Distribution of the Figures 0 and 1 in the Various Orders of Binary Representations of kth Powers of Integers

By W. Gross and R. Vacca

Leibniz' observation (in *Mathematischen Schriften*, edited by C. I. Gerhardt, Halle, v. 7, 1863, p. 235) of the periodic repetition of the figures 0 and 1 in the columns of tables formed by writing successively the binary representations of the values taken by polynomials in x of arbitrary degree when x is given all the values of the positive integers, prompted R. Vacca to carry out counts of the figures 0 and 1 in the columns of tables formed by the kth powers of integers (with k an integer), which he did by means of appropriate programs for the FINAC electronic computer at Istituto Nazionale per le Applicazioni del Calcolo in Rome. Obviously in the order 2^0 the number of figures equal to 1 is equal to the number of figures equal to 0 for any k. It was also observed that for $h > 0$ the number of figures equal to 1 in the order 2^h is less than or equal to the number of figures equal to 0, that for increasing values of h the ratio between the number of figures equal to 1 and the total number of figures within a periodic sequence tends to the value $\frac{1}{2}$ and that the value $\frac{1}{2}$ is reached periodically, in the table of kth powers, every k orders, or in all columns 2^h for which h is divisible by k and for $k > 2$.

The "experimental" counts above referred to led to the formulation of the following theorem, the proof of which is due to W. Gross.

Theorem. *Let us consider a generic natural number n in its binary representation*

$$n = \sum_0^\infty \epsilon(h, n)2^h, \qquad \textit{with } \epsilon(h, n) = 0 \textit{ or } 1 .$$

The binary representation of the kth power of n, with k a positive integer, is

$$n^k = \sum_0^\infty \epsilon_k(h, n)2^h, \qquad \textit{with } \epsilon_k(h, n) = 0 \textit{ or } 1 .$$

We observe first that

$$\epsilon_k(h, n + 2^{h+1}) = \epsilon_k(h, n)$$

or, in other words, that $\epsilon_k(h, n)$ is periodic with period 2^{h+1} as a function of n, due to the fact that $n \equiv m \bmod 2^{h+1}$ implies that $n^k \equiv m^k \bmod 2^{h+1}$ ($\epsilon(h, n)$, in fact, depends only on the residue of $n \bmod 2^{h+1}$). We shall denote by $N_k(h)$ the number of $\epsilon_k(h, n)$ which are equal to 1 within a period, that is

$$N_k(h) = \sum_{i=0}^{2^{h+1}-1} \epsilon_k(h, i) .$$

The values of the ratio $N_k(h)/2^{h+1}$ (which is obviously equal to $\frac{1}{2}$ for $k = 1$) can be listed as follows:

Received June 21, 1967.

For $k = 2$

$$N_2(h)/2^{h+1} = \tfrac{1}{2} \qquad \text{for } h = 0\,,$$
$$N_2(h)/2^{h+1} = \tfrac{1}{2}\,(1 - 2^{-s}) \qquad \text{for } h > 0, \text{ with } s = [h/2]\,.$$

For $k > 2$

$$N_k(h)/2^{h+1} = \tfrac{1}{2} \qquad \text{if } k \text{ is a divisor of } h\,,$$
$$N_k(h)/2^{h+1} = \tfrac{1}{2}\,(1 - 2^{-s}) \qquad \text{if } k \text{ is not a divisor of } h\,,$$

where

$$s = [(h + k)/k] \qquad \text{if } \mu = 0\,,$$
$$s = [(h + k - \mu - 2)/k] \qquad \text{if } \mu \neq 0\,,$$

having denoted by μ *the maximum exponent such that* 2^μ *is a divisor of* k. *Obviously the statement that* $\mu = 0$ *implies that* k *is odd. We have denoted by* $[x]$ *the integral part of* x.

Proof. Let us begin by introducing a function similar to $N_k(h)$, but in which the sum is only extended to odd numbers:

$$\nu_k(h) = \sum_{m=0}^{2^h-1} \epsilon_k(h, 2m + 1) \tag{1}$$

and let us express $N_k(h)$ in terms of $\nu_k(h)$, We observe, in this context, that any number i included in the interval $0 \leqq i \leqq 2^{h+1} - 1$ may be written in the unique form

$$i = 2^r(2m + 1)\,,$$

with $0 \leqq m \leqq 2^{h-r} - 1;\ 0 \leqq r \leqq h$, so that the sum $N_k(h)$ may be written in the form

$$N_k(h) = \sum_{r=0}^{h} \sum_{m=0}^{2^{h-r}-1} \epsilon_k(h, 2^r(2m + 1))\,. \tag{2}$$

We observe now that

$$i^k = 2^{rk}(2m + 1)^k$$

from which

$$\epsilon_k(h, 2^r(2m + 1)) = 0 \qquad \text{for } rk > h\,,$$

whereas

$$\epsilon_k(h, 2^r(2m + 1)) = \epsilon_k(h - rk, (2m + 1)) \qquad \text{for } rk \leqq h\,.$$

We may write therefore

$$N_k(h) = \sum_{r=0}^{[h/k]} \sum_{m=0}^{2^{h-r}-1} \epsilon_k(h - rk, 2m + 1)\,, \tag{3}$$

while, in virtue of definition (1), we have

$$\nu_k(h - rk) = \sum_{m=0}^{2^{h-rk}-1} \epsilon_k(h - rk, 2m + 1)\,. \tag{4}$$

Due to the periodicity of ϵ_k with respect to n, the internal sum of formula (3) has the value

$$\sum_{m=0}^{2^{h-r}-1} \epsilon_k(h - rk, 2m + 1) = 2^{r(k-1)}\nu_k(h - rk) . \tag{5}$$

Substituting the value given by (5) into (3) we obtain

$$N_k(h) = \sum_{r=0}^{[h/k]} 2^{r(k-1)}\nu_k(h - rk) . \tag{6}$$

The problem is, therefore, reduced to the computation of $\nu_k(h)$.

Let us observe now that $\nu_k(0) = 1$, so that in what follows we shall limit ourselves to the consideration of cases in which $h > 0$. Consider first the case of k odd and let us observe that, if x takes all the values of the odd numbers between 1 and $2^{h+1} - 1$, then x^k takes the same values mod 2^{h+1}. This is due to the fact that for x and y both odd

$$x^k \equiv y^k \qquad \text{mod } 2^{h+1}$$

if and only if

$$x \equiv y \qquad \text{mod } 2^{h+1}$$

which appears immediately obvious considering that

$$x^k - y^k = (x - y) \sum_{s=0}^{k-1} x^s y^{k-1-s}$$

and that the summation on the right contains an odd number of odd terms and is, therefore, an odd number, which proves the assertion.

Remembering that the $\epsilon(h, n)$ depend only on the residue of n mod 2^{h+1} and based on the observation above, we have that

$$\sum_{m=0}^{2^h-1} \epsilon_k(h, 2m + 1) = \sum_{m=0}^{2^h-1} \epsilon(h, 2m + 1)$$

or that

$$\nu_k(h) = \nu_1(h) , \tag{7}$$

and, as obviously

$$\nu_1(h) = 2^{h-1} ,$$

the final result for k odd is

$$\nu_k(h) = 2^{h-1} .$$

Take now $k = 2^\mu\rho$ with ρ odd (where μ is the number introduced in the statement of the theorem). We have $x^k = (x^\rho)^{2^\mu}$ and x^ρ for the reasons stated above takes mod 2^{h+1} all the values of the odd numbers between 1 and $2^{h+1} - 1$ while x varies between the same bounds, so that in this case we have

$$\sum_{m=0}^{2^h-1} \epsilon_k(h, 2m + 1) = \sum_{m=0}^{2^h-1} \epsilon_{2^\mu}(h, 2m + 1)$$

or

$$\nu_k(h) = \nu_{2^\mu}(h) .$$

Formula (7) is a particular case, for $\mu = 0$, of the above formula. We have, therefore, reduced the problem to the computation of ν_k, where k is a power of 2.

A well-known theorem of the theory of numbers states that for x odd we have

$$x^{2^\mu} \equiv 1 \quad \bmod 2^{\mu+2} \tag{8}$$

which means that 2^μ powers of an odd number in the binary representation contain $(\mu + 1)$ zeros on the left of the terminal 1. This entails that for $1 \leqq h \leqq \mu + 1$ we have $\nu_{2^\mu}(h) = 0$.

Let us proceed now to the case in which $h \geqq \mu + 2$. Observe that the number of odd numbers between 1 and $2^{h+1} - 1$ which satisfy (8) is $2^{h-1-\mu}$ and that half of them obviously has the value $\epsilon(h, n) = 1$.

If we prove, therefore, that when x takes the values of the mentioned odd numbers x^k takes each value exactly $2^{\mu+1}$ times, we will have shown that $\nu_k(h) = 2^{h-1}$.

In other words it is sufficient to prove that, for z odd, the congruence

$$x^{2^\mu} \equiv z^{2^\mu} \quad \bmod 2^{h+1} \tag{9}$$

has exactly $2^{\mu+1}$ solutions.

We shall use a well-known representation theorem which states what follows. Any odd number may be represented mod 2^{h+1} in the unique form

$$x \equiv (-1)^\alpha 5^\beta \quad \bmod 2^{h+1}$$

where α takes the values 0 and 1 and β takes those of a complete system of residues mod 2^{h-1}.

Write, then, in virtue of this representation

$$x \equiv (-1)^\alpha 5^\beta \quad \bmod 2^{h+1} ; \qquad z \equiv (-1)^{\alpha'} 5^{\beta'} \quad \bmod 2^{h+1} .$$

Substituting in (9) we have

$$5^{2^\mu \beta} \equiv 5^{2^\mu \beta'} \quad \bmod 2^{h+1}$$

and, because the representation is unique, this relationship is equivalent to

$$2^\mu \beta \equiv 2^\mu \beta' \quad \bmod 2^{h-1} . \tag{10}$$

The solutions of (10), as indicated by the general theory of congruences, coincide with those of

$$\beta \equiv \beta' \quad \bmod 2^{h-\mu-1} . \tag{11}$$

Formula (10) is satisfied therefore by the 2^μ values of β which satisfy (11)

$$\beta \equiv \beta' + k2^{h-\mu-1} \quad \bmod 2^{h-1} \qquad \text{with } k = 0, 1, \cdots, 2^\mu - 1 .$$

This number is doubled if we take into account the fact that α can take two values.

We have proven, therefore, that

$$\nu_{2^\mu}(h) = 2^{h-1} \qquad \text{for } h \geqq \mu + 2 .$$

Finally we have therefore

$$\begin{aligned}
&\nu_k(h) = 1 && \text{for } h = 0\,, \\
(12)\qquad &\nu_k(h) = 2^{h-1} && \text{for } k \text{ odd and } h > 0 \text{ and for } k = 2^\mu\rho \text{ and } h > \mu + 1\,, \\
&\nu_k(h) = 0 && \text{for } k = 2^\mu\rho \text{ and } 0 < h \leqq \mu + 1\,.
\end{aligned}$$

In order to compute $N_k(h)$ it is now sufficient to use (6) taking into account (12). Obviously we have $N_k(0) = 1$ and again we shall consider only cases for which $h > 0$.

Let us now consider the various cases.

(a) *k odd; k is not a divisor of h.* We have

$$N_k(h) = \sum_{r=0}^{[h/k]} 2^{r(k-1)} 2^{h-rk-1} = \sum_{r=0}^{[h/k]} 2^{h-1-r} = 2^h(1 - 2^{-[(h+k)/k]})\,.$$

(b) *k odd; h is a multiple of k.* We have

$$N_k(h) = \sum_{r=0}^{h/k-1} 2^{r(k-1)} 2^{h-rk-1} + 2^{h(k-1)/k} = \sum_{r=0}^{h/k-1} 2^{h-1-r} + 2^{h-h/k} = 2^h\,.$$

(c) *k even; $h \leqq \mu + 1$ (except the case $k = 2$, $h = 2$).* We have obviously

$$N_k(h) = 0\,.$$

(d) *k even; $h > \mu + 1$; k is not a divisor of h.* We have

$$N_k(h) = \sum_{r=0}^{[(h-\mu-2)/k]} 2^{r(k-1)} 2^{h-rk-1} = \sum_{r=0}^{[(h-\mu-2)/k]} 2^{h-1-r} = 2^h(1 - 2^{-[(h+k-\mu-2)/k]})\,.$$

(e) *k even and different from 2; $h > \mu + 1$; k is a divisor of h.* We have

$$N_k(h) = \sum_{r=0}^{[(h-\mu-2)/k]} 2^{r(k-1)} 2^{h-rk-1} + 2^{h(k-1)/k} = 2^h(1 - 2^{-[(h+k-\mu-2)/k]}) + 2^{h-h/k}$$

but in the conditions which apply to the present case we also have $k \geqq \mu + 2$ which implies $[(h + k - \mu - 2)/k] = h/k$ so that $N_k(h) = 2^h$.

The only case left is now

(f) *$k = 2$; h even.* We have, for $h > 2$

$$N_k(h) = \sum_{r=0}^{h/2-2} 2^r 2^{h-2r-1} + 2^{h/2} = \sum_{r=0}^{h/2-2} 2^{h-1-r} + 2^{h/2} = 2^h(1 - 2^{-h/2})\,,$$

whereas for $h = 2$ we have simply $N_k(h) = 2$ which coincides with the previous formula.

The theorem is proved for $k > 2$ by the formulas of cases from (a) to (e), whereas for $k = 2$ it is proved by the formulas of cases (d) and (f), if we observe that for h odd and $k = 2$ we have

$$[(h + k - \mu - 2)/k] = (h - 1)/2 = [h/2]\,.$$

Istituto di Matematica
Università di Bari
Bari, Italy

Compagnia Generale Automazione
20, Via Fumaroli
Roma, Italy