Classroom

In this section of *Resonance*, we invite readers to pose questions likely to be raised in a classroom situation. We may suggest strategies for dealing with them, or invite responses, or both. "Classroom" is equally a forum for raising broader issues and sharing personal experiences and viewpoints on matters related to teaching and learning science.

B Sury
Stat-Math Unit
Indian Statistical Institute
8th Mile Mysore Road
Bangalore 560 059, India.
Email:sury@isibang.ac.in

Revisiting Kummer's and Legendre's Formulae

In a beginning course in number theory, an elementary exercise is to compute the largest power of a prime p dividing $n!$. This number, called the p-adic valuation of $n!$, is easily proved to be

$$v_p(n!) = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots. \qquad (1)$$

Note that this is a finite series. The number $v_p(n!)$ comes up naturally in a few situations like the following. In the group of permutations of n objects, this would give the power of p which is the order of a p-Sylow subgroup. While discussing p-adic numbers as analogues of the usual real numbers, one looks at the analogue of the exponential series. The expression for $v_p(n!)$ leads one to determine that the exponential series has the radius of convergence $p^{-1/(p-1)}$.

Now, $v_p(n!)$ can also be computed in another manner by a beautiful observation due to the legendary mathematician Legendre. Legendre observed that the p-adic

Keywords
p-adic valuation, base-p expansion, Legendre's formula, Kummer's formula.

valuation of $n!$ can be read off from the base-p expansion of n. It is simply $\frac{n-s(n)}{p-1}$ where $s(n)$ is the sum of the digits of n in this expansion. A related result that Kummer proved is that, if $r \leq n$, then the p-adic valuation of the binomial coefficient $\binom{n}{r}$ is simply the number of 'carry-overs' when one adds r and $n-r$ in base-p. In [1], Honsberger deduces Kummer's theorem from Legendre's result and refers to Ribenboim's lovely book [2] for a proof of the latter. Ribenboim's proof is by verifying that Legendre's base-p formula agrees with the standard formula given in (1).

The p-adic valuation of $n!$ can be read off from the base-p expansion of n. It is simply $\frac{n-s(n)}{p-1}$ where $s(n)$ is the sum of the digits of n in this expansion.

Is it possible to prove Legendre's formula without recourse to the above formula (1)? We shall see that this is indeed possible and that the standard formula follows from such a proof. What is more, Kummer's formula also follows without having to use Legendre's result. The author's long-standing belief that these proofs are more natural than the ones quoted above was vindicated during a selection interview to an undergraduate programme, when an outstanding candidate Swarnendu Datta came up essentially with the same proof! Let us start by recalling Legendre's formula.

Legendre's Formula

Let p be a prime number and let $a_k \cdots a_1 a_0$ be the base-p expansion of a natural number n. We shall show that if Legendre's formula

$$v_p(n!) = \frac{n - s(n)}{p-1} = \frac{n - \sum_{i=0}^{k} a_i}{p-1} \qquad (2)$$

holds good for n, then it also holds good for $pn + r$ for any $0 \leq r < p$. Note that the base-p expansion of $pn+r$ is

$$a_k \cdots a_1 a_0 \; r.$$

Let us denote, for convenience, the number $\frac{m-s(m)}{p-1}$ by

$f(m)$ for any natural number m. Evidently,

$$f(pn+r) = \frac{pn - \sum_{i=0}^{k} a_i}{p-1} = n + f(n).$$

On the other hand, it follows by induction on n that

$$v_p((pn+r)!) \doteq n + v_p(n!). \tag{3}$$

For, if it holds good for all $n < m$, then

$$\begin{aligned} v_p((pm+r)!) &= v_p(pm) + v_p((pm-p)!) \\ &= 1 + v_p(m) + m - 1 + v_p((m-1)!) \;=\; m + v_p(m!). \end{aligned}$$

Since it is evident that $f(m) = 0 = v_p(m!)$ for all $m < p$, it follows that $f(n) = v_p(n!)$ for all n. This proves Legendre's formula.

Note also that the formula

$$v_p(n!) = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots$$

follows inductively on using (3).

Kummer's Algorithm

As before p is any prime number. For any natural numbers r and s, let us denote by $g(r,s)$ the number of 'carry-overs' when the base-p expansions of r and s are added. Kummer's result is that for $k \leq n$,

$$v_p\left(\binom{n}{k}\right) = g(k, n-k). \tag{4}$$

Once again, this is clear if $n < p$, as both sides are then zero. We shall show that if the formula holds good for n (and *every* $k \leq n$), it does so for $pn + r$ for $0 \leq r < p$ (and any $k \leq pn + r$). This would prove the result for all natural numbers.

Consider any binomial coefficient $\binom{pn+r}{pm+a}$ for $0 \leq a < p$.

First, suppose $a \leq r$.

Write $m = b_k \cdots b_0$ and $n-m = c_k \cdots c_0$ in base-p. Then the base-p expansions of $pm + a$ and $p(n-m) + (r-a)$ are, respectively,

$$\begin{aligned} pm + a &= b_k \cdots b_0\ a, \\ p(n-m) + (r-a) &= c_k \cdots c_0\ r - a. \end{aligned}$$

Evidently, the corresponding number of carry-overs is

$$f(pm + a, p(n-m) + (r-a)) = f(m, n-m).$$

By the induction hypothesis, $f(m, n-m) = v_p(\binom{n}{m})$.

Now $v_p\left(\binom{pn+r}{pm+a}\right)$ is equal to

$$v_p((pn+r)!) - v_p((pm+a)!) - v_p((p(n-m)+r-a)!)$$

$$= n + v_p(n!) - m - v_p(m!) - (n-m) - v_p((n-m)!)$$

$$= v_p\left(\binom{n}{m}\right).$$

Thus, we are through in the case when $a \leq r$.

Now suppose that $r < a$. Then $v_p\left(\binom{pn+r}{pm+a}\right)$ is equal to

$$v_p((pn+r)!) - v_p((pm+a)!) - v_p((p(n-m-1)+(p+r-a))!)$$

$$= n + v_p(n!) - m - v_p(m!) - (n-m-1) - v_p((n-m-1)!)$$

$$= 1 + v_p(n) + v_p((n-1)!) - v_p(m!) - v_p((n-m-1)!)$$

$$= 1 + v_p(n) + v_p\left(\binom{n-1}{m}\right).$$

We need to show that

$$\begin{aligned} &f(pm + a, p(n-m-1) + (p+r-a)) \\ &\quad = 1 + v_p(n) + f(m, n-m-1). \end{aligned} \tag{5}$$

Note that $m < n$. Write $n = a_k \cdots a_0$, $m = b_k \cdots b_0$ and $n - m - 1 = c_k \cdots c_0$ in base-p. If $v_p(n) = d$, then $a_i = 0$ for $i < d$ and $a_d \neq 0$. In base-p, we have

$$n = a_k \cdots a_d \, 0 \cdots 0$$

and, therefore,

$$n - 1 = a_k \cdots a_{d+1} a_d - 1 \; p - 1 \cdots \; p - 1.$$

Now, the addition $m + (n - m - 1) = n - 1$ gives $b_i + c_i = p - 1$ for $i < d$ (since they must be $< 2p - 1$). Moreover, $b_d + c_d = a_d - 1$ or $p + a_d - 1$.

Note the base-p expansions

$$\begin{aligned} pm + a &= b_k \cdots b_0 \; a, \\ p(n - m - 1) + (p + r - a) &= c_k \cdots c_0 \; p + r - a. \end{aligned}$$

We add these using the fact that there is a carry-over in the beginning and that $1 + b_i + c_i = p$ for $i < d$. Since there is a carry-over at the first step as well as at the next d steps, we have

$$pn + r = * \; * \; \cdots \; a_d \, 0 \cdots 0 \; r$$

where there are d zeroes before r, and

$$f(pm+a, p(n-m-1)+(p+r-a)) = 1+d+f(m, n-m-1).$$

This proves Kummer's assertion also.

Suggested Reading

[1] R Honsberger, *In Polya's Footsteps*, published and distributed by the Mathematical Association of America, pp.229-233, 1997.

[2] P Ribenboim, *The Book of Prime Number Records*, Springer-Verlag, pp.30-32, 1996.