Discrete Mathematics 94 (1991) 147–151
North-Holland

Note

On a paper of Agur, Fraenkel and Klein

Andrew Granville

Department of Mathematics, University of Toronto, Toronto, Ont., Canada M5S 1A1

Received 7 September 1988;
Revised 13 February 1989

Abstract

Granville, A., On a paper of Agur, Fraenkel and Klein, Discrete Mathematics 94 (1991) 147–151.

We count binary strings where the possible numbers of successive 0's and 1's are restricted.

For given sets A and B of positive integers define, for each $n \geqslant 1$, $S(A, B; n)$ to be the set of vectors $(x_1, x_2, \ldots, x_n)$ in $\{0, 1\}^n$ which do not contain a subvector $(x_j, x_{j+1}, \ldots, x_{j+c}, x_{j+c+1})$ of the form $(1, 0, 0, \ldots, 0, 0, 1)$ (with c zeros) for any $c \notin A$ or the form $(0, 1, 1, \ldots, 1, 1, 0)$ (with c ones) for any $c \notin B$ (here the indices of the x_i's are taken $(\bmod\, n)$). (A vector in $\{0, 1\}^n$ is called a 'binary string with n bits'). Let $\Psi(A, B; n)$ be the number of elements in $S(A, B; n) \setminus \{(0, 0, \ldots, 0), (1, 1, \ldots, 1)\}$. We prove the following.

Theorem. *For any given sets A and B of positive integers,*

$$\sum_{n \geqslant 1} \Psi(A, B; n) x^n = -x \frac{\mathrm{d}}{\mathrm{d}x} \log(1 - f(x) g(x))$$

where $f(x) = \sum_{a \in A} x^a$ and $g(x) = \sum_{b \in B} x^b$.

(N.B. f and g converge inside the unit disk, centred at the origin, and so, henceforth assume that $|x| < 1$. We call f the 'characteristic generating function' of the set A.)

In [1], Agur, Fraenkel and Klein considered the two examples $A = B = \{\text{integers } n \geqslant 2\}$ and $A = B = \{1, 2\}$ and came to an equivalent result by a different method (Equation (1) below gives this equivalence explicitly). The first example appears in connection with a model of information processing called 'majority rule'. They actually had a set of k integers $\{c_1, c_2, \ldots, c_k\}$ and k complex numbers $\{\gamma_1, \gamma_2, \ldots, \gamma_k\}$ such that $\Psi(A, B; n) = \sum_{i=1}^{k} c_i \gamma_i^n$ for each $n \geq 1$. We shall derive, from the Theorem, necessary and sufficient conditions for when such a result holds.

Corollary. *Let A, B, f and g be as in the Theorem. There exist integers $c_1, \ldots, c_k$ and complex numbers $\gamma_1, \ldots, \gamma_k$ such that $\Psi(A, B; n) = \sum_{i=1}^{k} c_i \gamma_i^n$ for each $n \geq 1$ if and only if $f(x)g(x)$ is a rational function.*

Remark. In the two examples above one has

$$f(x) = g(x) = x^2/(1-x) \quad \text{and} \quad f(x) = g(x) = x(1+x)$$

and so $f(x)g(x)$ is a rational function.

We now proceed to the following.

Proof of the Theorem. We will first consider strings in $S(A, B; n)$ which begin with 1 (i.e., $x_1 = 1$), and write them in the abbreviated form

$$1^t 0^{a_1} 1^{b_1} 0^{a_2} \cdots 1^{b_{m-1}} 0^{a_m} 1^u$$

which corresponds to the vector which starts with t ones, than a_1 zeros, b_1 ones, $\ldots$, a_m zeros and finally u ones. Such a string is counted by $\Psi(A, B; n)$ if and only if $m \geq 1$, $t \geq 1$, each $a_i \in A$ and $b_i \in B$ for $i = 1, 2, \ldots, m$ where $b_m = u + t \geq t$, and $\sum_{i=1}^{m} (a_i + b_i) = n$. Therefore the number of such strings is precisely the coefficient of x^n in

$$\begin{aligned} &\sum_{m \geq 1} \sum_{t \geq 1} x^t f(x) g(x) \cdots f(x) g(x) f(x) \sum_{u \geq 0, b = u+t \in B} x^u \\ &\quad = \sum_{m \geq 1} f(x)^m g(x)^{m-1} \sum_{t \geq 1} \sum_{b \geq t, b \in B} x^b \\ &\quad = \sum_{m \geq 1} f(x)^m g(x)^{m-1} \sum_{b \in B} b x^b \\ &\quad = \frac{f(x)}{1 - f(x)g(x)} \cdot x g'(x). \end{aligned}$$

By counting the strings in $S(A, B; n)$ that begin with a 0, in an analogous way, we get

$$\begin{aligned} \sum_{n \geq 0} \Psi(A, B; n) x^n &= \frac{x(f(x)g'(x) + f'(x)g(x))}{1 - f(x)g(x)} \\ &= -x \frac{\mathrm{d}}{\mathrm{d}x} \log(1 - f(x)g(x)). \qquad \square \end{aligned}$$

Proof of the Corollary. Sir Isaac Newton implicitly used the following identity in his work on symmetric polynomials: For any integers $c_1, c_2, \ldots, c_k$ and complex numbers $\gamma_1, \gamma_2, \ldots, \gamma_k$,

$$x\frac{\mathrm{d}}{\mathrm{d}x}\left\{\log\left[\prod_{i=1}^{k}(1-\gamma_i x)^{-c_i}\right]\right\} = \sum_{n\geq 1}\left[\sum_{i=1}^{k} c_i\gamma_i^n\right]x^n. \tag{1}$$

The corollary can be deduced immediately from comparing this identity to the Theorem, and then invoking the Fundamental Theorem of Calculus. □

When the characteristic generating function of a set A can be written as a rational function then one can deduce precise information about the structure of A.

Proposition. *Suppose that $f(x)$ is the characteristic generating function of the set A. Then $f(x)$ is a rational function if and only if A consists of the integers belonging to some finite union of arithmetic progressions with, at most, finitely many exceptions.*

I had hoped that a similar result might be deduced for a product of characteristic generating functions $f(x)g(x)$, so that if this were a rational function then the sets A and B might both be finite unions of arithmetic progressions with finitely many exceptions. This would have given a delightful conclusion to the Corollary! However, this conjecture is incorrect, as may be seen from the clever counterexample provided independently by Michael Albert and Neil Calkin, and by Paul Erdős:

Let A be the set of sums of even powers of 2 and let B be the set of sums of odd powers of 2 (include 0 in both sets). Now as any integer $n \geq 1$ can be written in a unique way as a sum of distinct powers of 2 so $\Psi(A, B; n) = 1$ and therefore $f(x)g(x) = 1/(1-x)$ is a rational function.

On the other hand, any integer n that lies in an interval of the form $[2^{2k-1}, 2^{2k})$ cannot belong to the set A, as 2^{2k-1} appears when we write n as a sum of distinct powers of 2. So, as these intervals grow to be arbitrarily large, A cannot contain all positive integers from some point onwards of *any* arithmetic progression.

Proof of the Proposition. Any finite union of arithmetic progressions may be rewritten as a finite union of *disjoint* arithmetic progressions, with a common modulus m (which is the least common multiple of the original moduli). (As an example, the union of 1 (mod 2) and 2 (mod 3) may be written as the union of 1, 2, 3, and 5 (mod 6).) Thus the characteristic generating function of such a set A is

$$\sum_{r\in R}\frac{x^r}{1-x^m}$$

where the set $R \subseteq \{0, 1, \ldots, m-1\}$ is composed of the least nonnegative integer in each of the arithmetic progressions. In order to add a finite set of integers S and to remove a finite set of integers T from A we need only add the polynomial $\sum_{s \in S} x^s - \sum_{t \in T} x^t$ to our generating function. Therefore if A is a finite union of arithmetic progressions except at most finitely many integers, then it has a characteristic generating function of the form $u(x)/(1-x^m)$ where $u(x)$ is some polynomial and m some positive integer.

On the other hand suppose that $f(x) = u(x)/v(x)$ where $u(x)$ and $v(x) = v_0 + v_1 x + \cdots + v_d x^d$ are polynomials without a common zero. Note that $v_0 \neq 0$ else $v(0) = 0$ and $u(0) = f(0)v(0) = 0$, implying that u and v do have a common zero. Let n_0 be the maximum of the degrees of $u(x)$ and $v(x)$. Let $p_a = 1$ if $a \in A$ and 0 otherwise, so that $f(x) = \sum_{i \geq 0} p_i x^i$. Also for any $n \geq n_0$ define the vector $c_n = (p_n, p_{n-1}, \ldots, p_{n-d})$.

Now, as the value of each p_i is either 0 or 1, we see that there are only finitely many distinct vectors c_n. Therefore, by the Pigeonhole Principle, we can find values k and $k+m$, with $m \geq 1$, $k \geq n_0 + 1$, for which $c_{k+m} = c_k$. We shall now prove that $c_{n+m} = c_n$ for each $n \geq k$, by induction on n: We are given the result for $n = k$ and so assume that $c_{n-1+m} = c_{n-1}$. Therefore $p_{n+m-i} = p_{n-i}$ for $i = 1, 2, \ldots, d$. Then by comparing the coefficients of x^n and x^{n+m} on both sides of the equation

$$v(x)f(x) = u(x),$$

we get

$$\sum_{i=0}^{d} v_i p_{n-i} = \sum_{i=0}^{d} v_i p_{n+m-i} = 0. \tag{2}$$

Thus

$$\begin{aligned} v_0 p_{n+m} &= -\sum_{i=1}^{d} v_i p_{n+m-i} && \text{by (2)} \\ &= -\sum_{i=1}^{d} v_i p_{n-i} && \text{by the induction hypothesis} \\ &= v_0 p_n && \text{by (2).} \end{aligned}$$

Then $p_{n+m} = p_n$ as $v_0 \neq 0$.

Finally, as $c_{n+m} = c_n$ for each $n \geq k$, so $p_{n+m} = p_n$ for each $n \geq k$ and so, if $a \geq k$ we see that $a \in A$ if and only if $a + m \in A$. The result follows immediately. □

At first sight it seems that the main difficulty in the above proof lies in showing that whenever the characteristic generating function of some set is the rational function $u(x)/v(x)$ then $v(x)$ divides $1 - x^m$ for some $m \geq 1$. Actually it is possible to generalize this (though with some difficulty) to the following result.

If $f_1(x), f_2(x), \ldots, f_k(x)$ are the characteristic generating functions of k sets of nonnegative integers, such that $f_1(x)f_2(x)\cdots f_k(x)$ is the rational function $u(x)/v(x)$ then $v(x)$ divides $(1-x^m)^k$ for some $m \geq 1$.

Unfortunately, as we saw from the above counterexample, this does not imply that each $f_i(x)$ takes the form $u_i(x)/(1-x^m)$.

Acknowledgements

I would like to thank Michael Albert, Neil Calkin and Paul Erdős for finding the counterexample mentioned herein.

References

[1] Z. Agur, A.S. Fraenkel and S.T. Klein, The number of fixed points of the majority rule, Discrete Math. 70 (1988) 295–302.