce curvilinear coordinates, usually a disagreeable business, he finds a clever way to simplify the derivations by strong use of the invariant properties of the vector operations. The philosophy of tensor analysis permeates some of these early chapters although the word tensor is not mentioned in the book. Laplace's equation forms the background, first for a detailed treatment of electrostatic capacity bringing in ellipsoidal coordinates and the method of inversion along the way, and second for an admirably full and concise studv of spherical harmonics and Bessel functions. The subject of boundary value problems and associated Green's functions, always difficult to present, is set forth with great clarity yet without over-simplification. This leads logically to a discussion of integral equations in general, a chapter which almost forms a book in itself. Both the Fredholm and Hilbert-Schmidt theory are presented. Important proofs are not sidestepped, for example the one affirming that every Hermitian integral equation possesses at least one characteristic number. Rayleigh's principle and some vibration problems come in at the end.

In the chapter on the calculus of variations the emphasis is on dynamical applications: Hamiltonian systems, principle of least action, Liouville's theorem. The unusually fine chapter on operational calculus is rigorous, compact, and practical all at the same time. It is developed without benefit of the complex inversion integral. Instead emphasis is placed on the principle of superposition and on the Heaviside formula for the inversion of a transform expressed in series form, the latter requiring a six page proof. One virtue of this chapter is that the author when wrestling with an infinite double integral does not gloss over the difficulties. There is a full discussion of applications to systems of ordinary differential equations using matrix methods. Partial differential equations are omitted although several transforms important in their operational solution are treated.

The format of the book is excellent while typographical errors are few and trivial. The numerous sets of exercises, liberally sprinkled with hints, seem to have been selected with as much thought as went into the writing of the book proper. They encompass an amazing amount of new material and point the way to offshoots of the main line of development.

University of Maryland, College Park, Maryland. John L. Vanderslice

PROBLEMS AND QUESTIONS

Edited by

C. W. Trigg, Los Angeles City College

Readers of this department are invited to submit for solution problems believed to be new and subject-matter questions that may arise in study, in research, or in extra-academic situations. Proposals should be accompanied by solutions, when available, and by such information as will assist the editor. Ordinarily, problems in well-known text-books should not be submitted.

In order to facilitate their consideration, solutions should be submitted on separate, signed sheets within three months after publication of the problems.

All manuscripts should be typewritten on 8½" by 11" paper, double-spaced and with margins at least one inch wide. Figures should be drawn in india ink and in exact size for reproduction.

Send all communications for this department to *C. W. Trigg, Los Angeles City College, 855 N. Vermont Ave., Los Angeles 27, Calif.*

PROPOSALS

42. *Proposed by V. C. Throckmorton, Los Angeles City College.*

Encountering a man on the porch of his home, a census taker asked, "What are the ages of the persons living here?" The man replied, "My age is the sum of the ages of my wife, son and daughter. Each of our ages is a square number. My father's age is the sum of my age and the ages of my wife and daughter. Although he has passed the prime of life, his age is a prime number." What ages did the unstartled census taker record, and what obvious remark did he make about the wife's age?

43. *Proposed by Victor Thébault, Tennie, Sarthe, France.*

Show that the necessary and sufficient condition that the altitude AA', the median BB', and one of the bisectors of the angle C of a triangle ABC be concurrent is that $\sin A/\cos B = \pm \tan C$.

44. *Proposed by M. T. Goodrich, Keene Teachers College, Keene, N. H.*

Find all right triangles such that the sides are integers and the perimeter is numerically equal to the area.

45. *Proposed by J. M. Hurt, University of Texas.*

Two smooth wires are fitted, one along the (vertical) Y-axis and the other along the arc of the hypocycloid,

$$x^{2/3} + y^{2/3} = a^{2/3},$$

lying in the first quadrant. If two beads, one on each wire, are released at the same time from rest at the point (O,a), at what rate will the beads be separating when the bead on the straight wire reaches the origin? Assume gravity is the only force acting.

46. *Proposed by P. D. Thomas, U. S. Coast and Geodetic Survey,* *Washington, D. C.*

Integrate

$$I(x) = \int \frac{a \sinh^3 x + b \cosh^3 x}{c \sinh x + f \cosh x}\, dx$$

where a, b, c, f are constants.

47. *Proposed by N. A. Court, University of Oklahoma.*

Four parallel cevians AA', BB', CC', DD' of a tetrahedron $(T) = ABCD$ are divided by the four points P, Q, R, S internally in the same ratio, k. For what value of k will the four points P, Q, R, S be coplanar?

48. *Proposed by Howard Eves, Oregon State College.*

The centroid of any arc of a transition spiral coincides with the external center of similitude of the osculating circles of the extremities of the arc. (A transition spiral is a curve whose curvature varies directly with the arc length.)

SOLUTIONS

Late Solutions

34. O. F. Barcus, Temple University; L. A. Ringenberg, Eastern Illinois State College.

35. K. L. Cappel, San Francisco, Calif.

Spheres Associated with the Orthocentric Tetrahedron

1. [Sept. 1947] *Proposed by Victor Thébault, Tennie, Sarthe, France.*

In an orthocentric tetrahedron the spheres passing through three vertices and the feet of the corresponding altitudes have for radius the diameter of the first twelve-point sphere and intersect by threes on the second twelve-point sphere.

Solution by P. D. Thomas, U. S. Coast and Geodetic Survey, Washington, D. C. (Numbers in parentheses refer to articles in N. A. Court, *Modern Pure Solid Geometry*, Macmillan, 1935.) Any three vertices of an orthocentric tetrahedron and the orthocenters of the three respectively opposite faces are cospherical (217). The four spheres thus obtained clearly meet by threes in the orthocenters of the faces of the tetrahedron and are equal, since each radius is equal to one of the three equal bimedians of the orthocentric tetrahedron (210). But the three equal bimedians of an orthocentric tetrahedron are diameters of the first twelve-point sphere (797, 798) and the orthocenters of the faces of an orthocentric tetrahedron lie on its second twelve-point sphere (800, 801). Hence the proposition.

Permutations of Distinct Digits

5. [Sept. 1947] *Proposed by Victor Thébault, Tennie, Sarthe, France.*

Using once each of the digits 0,1,2,3,4,5,6,7,8,9 form a number which when increased by one million becomes a perfect square.

Editorial Note. Partial solutions of this problem appeared in THIS MAGAZINE, **21**, 232, (March 1948) and **22**, 49, (Sept. 1948). The 44 values of $M(=\sqrt{N+10^6})$, where N is a permutation of ten distinct digits, are:

35902	52966	56555	61217	70154	76411	86237	90287	96985
38836	53407	56728	61604	70541	76996	87661	91477	97055
40445	53801	58996	62198	75518	82099	87695	92402	97849
45433	54181	59183	64106	75529	85337	88354	95941	98182
50741	54584	59797	66266	76213	85805	88532	95966	

Furthermore, the 26 values of M, where N is a permutation of nine distinct digits, zero excluded, are:

12817	17279	18134	18863	22013	23768	24182	27269	28819
13958	17882	18181	18928	22787	23986	25282	27289	29024
13976	18116	18793	19441	23104	24029	26713	28781	

From these values various oddities may be extracted. For example, there are but two values of M which are palindromes, 66266 and 18181, one being in each group. Also, each group contains a unique M consisting of a permutation of five consecutive digits. Thus $(87695)^2 = 7689413025+10^6$ and $(23104)^2 = 532794816 + 10^6$. The ten digits occur once each in the latter two values of M taken together.

A Falling Chain

13. [Jan. 1948] *Proposed by J. S. Miller, Dillard University*

A uniform chain of length L and total mass M is held vertically with its lower end just touching a platform balance. The fixed upper end is released and the chain "accumulates" on the scale pan. What is the maximum reading of the balance?

Editorial Note. This problem is a special case of problem 17, page 252, J. H. Jeans, *Theoretical Mechanics*, Ginn, 1935. A solution to it appeared in *National Mathematics Magazine*, **19**, 320, (March 1945). There the maximum apparent weight (reading of the balance) was shown to be three times the actual weight of the chain. This occurs just as the last of the chain strikes the balance pan.

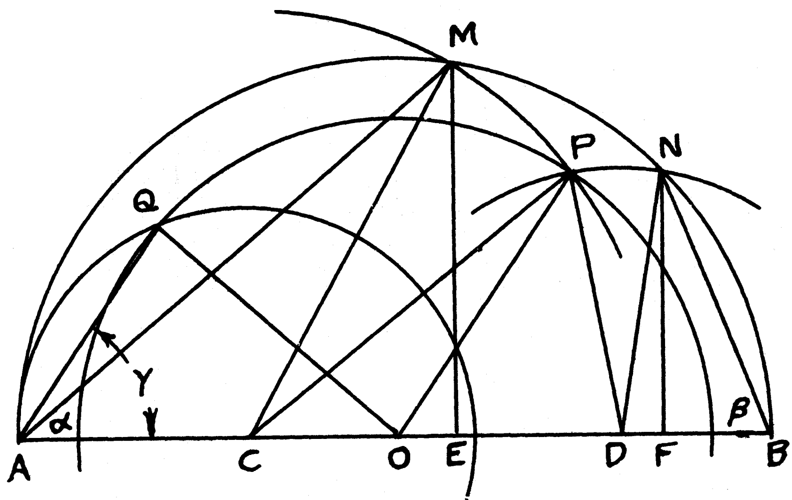
A Circular Locus

15. [Jan. 1948] *Proposed by Victor Thébault, Tennie, Sarthe, France.*

A rectangular segment EF of constant length slides on the diameter $AB = 2R$ of a semicircle (O). The perpendiculars to AB at E and F meet (O) at M and N. Show that the points of intersection P and P' of the circles passing respectively through M and N and having their centers at fixed points C and D on AB are on a circle concentric to (O), and that, if $CD = R$, we have the relation

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$$

between the angles α, β, γ made by AM, BN, AQ with AB, Q being the point where the circle of center C and radius CA intersects the locus of P.



Solution by Howard Eves, Oregon State College. By Stewart's theorem

$$\begin{aligned}(OP)^2(CD) &= (DP)^2(CO) + (CP)^2(OD) - (CD)(CO)(OD)\\ &= [(DF)^2 + (FN)^2](CO) + [(CE)^2 + (EM)^2](OD) - (CD)(CO)(OD)\\ &= [(OF-OD)^2 + R^2 - (OF)^2](CO) + [(CO-EO)^2 + R^2 - (EO)^2](OD)\\ &\quad - (CD)(CO)(OD)\\ &= [(OD)^2 + R^2 - 2(OF)(OD)](CO) + [(CO)^2 + R^2 - 2(CO)(EO)](OD)\\ &\quad - (CD)(CO)(OD)\\ &= [(OD)^2 + R^2](CO) + [(CO)^2 + R^2](OD) - 2(OD)(CO)(EF)\\ &\quad - (CD)(CO)(OD)\\ &= R^2(CD) - 2(OD)(CO)(EF). \qquad (1)\end{aligned}$$

Thus OP is independent of the position of segment EF on the diameter AB, and the first part of the problem is established.

Now $\cos^2\alpha = (AE)^2/(AM)^2 = AE/2R$, since $(AM)^2 = 2R(AE)$. Similarly, $\cos^2\beta = FB/2R$. Suppose $CD = R$. Then, from (1),

$$(OP)^2 = R^2 - 2(OD)(CO)(EF)/R = R^2 - 2(AC)(CO)(EF)/R. \qquad (2)$$

But, by the law of cosines,

$$(OP)^2 = (OQ)^2 = R^2 + (AQ)^2 - 2R(AQ)\cos\gamma \qquad (3)$$

Subtracting (3) from (2) we find

$$\begin{aligned}O &= 2R(AQ)\cos\gamma - (AQ)^2 - 2(AC)(CO)(EF)/R\\ &= 4R(AC)\cos^2\gamma - 4(AC)^2\cos^2\gamma - 2(AC)(CO)EF)/R,\end{aligned}$$

since $AQ = 2(AC)\cos\gamma$. Thus

$$\cos^2\gamma = 2(AC)(CO)(EF)/4R(AC)(R-AC) = EF/2R,$$

and finally we see that

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = (AE + FB + EF)/2R = 2R/2R = 1.$$

It may be observed that the second part of this problem furnishes a simple solution to the problem: Given two direction angles α and β of a line in space, to find, by euclidean tools, the third direction angle γ.

A Divisibility Property

17. [Sept. 1948] *Proposed by Leo Moser, University of Manitoba.*

Given an integer of n non-zero digits, show that it is always possible to replace a certain r $(0 \leq r < n)$ of these digits by zeros in such a way that the resulting number is divisible by n.

Solution by E. P. Starke, Rutgers University. Let the given integer be $k = a_{n-1}r^{n-1} + a_{n-2}r^{n-2} + \dots + a_1 r + a_0$, where r is the base of the system of numeration. Consider the n integers $k_1 = a_0$, $k_2 = a_1 r + a_0$, $\cdots$, $k_{n-1} = a_{n-2}r^{n-2} + \cdots + a_1 r + a_0$, $k_n = k$. If some k_j is divisible by n, obviously the problem is solved by replacing the first $n-j$ digits of k with zeros, thus obtaining k_j. If no k_j is divisible by n, then at least two of the k_j's must be congruent to each other, say $k_j \equiv k_i \pmod{n}$, $i < j$. Then the problem is solved by replacing the first $n-j$ digits and the last i digits of k with zeros. By hypothesis, the remaining number has at least one non-zero digit.

If the number of digits of k is less than n, the proposition need not be true. Consider the $r-2$ digit number $111 \cdots 111$ and take $n = r-1$. If any j digits $(0 \leq j < n-1)$ are replaced by zeros, the resulting number is congruent (mod n) to $n-j-1$ which cannot be zero.

Cubics with Integer Roots, Bend Points and Flex Points

28. [Jan. 1949] *Proposed by W. R. Talbot, Jefferson City, Missouri.*

Find the form of the roots of a cubic function for which the abscissas of the roots, bend points, and inflexion point are distinct integers.

Solution by Alan Wayne, Flushing, N. Y. Let $f(x) = (x-a)(x-b)(x-c)$ be the cubic function which has the desired properties, with a, b, and c distinct integers. The zeros of $f'(x)$ and $f''(x)$ are then also distinct. In $f(x)$, set $x = y + a$, so that $F(y) = y(y-h)(y-k)$, with $h = b-a$ and $k = c-a$. Then $F'(y) = 3y^2 - 2(h+k)y + hk$, and $F''(y) = 6y - 2(h+k)$. Moreover, the zeros of $f'(x)$ and $f''(x)$ are integers if and only if those of $F'(y)$ and $F''(y)$ are integers.

For integer zeros of $F''(y)$ we must have $(h+k)/3$ an integer. From this fact and the form of $F'(y)$ it follows that $hk/3$ must also be an integer. Hence $h = 3r$ and $k = 3s$, where r and s are integers. Thus,

$F'(y) = 3y^2 - 6(r+s)y + 9rs$ and has integer zeros if and only if $(r^2 - rs + s^2)^{1/2}$ is an integer; that is, if and only if $r = \pm(m^2 - n^2)$ and $s = \pm(2mn - n^2)$, where m and n are integer parameters. (Dickson, *History of the Theory of Numbers*, 1934, Vol. II, pages 405-406.)

The desired roots of $f(x) = 0$ are, therefore, of the form a, $a \pm 3(m^2 - n^2)$, and $a \pm 3(2mn - n^2)$. The abscissas of the bend points are $a \pm (2m^2 + mn - n^2)$ and $a \pm 3n(m-n)$, and the abscissa of the inflexion point is $a \pm (m^2 + 2mn - 2n^2)$. Here a, m, n, are any integers such that $m \neq 0, \pm n, 2n$ and $n \neq 0, 2m$. The restrictions upon m and n are necessary to insure distinctness of the five abscissas.

A cubic function with the desired properties and not too large coefficients is $x^3 + 3x^2 - 144x + 140$, obtained by taking the positive sign and $a = 1$, $m = 2$, $n = 3$.

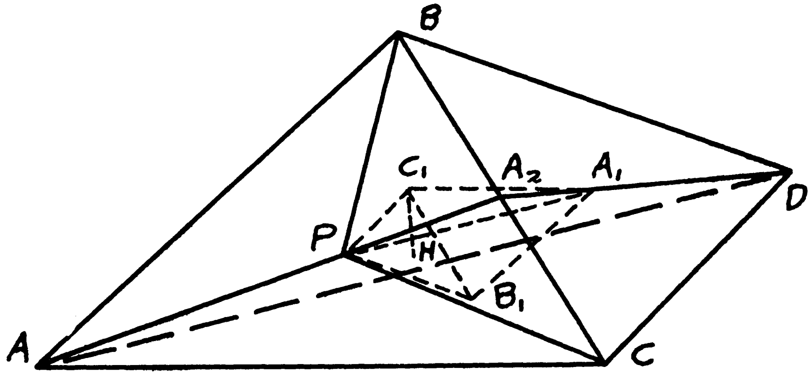
Two closely related problems occur in *American Mathematical Monthly*, **46**, 170, (1939) and **56**, 412, (1949).

Also solved by the proposer.

A Maximum Property of the Medial Tetrahedron

31. [Jan. 1949] *Proposed by Victor Thébault, Tennie, Sarthe, France.*

Through a point P inside the face ABC of a tetrahedron $ABCD$ draw parallels to the edges DA, DB, DC which meet the planes of the faces BCD, CDA, DAB in A_1, B_1, C_1, respectively. Determine the position of P so that the volume of the tetrahedron $PA_1B_1C_1$ will be a maximum.



Solution by L. M. Kelly, Michigan State College. Represent the areas of triangles ABC, $\cdots$ by $\overline{ABC}$, $\cdots$, respectively, and the volume of $PA_1B_1C_1$ by V. Let AP meet BC in A_2. Then from the similar triangles PA_1A_2 and ADA_2 we have $PA_1/PA_2 = DA/AA_2$, so $PA_1 = (PA_2/AA_2)(DA) = (\overline{PBC}/\overline{ABC})(DA)$. Similarly, it may be shown that $PB_1 = (\overline{PCA}/\overline{ABC})(DB)$ and $PC_1 = (\overline{PAB}/\overline{ABC})(DC)$.

Now $V = \frac{1}{3}(\overline{PA_1B_1})(CH) = \frac{1}{3}[\frac{1}{2}(PA_1)(PB_1)\sin A_1PB_1](PC_1)\sin C_1PH$ $= (\overline{PAB})(\overline{PBC})(\overline{PCA})(DA)(DB)(DC)\sin A_1PB_1 \sin C_1PH/6(\overline{ABC})^3$, where H is the foot of the altitude from C_1. Now as P varies in the face ABC, the angles A_1PB_1 and C_1PH remain constant, so V is a maximum when the product $(\overline{PAB})(\overline{PBC})(\overline{PCA})$ is a maximum. But the sum of these three areas is $\overline{ABC}$, a constant, and it is well known that if the sum of three positive numbers is fixed, their product is a maximum when the numbers are equal. Therefore $\overline{PAB} = \overline{PBC} = \overline{PCA}$ and P is the centroid of ABC. Furthermore, $PA_1B_1C_1$ is the medial tetrahedron of $ABCD$.

MATHEMATICAL MISCELLANY

Edited by

Charles K. Robbins

Let us know (briefly) of unusual and successful programs put on by your Mathematics Club, of new uses of mathematics, of famous problems solved, and so on. Brief letters concerning the MATHEMATICS MAGAZINE or concerning other "matters mathematical" will be welcome. Address: CHARLES K. ROBBINS, Department of Mathematics, Purdue University, Lafayette, Indiana.

On the Danger of Induction

In pointing out the danger of jumping to general conclusions from a few particular results the Euler polynomial $f(x) = x^2 + x + 41$ which is prime for $x = 1,2,\cdots,39$ (but composite for $x = 40$) is often cited. We will give here some other examples of false theorems which are true for the first few cases, and which seem somewhat less artificial than Euler's result.

Let $f(n)$ be the number of regions determined in the inside of a circle by joining n points on the circumference in all possible ways by straight lines, no three of which are concurrent inside the circle. It is easy to check that $f(n) = 1,2,4,8,16$, for $n = 1,2,3,4,5$; and it seems natural to suppose that $f(n) = 2^{n-1}$. However if we go to $n = 6$ we find that $f(6) = 31$. We leave it to the reader to show that the correct result for $n > 1$ is actually $f(n) = \sum_{j=0}^{n} \binom{n-1}{j}$

Our next example deals with prime representing functions. Consider $g(n) = 1 + \sum_{j=1}^{n} \phi(j)$, where $\phi(j)$ is Euler's totient function. For $n = 1,2,3,4,5,6$, this yields 2,3,5,7,11,13, which are the first 6 primes. This together with the fact that $\phi(n)$ is an arithmetic function connected with the theory of prime numbers and the fact that both $g(n)$ and p_n have only a single even element, makes the conjecture $g(n) = p_n$ plausible. However $g(7) = 19$ shatters the illusion. $g(8) = 23$ and $g(9) = 19$, so we might still suspect that $g(n)$ is always prime, but this too is false since $g(10) = 33$.

Finally there is the somewhat humorous experimental 'proof' that "Every odd number is a prime!" A prime is a number divisible only by 1 and itself. Certainly this is true for $n = 1$. We proceed step by step and in each case prove primality by dividing the odd number by all numbers less than it. In this way we find that 3,5, and 7, are primes. Apparently $9 = 3 \times 3$ but this would spoil the theory which has covered the facts perfectly so far, so we ascribe the discrepancy to experimental error and continue undaunted. We find the theory to hold for 11 and 13 and then we test a few odd numbers chosen at random like 23, 37, 41, $\cdots$. Thus the theorem has been proved.

University of North Carolina. Leo Moser

This month's letter:

"The enclosed check is to renew my subscription. I want to thank you for two of the things you have done with the Magazine.

First, I am glad you try to make the articles intelligible, and not a mere display of the learning of their authors.

Second, I am glad you are helping accustom the mathematical fraternity to the appearance of pages done by offset from typed material."

Very truly yours,

William R. Ranson.

Tufts College, Medford, Massachussetts.

(Received by Professor Glenn James, with permission to quote.)

Co-ed who is taking Algebra and who is being helped by her instructor outside of class: "Professor I need to know more about the biological theorem."

A Chart of Integral Right Triangles

When formulating problems in certain branches of elementary mathematics and physics, it is frequently desirable to make use of the properties of rational right triangles in order to minimize numerical work and to provide elegance in statement and solution. After a while the continual recrudescence of such old standbys as 3-4-5, 5-12-13, and 8-15-17 becomes monotonous, besides which more variety of shape may be required in order to fit imposed physical conditions.

The general solution of the *Pythagorean Problem*, namely the determination of right-angled triangles with integral sides, was known to Euclid, who gave rules[1] equivalent to this:
If m, n are integers, $(m > n)$ then

$$x = mn, \qquad y = \frac{m^2 - n^2}{2}, \qquad z = \frac{m^2 + n^2}{2} \tag{1}$$

obey the law

$$x^2 + y^2 = z^2.$$

When m and n are mutually prime odd numbers, x, y, z are relatively prime integers, and the triangle is called *primitive*. It is easily proved that all possible primitive triangles are given by equations (1). But as most readers will be aware, the orderly substitution of appropriate ascending values for m and n does not result in a series of

[1] Euclid's 'Elements', Book X, Prop. 29, Lemma 1.

[2] Even when the table is filled in by such a simple and elegant process as that communicated by Sir Flinders Petrie to Nature, 132, 411 (Sept. 9, 1933).

triangles having increasing or decreasing acute angles. Consequently, since in problem construction one usually starts the search with definite notions as to the *shape* of the required triangle, recourse to the above method may prove laborious[2] and unsatisfactory.

Being confronted from time to time with this situation, the writer prepared for his own use a chart of thirty primitive right triangles, drawn to scale, and set out in order of magnitude according to the smallest angle — in other words, arranged according to shape. The angles cover fairly well the interval from a little over 6° to a few minutes under 45°. This chart is reproduced in the accompanying figure for the benefit of others who may find such a tabulation useful. The angles, correct to the nearest minute, are given, and an index is provided so that when small numbers are more important than conformity to a certain shape, a suitable triangle may readily be selected. When the numbers are composite, their prime factors are indicated below. These will be found helpful when it is desired to combine two of the triangles into a single oblique one with a rational altitude (and hence a rational area) by 'matching legs' (the so-called *Heronian* case). A glance at the factors will suffice to tell whether the lowest common multiple of any likely-looking pair of sides is going to be uncomfortably large or not.

No novelty except that of convenience is claimed for the chart. It was prepared because, although very many more complete tables have been published[3], the writer is unaware of any easily accessible collection drawn up from this utilitarian point of view.

One bypath, however, may be of interest. It is frequently desirable, in mechanical applications especially, to obtain an integral right triangle, not given in the chart, with one acute angle falling between two given limits, say θ_1 and θ_2. We can approximate the angle as closely as we wish by noting that if θ be an angle in such a triangle we may write, from equations (1),

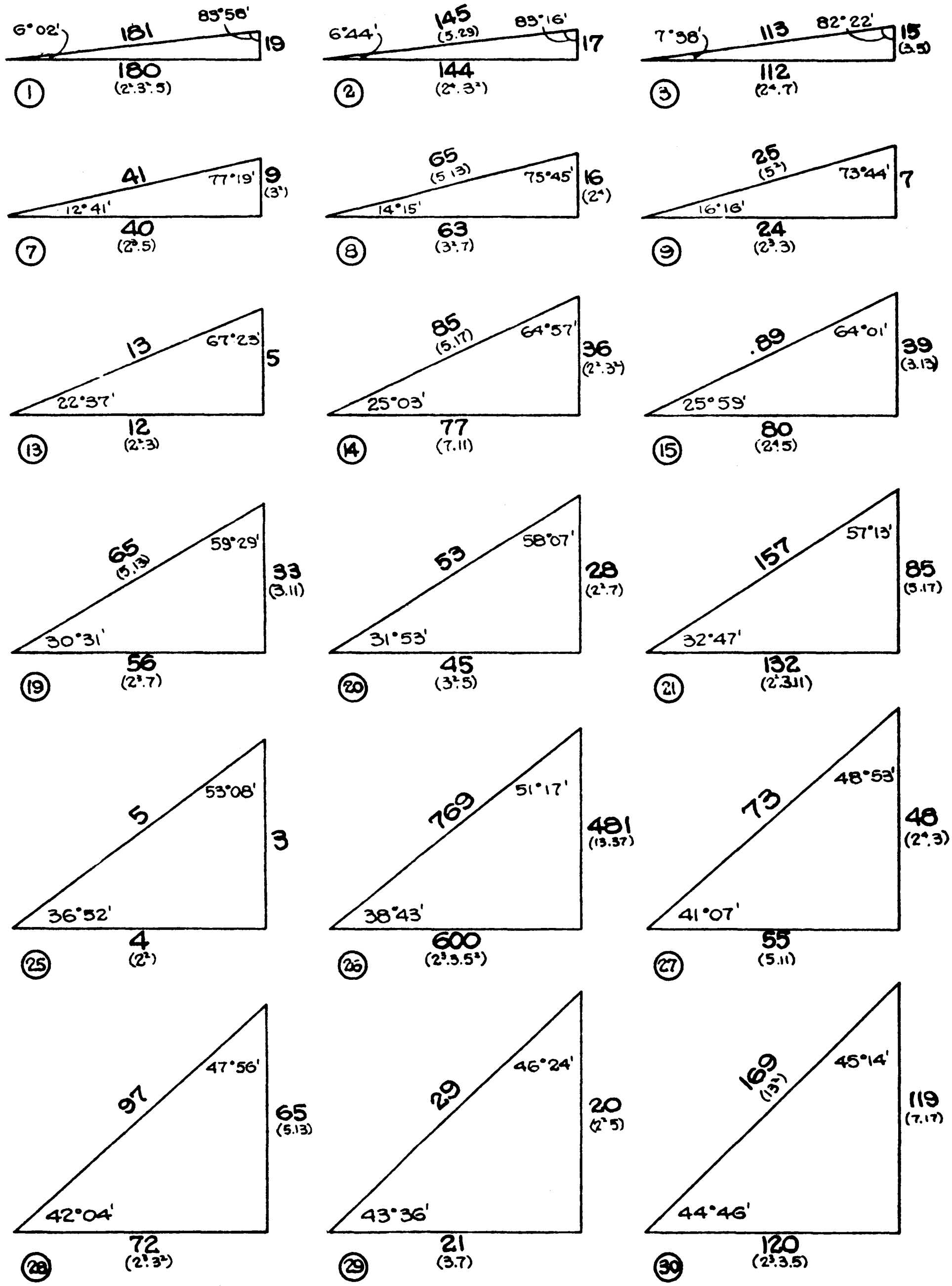
$$\tan\theta = \frac{x}{y} = \frac{2mn}{m^2 - n^2} = \frac{2\left(\frac{n}{m}\right)}{1 - \left(\frac{n}{m}\right)^2}$$

whence clearly

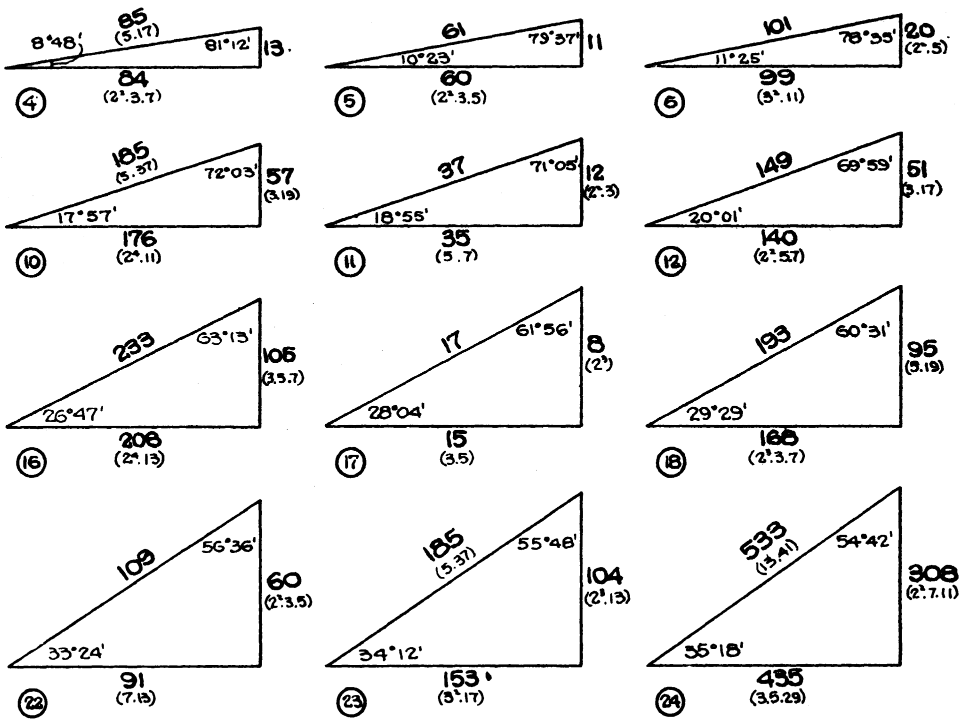
$$\tan\frac{\theta}{2} = \frac{n}{m}.$$

Thus the problem merely resolves into that of finding a rational

[3] Those interested in the topic will find a wealth of readable information, as well as copious references to the literature, in Professor L. E. Dickson's 'History of the Theory of Numbers' (Carnegie Institution of Washington Publications No. 256) Vol. 2, Chap. 4, pp. 165-190.



① 6°02' 181 83°58' 19 180 (2².3².5)
② 6°44' 145 (5.29) 83°16' 17 144 (2⁴.3²)
③ 7°38' 113 82°22' 15 (3.5) 112 (2⁴.7)
⑦ 41 77°19' 9 (3²) 12°41' 40 (2³.5)
⑧ 65 (5.13) 75°45' 16 (2⁴) 14°15' 63 (3².7)
⑨ 25 (5²) 73°44' 7 16°16' 24 (2³.3)
⑬ 13 67°23' 5 22°37' 12 (2².3)
⑭ 85 (5.17) 64°57' 36 (2².3²) 25°03' 77 (7.11)
⑮ 89 64°01' 39 (3.13) 25°59' 80 (2⁴.5)
⑲ 65 (5.13) 59°29' 33 (3.11) 30°31' 56 (2³.7)
⑳ 53 58°07' 28 (2².7) 31°53' 45 (3².5)
㉑ 157 57°13' 85 (5.17) 32°47' 132 (2².3.11)
㉕ 5 53°08' 3 36°52' 4 (2²)
㉖ 769 51°17' 481 (13.37) 38°43' 600 (2³.3.5²)
㉗ 73 48°53' 48 (2⁴.3) 41°07' 55 (5.11)
㉘ 97 47°56' 65 (5.13) 42°04' 72 (2³.3²)
㉙ 29 46°24' 20 (2².5) 43°36' 21 (3.7)
㉚ 169 (13²) 45°14' 119 (7.17) 44°46' 120 (2³.3.5)



Index by Greatest Side

Sides			No.	Sides			No.	Sides			No.
3	4	5	25	33	56	65	19	51	140	149	12
5	12	13	13	48	55	73	27	85	132	157	21
8	15	17	17	13	84	85	4	119	120	169	30
7	24	25	9	36	77	85	14	19	180	181	1
20	21	29	29	39	80	89	15	57	176	185	10
12	35	37	11	65	72	97	28	104	153	185	23
9	40	41	7	20	99	101	6	95	168	193	18
28	45	53	20	60	91	109	22	105	208	233	16
11	60	61	5	15	112	113	3	308	435	533	24
16	63	65	8	17	144	145	2	481	600	769	26

Conspectus of Integral Right Triangles

Of use in framing problems in Elementary Mensuration, Trigonometry, Analytical Geometry, Mechanics, etc.

NOTE: SMALL FIGURES IN BRACKETS ARE PRIME FACTORS fraction in its lowest terms, both numerator and denominator being odd, such that

$$\tan \frac{\theta_1}{2} < \frac{n}{m} < \tan \frac{\theta_2}{2} .$$

An ordinary straight slide-rule greatly facilitates the selection of m and n when, as is often the case, the limiting angles are not too close to one another[4]. The procedure may be illustrated by finding a triangle with θ between $\theta_1 = 21°30'$ and $\theta_2 = 22°00'$. Here $\frac{n}{m}$ must clearly lie between tan $10°45' = 0.190$ and tan $11°00' = 0.194$. Now if we set the index of the C-scale to some position intermediate between 190 and 194 on the D-scale, and then run an eye up the odd integers on the latter scale, we quickly come to 9 which is sufficiently close to 47 (on the C-scale) to make $0.190 < \frac{9}{47} < 0.194$. Since 9 and 47 are odd and relatively prime, we may now take $m = 47$, $n = 9$. Substitution in equations (1) gives the triangle 423-1064-1145. A check through, say, the tangent, shows that $\theta = 21°41'$, thus fulfilling the conditions imposed.

[4]For examples of closer approximation using continued fractions, see E. Sang, "On the Theory of Commensurables", Trans. Roy. Soc. Edin., 23, pp. 721-760, (1864).

McGill University
Montreal, Canada

W. Bruce Ross