Sums of Powers of Integers

A. F. Beardon

1. INTRODUCTION. Our starting point is the well-known identity

$$1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2. \tag{1.1}$$

Sums of the form

$$\sigma_k(n) = 1^k + 2^k + \cdots + n^k$$

have been studied for hundreds of years and even now there is still a steady stream of notes published on the subject, many of which can be found by browsing through back issues of the *Monthly* and the *Mathematical Gazette*. Most of these articles are concerned with generalizing (1.1), now written as

$$\sigma_3 = \sigma_1^2,$$

either by expressing a power of σ_1 as a linear combination of powers of other σ_i, for example,

$$\sigma_1^3 = \frac{1}{4}\sigma_3 + \frac{3}{4}\sigma_5, \qquad \sigma_1^5 = \frac{1}{16}\sigma_5 + \frac{5}{8}\sigma_7 + \frac{5}{16}\sigma_9, \tag{1.2}$$

or with identities involving σ_k and binomial coefficients, for example,

$$\sigma_2(n) = 2\binom{n+1}{3} + \binom{n+1}{2}, \quad \sigma_5(n) = \binom{n+1}{2} + 30\binom{n+2}{4} + 120\binom{n+3}{6},$$

or with showing that $\sigma_3^m = \sigma_1^{2m}$ is the only identity of the form

$$\sigma_{i_1} \dots \sigma_{i_r} = \sigma_{j_1} \dots \sigma_{j_s}.$$

For some of the history of the subject, and for a selection of these articles, we mention **[1]**, **[3]**, **[5]**, **[7]**, **[9]**, **[11]**, **[12]**, **[13]** and **[16]**, and especially **[6]**, **[8]** and **[10]**.

Here, we shall take a quite different approach and generalise (1.1) to the extent that we describe *all polynomial relations that exist between any two of the* σ_i. As (1.1) simply asserts that the points $(\sigma_1(n), \sigma_3(n))$, $n = 1, 2, \dots,$ lie on the parabola $y = x^2$, we are led naturally to (elementary) methods of algebraic geometry. The set of points (x, y) satisfying $T(x, y) = 0$, where T is a polynomial in two real variables, is a plane algebraic curve so that, writing

$$\Sigma_{ij} = \{(\sigma_i(n), \sigma_j(n)) : n = 1, 2, \dots\}, \tag{1.3}$$

the problem is to find all plane algebraic curves that contain the set Σ_{ij}.

It is well known that

$$\sigma_1 = \frac{n(n+1)}{2}, \qquad \sigma_2 = \frac{n(n+1)(2n+1)}{6}, \quad \sigma_3 = \frac{n^2(n+1)^2}{4},$$

and (by eliminating n from these) it is easy to find a polynomial relation between

σ_1 and σ_2, and between σ_2 and σ_3; these relations are

$$T(\sigma_1, \sigma_2) = 0, \qquad T(x, y) = 8x^3 + x^2 - 9y^2, \tag{1.4}$$

and

$$T(\sigma_2, \sigma_3) = 0, \qquad T(x, y) = 81x^4 - 18x^2y + y^2 - 64y^3, \tag{1.5}$$

respectively. Other, less obvious, relations are

$$T(\sigma_3, \sigma_5) = 0, \qquad T(x, y) = 16x^3 - x^2 - 6xy - 9y^2 \tag{1.6}$$

and

$$T(\sigma_2, \sigma_4) = 0, \qquad T(x, y) = 972x^5 - 7x^3 - 90x^2y - 375xy^2 - 500y^3. \tag{1.7}$$

The technique of eliminating a parameter enables us to prove much more than this, and we show

(1) for each pair of integers i and j with $1 \le i < j$, there is a unique irreducible polynomial T_{ij} in two variables, with integer coefficients, such that $T_{ij}(\sigma_i, \sigma_j) = 0$. By considering rings and ideals of polynomials, we can also show

(2) T_{ij} is the primitive relation between σ_i and σ_j in the sense that all other relations between these are trivial consequences of this one, and

(3) there is a (finite) algorithm for constructing any particular T_{ij}.

We remark in passing that we shall also see that there is no such result for polynomial relations among three or more of the σ_k.

To illustrate the results just described, consider the sums σ_1 and σ_3, and suppose that T is a real polynomial in two variables such that, for each $n, T(\sigma_1(n), \sigma_3(n)) = 0$. As the polynomial $T(t, t^2)$ vanishes when $t = n(n+1)/2$, $n = 1, 2, \dots,$ it is identically zero and so $T(x, y)$ vanishes on the parabola $y = x^2$. One can show that this forces T to have a factor $y - x^2$, whence any relation $T(\sigma_1, \sigma_3) = 0$ is a trivial consequence of the primitive relation $\sigma_3 = \sigma_1^2$.

This paper is written to be available to as wide a readership as possible. Some historical references are given (our earliest source dates back to 1615), but no attempt has been made to identify the original sources.

2. THE COEFFICIENTS OF σ_k. It is a fundamental fact that $\sigma_k(n)$ is a polynomial in n of degree $k+1$ and it is worthwhile to review Pascal's elementary proof of this (given in 1654). It is simply that

$$\begin{aligned}
(n+1)^{k+1} - 1 &= \sum_{m=1}^{n} \left[(m+1)^{k+1} - m^{k+1}\right] \\
&= \sum_{m=1}^{n} \sum_{r=0}^{k} \binom{k+1}{r} m^r \\
&= \sum_{r=0}^{k} \binom{k+1}{r} \sigma_r(n),
\end{aligned}$$

from which it follows (by induction) that $\sigma_k(n)$ is a polynomial in n of degree $k+1$. This means that we can now legitimately consider $\sigma_k(z)$ *as a polynomial in a complex variable* z and we shall soon see that $\sigma_k(0) = 0$.

The Bernoulli numbers B_n, $n \ge 0$, first appeared in the posthumous work *Ars Conjectandi* by Jakob Bernoulli in 1713, although they were known by Faulhaber much earlier than this (see **[6]** and Chapter 10 in **[8]**). They were introduced in order to provide an explicit formula for the coefficients of the polynomial σ_k, and are defined inductively by the recurrence relation

$$B_0 = 1, \quad \sum_{j=0}^{m} \binom{m+1}{j} B_j = 0 \tag{2.1}$$

([14], p. 229); we then have

$$1^k + 2^k + \cdots + (n-1)^k = \frac{1}{k+1} \sum_{j=0}^{k} \binom{k+1}{j} B_j n^{k+1-j} \tag{2.2}$$

([14], p. 234).

A calculation using (2.1) shows that

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0$$

(in fact, $B_3 = B_5 = B_7 = \cdots = 0$), so that, if $k \geq 3$, then

$$1^k + 2^k + \cdots + (n-1)^k = \frac{n^{k+1}}{k+1} - \frac{n^k}{2} + \frac{kn^{k-1}}{12} + O(n^{k-3}),$$

where here, $O(n^t)$ denotes a polynomial of degree at most t. Adding n^k to both sides, we obtain

$$\sigma_k(n) = \frac{n^{k+1}}{k+1} + \frac{n^k}{2} + \frac{kn^{k-1}}{12} + O(n^{k-3}); \tag{2.3}$$

note that *there is no term in* n^{k-2} (a consequence of $B_3 = 0$); this will be used later.

Of course, (2.2) shows that, for any complex number z,

$$\sigma_k(z) = \frac{1}{k+1} \sum_{j=0}^{k} \binom{k+1}{j} B_j (z+1)^{k+1-j}$$

which, with (2.1), yields

$$\sigma_k(0) = 0. \tag{2.4}$$

As an application of this, suppose that $T(\sigma_i, \sigma_j) = 0$ is any polynomial relation between σ_i and σ_j. Then, from (2.4),

$$0 = T\big(\sigma_i(0), \sigma_j(0)\big) = T(0,0);$$

thus *the constant term in* T *is zero*. Also, by putting $n = 1$, we see that *the sum of the coefficients of* T *is zero* (this is a useful check on our arithmetic).

3. FAULHABER POLYNOMIALS. It is well known that $\sigma_k(n)$ is a polynomial in n of degree $k+1$, but it is less well known that, when k is odd, σ_k is a polynomial in σ_1 of degree $\frac{1}{2}(k+1)$. The simplest case of this is $\sigma_3 = \sigma_1^2$, and the next two cases are

$$\sigma_5 = \sigma_1^2(4\sigma_1 - 1)/3, \quad \sigma_7 = \sigma_1^2\big(6\sigma_1^2 - 4\sigma_1 + 1\big)/3 \tag{3.1}$$

(proofs of these are given below). As these two formulae suggest, when k is odd, σ_1^2 divides σ_k. If k is even, then $\sigma_k(n)$ is of odd degree in n and so cannot be a polynomial in σ_1; however, in this case, σ_2 divides σ_k, and the quotient is a polynomial in σ_1. These results were known to Faulhaber and have been rediscovered many times since; even so, for the sake of completeness we give a formal statement and proof.

Theorem 3.1. *(i) For $k = 3, 5, \ldots$ there exists a polynomial F_k, of degree $\frac{1}{2}(k+1)$ and with a double zero at the origin, such that $\sigma_k = F_k(\sigma_1)$.*

(ii) For $k = 2, 4, \ldots$ there exists a polynomial F_k, of degree $\frac{1}{2}(k-2)$, such that $\sigma_k = \sigma_2 F_k(\sigma_1)$.

Proof: This is easy. Following the idea in Pascal's proof (namely, telescoping sums), we have

$$\begin{aligned}\sigma_1(n)^k &= \sum_{m=1}^{n}\left[\left(\frac{m(m+1)}{2}\right)^k - \left(\frac{(m-1)m}{2}\right)^k\right]\\ &= \sum_{m=1}^{n}\sum_{r=0}^{k}\binom{k}{r}\left(\frac{m}{2}\right)^k m^r\left[1-(-1)^{k-r}\right]\\ &= \frac{1}{2^k}\sum_{r=0}^{k}\binom{k}{r}\sigma_{k+r}(n)\left[1-(-1)^{k-r}\right]. \qquad (3.2)\end{aligned}$$

Now (3.2) holds for all k, but assume now that k is odd. Then the only terms in this sum that make a non-zero contribution are those with r even, and this show that σ_1^k is a linear combination of $\sigma_1, \sigma_3, \ldots, \sigma_{2k-1}$. As $\sigma_3 = \sigma_1^2$, this provides the basis of a proof of (i) by induction; we omit the details.

We can prove (ii) in a similar way, so suppose now that k is even, and note that

$$\begin{aligned}(2n+1)\sigma_1(n)^k &= \sum_{m=1}^{n}\left[(2m+1)\left(\frac{m(m+1)}{2}\right)^k - (2m-1)\left(\frac{(m-1)m}{2}\right)^k\right]\\ &= \sum_{m=1}^{n}\sum_{r=0}^{k}\binom{k}{r}\left(\frac{m}{2}\right)^k m^r\left[(2m+1)-(2m-1)(-1)^{k-r}\right]\\ &= \frac{1}{2^k}\sum_{r=0}^{k}\binom{k}{r}\Big(2\sigma_{k+r+1}(n)\left[1+(-1)^{k-r+1}\right]\\ &\qquad\qquad + \sigma_{k+r}\left[1+(-1)^{k-r}\right]\Big). \qquad (3.3)\end{aligned}$$

In the sum on the right, only the terms σ_q, with q even, survive, so that $(2n+1)\sigma_1(n)^k$ is a linear combination of $\sigma_2, \sigma_4, \ldots, \sigma_{2k}$. Again, this provides the basis of a proof of (ii) by induction.

Notice that taking k to be 3, and then 5, in (3.2), we obtain the two identities in (1.2). Next, eliminating σ_3 from (1.1) and (1.2), we obtain the first identity in (3.1). Finally, taking $k = 4$ in (3.2) gives

$$\sigma_1^4 = \frac{1}{2}\sigma_5 + \frac{1}{2}\sigma_7,$$

and eliminating σ_5 from this and the first identity in (3.1), we obtain the second identity in (3.1).

The first few of the polynomials F_k in Theorem 3.1 are

$$\begin{aligned}F_3(t) &= t^2,\\ F_4(t) &= (6t-1)/5,\\ F_5(t) &= t^2(4t-1)/3\\ F_6(t) &= (12t^2-6t+1)/7\\ F_7(t) &= t^2(6t^2-4t+1)/3.\end{aligned}$$

The formulae for F_3, F_5 and F_7 are restatements of (1.1) and (3.1); the formulae for F_4 and F_6 are obtained by putting $k = 2$ and $k = 3$ in (3.3). Theorem 3.1 was known to Faulhaber in 1615, and it is suggested in **[6]** that the polynomials F_j (strictly, a mild variant of these) are called the *Faulhaber polynomials*. For more details, see **[6]**, **[7]** and **[8]** but, briefly, these ideas date back to Johann Faulhaber (1631). Later contributions were made by Fermat (1636), Pascal (1654), Bernoulli (1713), Euler (1755) and Jacobi (1834).

Because of the polynomial relation (1.4) between σ_1 and σ_2^2, Theorem 3.1 has the following corollary.

Corollary 3.2. *(i) If k is odd, σ_k is a polynomial in σ_1;*
(ii) if k is even, σ_k^2 is a polynomial in σ_1.

The apparent lack of symmetry between the cases k odd and k even can be overcome by the substitution $y = x + 1/2$. Then

$$\sigma_1(x) = \frac{y^2 - 1/4}{2}, \qquad \sigma_2(x) = \frac{y(y^2 - 1/4)}{3},$$

and, more generally, we see from Theorem 3.1 that

(1) if k is odd then $\sigma_k(x)$ is an even function of $x + 1/2$;
(2) if k is even then $\sigma_k(x)$ is an odd function of $x + 1/2$.

Although we will have no use for the following formulae, we end this section by recording that the F_k can be defined by generating functions (see **[9]**):

$$\frac{\cosh\left[(x/2)\sqrt{1+8t}\right] - \cosh(x/2)}{2\sinh(x/2)} = \sum_{r=0}^{\infty} F_{2r+1}(t)\frac{x^{2r+1}}{(2r+1)!},$$

and

$$\frac{\sinh\left[(x/2)\sqrt{1+8t}\right]}{2\sqrt{1+8t}\sinh(x/2)} = \frac{1}{2} + \frac{1}{3}\sum_{r=1}^{\infty} tF_{2r}(t)\frac{x^{2r}}{(2r)!}.$$

4. THE EXISTENCE OF RELATIONS. Some algebraic curves are given parametrically by, say, $x = f(t)$ and $y = g(t)$, where f and g are polynomials, and the technique of eliminating t to obtain the polynomial relation between x and y is sometimes referred to as the theory of elimination (see, for example, **[2]** pp. 179–181 and **[4]**, Chapter 3). As each $\sigma_k(n)$ is a polynomial in n, we can apply this theory to obtain a polynomial relation between any two of the σ_i. Indeed, this was the way we produced the relations (1.4) and (1.5) between σ_1 and σ_2, and between σ_2 and σ_3. The relation (1.6) between σ_3 and σ_5 can be obtained by eliminating not n but σ_1. According to Theorem 3.1 and the explicit expressions for the F_k, we have

$$\frac{\sigma_5}{\sigma_3} = \frac{4\sigma_1 - 1}{3},$$

whence

$$16\sigma_3 = (4\sigma_1)^2 = \left(\frac{3\sigma_5 + \sigma_3}{\sigma_3}\right)^2$$

which yields (1.6). Likewise, the relation (1.7) can be found directly by eliminating

σ_1 from the identities

$$\frac{\sigma_4}{\sigma_2} = \left(\frac{6\sigma_1 - 1}{5}\right), \qquad \sigma_2^2 = \frac{\sigma_1^2(8\sigma_1 + 1)}{9}.$$

To obtain a general result, we need a general theory and it is this that we now describe. Suppose that the two polynomials

$$f(x) = a_0x^n + \cdots + a_{n-1}x + a_n, \qquad g(x) = b_0x^m + \cdots + b_{m-1}x + b_m$$

have a common zero, say x_0. Then each of the equations

$$f(x) = xf(x) = \cdots = x^{m-1}f(x) = 0 = g(x) = xg(x) = \cdots = x^{n-1}g(x)$$

is of the form $p(x) = 0$, where p is a polynomial of degree (at most) $m + n - 1$, and, as each of these equations is satisfied when $x = x_0$, the determinant of the coefficients must vanish. This $(m + n) \times (m + n)$ determinant is the *resultant* $R(f,g)$ of f and g and, explicitly,

$$R(f,g) = \begin{vmatrix} a_0 & \cdots & \cdots & a_n & & & \\ & \ddots & & & \ddots & & \\ & & a_0 & \cdots & \cdots & a_n \\ b_0 & \cdots & \cdots & b_m & & & \\ & \ddots & & & \ddots & & \\ & & b_0 & \cdots & \cdots & b_m \end{vmatrix},$$

where the omitted elements are zero, and the diagonal of $R(f, g)$ contains m occurrences of a_0 and n of b_m. For more details, see, for example, **[2]**, **[4]** and **[15]**, pp. 83–88. For emphasis, we repeat that the existence of a common zero of f and g implies that $R(f, g) = 0$.

Let us now illustrate this use of the resultant by verifying the relation (1.4) between σ_1 and σ_2. As

$$2\sigma_1(n) = n^2 + n, \qquad 6\sigma_2(n) = n(n + 1)(2n + 1) = 2n^3 + 3n^2 + n,$$

the polynomials

$$f(t) = t^2 + t - 2\sigma_1(n), \qquad g(t) = 2t^3 + 3t^2 + t - 6\sigma_2(n)$$

have a common zero, namely, $t = n$. We deduce that, for each n,

$$\begin{vmatrix} 1 & 1 & -2\sigma_1(n) & 0 & 0 \\ 0 & 1 & 1 & -2\sigma_1(n) & 0 \\ 0 & 0 & 1 & 1 & -2\sigma_1(n) \\ 2 & 3 & 1 & -6\sigma_2(n) & 0 \\ 0 & 2 & 3 & 1 & -6\sigma_2(n) \end{vmatrix} = 0$$

and this simplifies to give (1.4).

A similar argument holds for any pair σ_i and σ_j so there is at least one non-trivial polynomial P with $P(\sigma_i, \sigma_j) = 0$. Pascal's argument shows that the coefficients of n in the polynomial σ_k are rational numbers; hence, after clearing denominators, we may assume that the coefficients of P are integers. Even more is true, for this argument shows that P is of degree $j + 1$ in σ_i and $i + 1$ in σ_j, and (by considering the expansion of the determinant) we even see that the term involving σ_i^{j+1} is independent of σ_j, and likewise with i and j interchanged; for example, the determinant above is of the form

$$\begin{vmatrix} 2 & 3 \\ 0 & 2 \end{vmatrix}(-2\sigma_1)^3 + \cdots + \begin{vmatrix} 1 & 1 & -2\sigma_1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}(-6\sigma_2)^2.$$

We have now proved the following result.

Proposition 4.1. *Given integers i and j with $1 \leq i < j$, there is a polynomial $P(x, y)$, with integer coefficients and zero constant term, and of degree $j + 1$ in x and $i + 1$ in y with the leading terms in x and y having constant coefficients, such that $P(\sigma_i, \sigma_j) = 0$.*

This is not the end of the story, however. If we use this method in the case $i = 3$ and $j = 5$, we obtain a polynomial $P(x, y)$, of degree 6 in σ_3 and 4 in σ_5, such that $P(\sigma_3, \sigma_5) = 0$. It is tiresome (but instructive) to find this polynomial P explicitly, but in any event it is more complicated than the known relation (1.6) between σ_3 and σ_5 of *lower* degree. Clearly, the resultant obtained by eliminating n from the expressions $\sigma_3(n)$ and $\sigma_5(n)$ gives rise to a 10×10 determinant. If, on the other hand, we write σ_3 and σ_5 as polynomials in σ_1, and then eliminate σ_1, we obtain a 5×5 determinant which, after simplification, yields (1.6). The same reasoning applies to any pair σ_i and σ_j, where i and j are both odd integers, so that, in this case, it is better to eliminate σ_1 rather than n.

Now suppose that i and j are both even. We can write σ_i^2 and σ_j^2 as polynomials in σ_1 and then use the resultant to obtain a relation between σ_i^2 and σ_j^2 expressed as an $(i + j + 2) \times (i + j + 2)$ determinant which will be of degree $2(j + 1)$ in σ_i and $2(i + 1)$ in σ_j. If, on the other hand, we eliminate n between the expressions for σ_i and σ_j in terms of n, we obtain a determinant of the same size but with entries involving σ_i and σ_j (instead of σ_i^2 and σ_j^2). To illustrate this, observe that, using (2.2),

$$\sigma_2 = \frac{2n^3 + 3n^2 + n}{6}, \sigma_4 = \frac{6n^5 + 15n^4 + 10n^3 - n}{30},$$

whereas, using F_4 and (1.4),

$$\sigma_2^2 = \frac{8\sigma_1^3 + \sigma_1^2}{9}, \sigma_4^2 = \left(\frac{8\sigma_1^3 + \sigma_1^2}{9}\right)\left(\frac{6\sigma_1 - 1}{5}\right)^2.$$

In the case when i and j are both even, then, it is clearly better to eliminate the variable n rather than σ_1. We leave the reader to consider the case when i and j have opposite parity.

5. LÜROTH'S THEOREM AND THE RESULTANT. Consider again the case when i and j are odd integers. We can either use the resultant to eliminate n and so obtain a relation $R(\sigma_i, \sigma_j) = 0$ between σ_i and σ_j, or we can express both as a function of σ_1 and then use the resultant to eliminate σ_1 and so obtain a second relation $R^*(\sigma_i, \sigma_j) = 0$ between σ_i and σ_j. In this section we shall describe the precise relationship between the two relations $R(\sigma_i, \sigma_j)$ and $R^*(\sigma_i, \sigma_j)$. The material in this section is related to Lüroth's Theorem ([**15**], pp. 198–200). The interested reader can consult [**15**] for a precise statement of this, but it is not necessary for it suffices to give the geometric interpretation described in [**15**]. Suppose that an algebraic curve is parametrised by $x = f(t)$ and $y = g(t)$, where f and g are polynomials. If each point of the curve corresponds to, say, d values of

t, then Lüroth's Theorem guarantees that there is a polynomial ϕ of degree d, and polynomials f_1 and g_1, such that

$$x = f(t) = f_1(\phi(t)), y = g(t) = g_1(\phi(t)); \tag{5.1}$$

it follows that we can take $s = \phi(t)$ as a new variable and parametrise the curve by the polynomials $x = f_1(s), y = g_1(s)$ of lower degree. In these circumstances, we can find the equation of the algebraic curve either by eliminating t, or by eliminating s. The two equations (arising from the resultants) are denoted by $R(f, g)$ and $R(f_1, g_1)$, respectively, and the relation between these is given in the following result.

Theorem 5.1. *In the above notation, $R(f, g) = cR(f_1, g_1)^d$ for some constant c.*

In our earlier discussion, we have expressed σ_i both as a polynomial in n and as a polynomial in σ_1, and the above discussion applies with $d = 2$ and

$$\phi(t) = \sigma_1(t) = t(t+1)/2.$$

We deduce that if i and j are odd integers, the resultant obtained by eliminating n is simply a scalar multiple of the square of the resultant obtained by eliminating σ_1.

The proof of Theorem 5.1. We shall work with complex numbers (so that all polynomials factorise into linear factors). Suppose first that f and g are any complex polynomials, say

$$f(z) = a(z - z_1) \cdots (z - z_n), \qquad g(z) = b(z - w_1) \cdots (z - w_m).$$

As each coefficient of f is the product of a with a symmetric function of the roots z_j, and similarly for g, the resultant $R(f, g)$ is of the form $a^m b^n P(z_1, \ldots, z_n, w_1, \ldots, w_m)$ for some polynomial P. Thinking of the roots z_i and w_j as variables, we note that if $z_i - w_j = 0$, then f and g have a common root and so $R(f, g) = 0$. Continuing this line of argument (the details can be found in **[15]**, p. 86), we find that

$$R(f, g) = a^m b^n \prod_{i,j} (z_i - w_j).$$

We shall now apply these ideas to prove Theorem 5.1.

Let f, g, f_1, g_1 and ϕ now be the polynomials in (5.1) and suppose that these have degrees nd, md, n, m and d, respectively. We denote the zeros of f_1 by $\alpha_1, \ldots, \alpha_n$, and the solutions of $\phi(z) = \alpha_j$ by $x_{1j}, \ldots, x_{dj}$; then the zeros of f are precisely the numbers x_{ij}, where $i = 1, \ldots, d$ and $j = 1, \ldots, n$. Likewise, we denote the zeros of g_1 by $\beta_1, \ldots, \beta_m$, and the solutions of $\phi(z) = \beta_s$ by $y_{1s}, \ldots, y_{ds}$; then the zeros of g are the numbers y_{rs}. Finally, we use A_1, A_2, A_3, A_4 to denote constants (which we do not bother to evaluate).

According to the first paragraph in the proof,

$$R(f, g) = A_1 \prod_{i,j,r,s} (x_{ij} - y_{rs}), \quad R(f_1, g_1) = A_2 \prod_{j,s} (\alpha_j - \beta_s).$$

However, we also have

$$\phi(z) - \beta_s = A_3(z - y_{1s}) \cdots (z - y_{ds}) = A_3 \prod_r (z - y_{rs}),$$

so that, for each i, $1 \le i \le d$,

$$\alpha_j - \beta_s = \phi(x_{ij}) - \beta_s = A_3 \prod_r (x_{ij} - y_{rs}).$$

This holds for each i so, taking the product of both sides over $i = 1, \dots, d$, we obtain

$$(\alpha_j - \beta_s)^d = A_3^d \prod_{i,r} (x_{ij} - y_{rs})$$

from which the result follows immediately.

6. IDEALS OF POLYNOMIALS. In this section we characterise, for a given pair i and j satisfying $1 \le i < j$, the totality of polynomials T with integer coefficients for which $T(\sigma_i, \sigma_j) = 0$. To achieve this we borrow an idea from algebraic geometry and study the family of polynomials that vanish on a given set; for more details, we recommend [**4**]. We denote the integers by $\mathbf{Z}$, and the class (or ring) of polynomials with integer coefficients and in the two variables x and y by $\mathbf{Z}[x, y]$. The key fact that we need about $\mathbf{Z}[x, y]$ is that it is a unique factorisation domain; this means that any polynomial in $\mathbf{Z}[x, y]$ may be factored into a product of irreducible polynomials, and that up to order and factors of -1, this factorisation is unique (see [**2**], pp. 172–176).

Our declared aim is to study the family

$$\Gamma = \{T \in \mathbf{Z}[x, y]: T(\sigma_i, \sigma_j) = 0\}, \tag{6.1}$$

which we prefer to write as

$$\Gamma = \{T \in \mathbf{Z}[x, y]: T = 0 \text{ on } \Sigma_{ij}\},$$

where Σ_{ij} is given in (1.3) as the set of points $\{(\sigma_i(n), \sigma_j(n)) \in \mathbb{R}^2: n = 1, 2, \dots\}$.

For the moment, let E be any non-empty subset of $\mathbb{R}^2$, and define

$$\mathscr{I}(E) = \{T \in \mathbf{Z}[x, y]: T = 0 \text{ on } E\}.$$

The set $\mathscr{I}(E)$ is known in ring theory as an *ideal*, for it is closed under addition and the product $T_1 T_2$ is in $\mathscr{I}(E)$ whenever one of the T_i is. We want to investigate circumstances under which $\mathscr{I}(E)$ consists of all polynomial multiples of a single polynomial $T_0(x, y)$ (then $\mathscr{I}(E)$ is said to be the *principal ideal* generated by T_0). In general, this will not be so; it is not when, for example, E is the intersection of the two co-ordinate axes. However, we do have the following result.

Lemma 6.1. *Let f and g be non-constant polynomials with rational coefficients in one real variable, and let $E = \{(f(n), g(n)): n = 1, 2, \dots\}$. Then $\mathscr{I}(E)$ is generated by a non-trivial irreducible polynomial in $\mathbf{Z}[x, y]$.*

Proof: We have already seen that there is a polynomial P in two variables such that for all integers n, $P(f(n), g(n)) = 0$. This P can be obtained from the resultant of f and g (as in the proof of Proposition 4.1) so that if f and g have rational coefficients P may be taken to have integer coefficients. We claim that we may also assume that P is irreducible, for suppose that $P = P_1 \dots P_l$, where the P_j are the irreducible factors of P. Then there must be (at least) one factor P_j such that the polynomial $P_j(f(n), g(n))$ vanishes for infinitely many integer values of n. It follows that the polynomial $P_j(f(x), g(x))$ has infinitely many zeros and so vanishes for all x, and hence for all integers n. In conclusion, there is a irreducible polynomial P in $\mathbf{Z}[x, y]$ such that $P(f, g) = 0$.

We shall now show that P divides any polynomial T in $\mathbf{Z}[x, y]$ for which $T(f, g) = 0$. First, we express P and T as polynomials in y (whose coefficients are polynomials in x) and we then compute the resultant of P and T by eliminating the variable y. The resultant is a polynomial $R(x)$, and it is known that there are polynomials A and B in $\mathbf{Z}[x, y]$ such that

$$A(x, y)P(x, y) + B(x, y)T(x, y) = R(x),$$

where (as polynomials in y) $\deg(B) < \deg(P)$ and $\deg(A) < \deg(T)$ (see [**2**], Proposition 4.2.4, p. 179, and p. 192). As $P(f(n), g(n))$ and $T(f(n), g(n))$ vanish for each integer n, we see that $R(n) = 0$ for each integer n. It follows that R is the zero polynomial, and hence $P(x, y)$ divides $B(x, y)T(x, y)$. As P is irreducible, it divides B or T, and as $\deg(B) < \deg(P)$, it cannot divide B. It follows that P divides T, and that $\mathscr{I}(E)$ is the principal ideal generated by P.

The discussion to date yields our main result which follows.

Theorem 6.2. *Let i and j be integers with $1 \le i < j$. Then there is a non-constant irreducible polynomial T_{ij} in $\mathbf{Z}[x, y]$ such that $T_{ij}(\sigma_i, \sigma_j) = 0$. Further, T_{ij} divides P for any P in $\mathbf{Z}[x, y]$ for which $P(\sigma_i, \sigma_j) = 0$.*

Of course, this result says that any polynomial relation between σ_i and σ_j is a trivial consequence of the relation $T_{ij}(\sigma_i, \sigma_j) = 0$; for example, $T_{13}(x, y) = y - x^2$ and so if $P(\sigma_1, \sigma_3) = 0$, then P contains a factor $y - x^2$ and so the relation holds because of the existence of this factor. If P is any non-constant irreducible polynomial in $\mathbf{Z}[x, y]$, and if $P(\sigma_i, \sigma_j) = 0$, then $P = \pm T_{ij}$ and this observation enable us to identify T_{ij} in certain cases. For example, to show that

$$T_{23}(x, y) = 81x^4 - 18x^2y + y^2 - 64y^3, \qquad T_{35}(x, y) = 16x^3 - x^2 - 6xy - 9y^2,$$

we have, because of (1.5) and (1.6), only to prove that these polynomials are irreducible over $\mathbf{Z}$ and this can easily be done by assuming the contrary and reaching a contradiction. Finally, Faulhaber's observations mean that, for odd k, $T_{1k}(x, y)$ is an integer multiple of $y - F_k(x)$, whereas for even k, $T_{1k}(x, y)$ is an integer multiple of $y^2 - x^2(8x + 1)F_k(x)^2$.

Using the resultant we can eliminate n from $\sigma_i(n)$ and $\sigma_j(n)$ (in a finite number of steps) and obtain an explicit polynomial P in $\mathbf{Z}[x, y]$ for which $P(\sigma_i, \sigma_j) = 0$. This P must contain T_{ij} as a factor, and since the factorisation of P (over $\mathbf{Z}$) can also be carried out in a finite number of steps (see [**15**], p. 77), it follows that *each T_{ij} is computable in a finite number of steps*. There are, of course, other ways of finding the T_{ij}, for example, by using the Groebner basis method described in [**4**]. In the (few) examples I have tried with i and j odd, the polynomial relation between σ_i and σ_j obtained by eliminating σ_1 is irreducible over $\mathbf{Z}$, and hence is T_{ij}. I have been unable to prove that this is true for all odd i and j, but if it were to be true, it would provide a beautiful relationship between the ideal of polynomials annihilating (σ_i, σ_j) and Faulhaber's contribution 350 years ago.

7. SEPARABILITY. There is one feature of the relations between the σ_i that we have not yet commented on. We say that the relation $T_{ij}(\sigma_i, \sigma_j) = 0$, where $1 \le i < j$, is *separable* if it can be expressed in the form

$$P(\sigma_i) = Q(\sigma_j), \tag{7.1}$$

where P and Q are polynomials in one variable. Faulhaber's results show that every T_{ij} with $i = 1$ is separable, and all the examples we have of T_{ij} with $2 \le i < j$ are not separable. This suggests that T_{ij} is separable if and only if $i = 1$ and we shall now show that this is so. This means, of course, that the only cases in which T_{ij} is separable are those found by Faulhaber in the seventeenth century.

Theorem 7.1. *The relation $T_{ij}(\sigma_i, \sigma_j) = 0$ is separable if and only if $i = 1$.*

Proof: We assume that (7.1) holds with $2 \le i < j$ and we seek a contradiction. We write

$$P(x) = Ax^r + O(x^{r-1}), \qquad Q(x) = Bx^s + O(x^{s-1}), \tag{7.2}$$

where $AB \ne 0$, and, by comparing degrees of the two polynomials in (7.1), we have

$$(i + 1)r = (j + 1)s = N, \tag{7.3}$$

say. Observe that as $j \ge 3$, we have $N \ge 4$.

Now $Q(\sigma_j)$ is a polynomial of degree N in the variable n and, as

$$(j + 1)(s - 1) = N - (1 + j) \le N - 4,$$

we have, from (2.3) and (7.2),

$$\begin{aligned} Q(\sigma_j(n)) &= B\left[\left(\frac{n^{j+1}}{j+1} + \frac{n^j}{2} + \frac{jn^{j-1}}{12}\right) + O(n^{j-3})\right]^s + O(n^{N-4}) \\ &= B\left(\frac{n^{j+1}}{j+1} + \frac{n^j}{2} + \frac{jn^{j-1}}{12}\right)^s + O(n^{N-4}) \end{aligned}$$

because $(j + 1)(s - 1) + (j - 3) = N - 4$. The same holds for $P(\sigma_i)$; thus

$$A\left(\frac{n^{i+1}}{i+1} + \frac{n^i}{2} + \frac{in^{i-1}}{12}\right)^r = B\left(\frac{n^{j+1}}{j+1} + \frac{n^j}{2} + \frac{jn^{j-1}}{12}\right)^s + O(n^{N-4}).$$

Using (7.3), and equating the coefficients of x^N, we obtain

$$\frac{A}{(i+1)^r} = \frac{B}{(j+1)^s},$$

and this leads to

$$\begin{aligned} &\left(n^{i+1} + \frac{(i+1)n^i}{2} + \frac{i(i+1)n^{i-1}}{12}\right)^r \\ &\qquad = \left(n^{j+1} + \frac{(j+1)n^j}{2} + \frac{j(j+1)n^{j-1}}{12}\right)^s + O(n^{N-4}). \end{aligned}$$

Working now to an error term of order n^{N-3}, and using (7.3), this simplifies to

$$\begin{aligned} &24n^N + 12Nn^{N-1} + 2iNn^{N-2} + 3N(N - i - 1)n^{N-2} \\ &\qquad = 24n^N + 12Nn^{N-1} + 2jNn^{N-2} + 3N(N - j - 1)n^{N-2}. \end{aligned}$$

As this implies that $i = j$ (contrary to our assumption) Theorem 7.1 is proved.

8. CONCLUDING REMARKS. The reader may have noticed that much of the above discussion does not depend on the particular nature of the polynomials

$\sigma_k(n)$ as sums of powers of integers. Indeed, if

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n, \qquad g(x) = b_0 + b_1 x + \cdots + b_m x^m,$$

are any real polynomials, then the polynomials

$$F(t) = [a_0 - f(y)] + a_1 t + \cdots + a_n t^n, \quad G(t) = [b_0 - g(y)] + b_1 t + \cdots + b_m t^m,$$

have a common zero, namely $t = y$, and so $R(F, G) = 0$. This means, of course, that there is some polynomial $T(x, y)$ such that $T(f, g) = 0$.

It is amusing to apply these ideas to the Tchebychev polynomials T_n, which are defined by

$$T_n(\cos\theta) = \cos(n\theta).$$

As

$$T_n(T_m(\cos\theta)) = \cos(mn\theta) = T_m(T_n(\cos\theta)),$$

we see that $T_n(T_m(t)) = T_m(T_n(t))$ for all t; thus the polynomial relation connecting T_m and T_n is

$$T(T_m, T_n) = 0, \qquad T(x, y) = T_n(x) - T_m(y).$$

In this case, then, all of the primitive relations are separable.

As a final example, consider the Legendre polynomials ϕ_n defined by $\phi_0(x) = 1$ and

$$\phi_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}\left([1 - x^2]^n\right).$$

Now

$$\phi_2(x) = \frac{3x^2 - 1}{2}, \qquad \phi_3(x) = \frac{5x^3 - 3x}{2},$$

so that $4\phi_3(x)^2$ is polynomial in x^2 and $x^2 = (2\phi_2(x) + 1)/3$. This leads directly to the relation

$$T(\phi_2, \phi_3) = 0, \quad T(x, y) = 50x^3 - 15x^2 - 12x + 4 - 27y^2,$$

a result that can also be obtained by using the resultant.

REFERENCES

1. Anderson, I., Sums of squares and binomial coefficients, *Math. Gazette*, 65 (1981), 87–92.
2. Brieskorn, E. and Knörrer, H., *Plane algebraic curves*, Birkhäuser, 1986.
3. Colledge, T., *Pascal's triangle*, Stradbroke: Tarquin, 1992.
4. Cox, D., Little, J. and O'Shea, D., *Ideals, varieties and algorithms*, Springer-Verlag, 1992.
5. Edmonds, S. M., Sums of powers of the natural numbers, *Mathematical Gazette*, 41 (1957), 187–189.
6. Edwards, A. W. F., Sums of powers of integers: a little of the history, *Math. Gazette*, 66 (1982), 22–29.
7. Edwards, A. W. F., A quick route to sums of powers, *Amer. Math. Monthly*, 93 (1986), 451–455.
8. Edwards, A. W. F., *Pascal's arithmetical triangle*, Charles Griffin & Co., 1987.
9. Gessel, I., A formula for power sums, *Amer. Math. Monthly*, 95 (1988), 961–962.
10. Knuth, D. E., Johann Faulhaber and sums of powers, *Mathematics of Computation*, 61 (1993), 277–294.
11. Parker, R. V., Sums of powers of the integers, *Mathematical Gazette*, 42 (1958), 91–95.
12. Penney, D. E. and Pomerance, C., Multiplicative relations for sums of initial k-th powers, *Amer. Math. Monthly*, 92 (1985), 729–731.

13. Piza, P. A., Powers of sums and sums of powers, *Mathematics Magazine*, 25 (1952), 137–142.
14. Rose, H. E., *A course in number theory*, Oxford Univ. Press, 1989.
15. Van de Waerden, B. L., *Modern Algebra*, Volume 1, Ungar Pub. Co., 1949.
16. Woodall, D. R., Finite sums, matrices and induction, *Math. Gazette*, 65 (1981), 92–103.

Department of Pure Mathematics & Mathematical Statistics
University of Cambridge
16 Mill Lane
Cambridge CB2 1SB
ENGLAND
a.f.beardon@pmms.cam.ac.uk

PICTURE PUZZLE
(from the collection of Paul Halmos)

The picture was taken in 1980. Could you recognize him without the beard?
(see page 229)