The following proof was obtained by Jessica Bernier and William Yslas Vélez, October 2008

Theorem: $\sum_{k=1}^{n} k^r$ ($r \in N$) is a polynomial in the variable n of degree r + 1, where N ={*1,2,3,…*}, the set of natural numbers.

<u>Proof</u>:

We will prove the theorem by induction on r. Let's begin with the base case, r = 1. For this case, we want to show that $\sum_{k=1}^{n} k$ is a quadratic polynomial. From Fig. 1, we see that we can express the sum as:

$$\sum_{k=1}^{n} k = \int_0^n x\,dx + \frac{n}{2} \tag{1}$$

In other words, $\sum_{k=1}^{n} k$ is the right-hand Riemann sum approximation of the integral $\int_0^n x\,dx$, which exceeds the value of the integral by the collective area above the line $y = x$ (this area is $n/2$).

Since the integral $\int_0^n x\,dx$ evaluates to $\frac{1}{2}x^2\Big|_0^n = \frac{1}{2}n^2$, we arrive at the equation

$$\sum_{k=1}^{n} k = \tfrac{1}{2}n^2 + \tfrac{1}{2}n = \frac{n(n+1)}{2},$$

which is indeed a quadratic polynomial.

Now, we will prove the theorem is true for a r + 1 by using "strong induction". In the usual form of induction we only use the assumption that the assertion is true for r to prove that the assertion is true for $r + 1$. In "strong induction" we use the assumption that the statement is true for all s, where $0 < s < r + 1$.

Let's return to the proof. In analogy to the base case, we can express $\sum_{k=1}^{n} k^r$ as the right-hand Riemann sum approximation of the integral $\int_0^n x^r\,dx$ (see Fig. 2). Thus,

$$\sum_{k=1}^{n} k^r = \int_0^n x^r\,dx + \sum_{k=1}^{n}\left[\int_{k-1}^{k}(k^r - x^r)\,dx\right],$$

and first integral evaluates to $\frac{n^{r+1}}{r+1}$. The integral inside the summation is a little trickier to integrate. (In case you're wondering where this term came from, the integral is the area between k^r and x^r, and is only one unit wide, which makes sense when we closely examine Fig. 2. We must then sum each individual area to get the total area above the line $y = x^r$. But, one step at a time. Let's tackle the integral first.) We know:

$$\int_{k-1}^{k}(k^r - x^r)\,dx = \left[k^r x - \frac{x^{r+1}}{r+1}\right]_{k-1}^{k}$$

$$= \left(k^{r+1} - \frac{k^{r+1}}{r+1}\right) - \left(k^r(k-1) - \frac{(k-1)^{r+1}}{r+1}\right) \qquad (2)$$

$$= k^r + \frac{(k-1)^{r+1} - k^{r+1}}{r+1}$$

after a bit of algebra. Now, we want to group all the like powers together, so that we have only one term with k^{r+1} in it, one term with k^r, and so on (the motivation for this will become apparent later). To do this, observe that we can expand the polynomial $(k-1)^{r+1}$ as:

$$(k-1)^{r+1} = \sum_{p=0}^{r+1}\binom{r+1}{p}k^{r+1-p}(-1)^p \ ,$$

where

$$\binom{r+1}{p} = \frac{(r+1)!}{p!(r+1-p)!} \qquad (3)$$

is the binomial coefficient. Now, we rewrite the polynomial $(k-1)^{r+1}$ as:

$$(k-1)^{r+1} = \binom{r+1}{0}k^{r+1}(-1)^0 + \binom{r+1}{1}k^r(-1)^1 + \binom{r+1}{2}k^{r-1}(-1)^2 + \ldots$$

$$+ \binom{r+1}{r}k^1(-1)^r + \binom{r+1}{r+1}k^0(-1)^{r+1}$$

$$= k^{r+1} - (r+1)k^r + \binom{r+1}{2}k^{r-1} - \ldots + \binom{r+1}{r}k^1(-1)^r + (-1)^{r+1}$$

(some binomial coefficients have been evaluated for clarity). Now, when we plug this expression back into Eq. 2, we have

$$k^r + \frac{\left[k^{r+1} - (r+1)k^r + \binom{r+1}{2}k^{r-1}(-1)^2 + \binom{r+1}{3}k^{r-2}(-1)^3 \ldots + \binom{r+1}{r}k^1(-1)^r + (-1)^{r+1}\right] - k^{r+1}}{r+1}$$

We see the k^{r+1} and k^r terms cancel, leaving

$$\int_{k-1}^{k}(k^r - x^r)\,dx = \frac{\binom{r+1}{2}k^{r-1}(-1)^2 + \binom{r+1}{3}k^{r-2}(-1)^3 + \ldots + \binom{r+1}{r}k^1(-1)^r + (-1)^{r+1}}{r+1}$$

as our final expression for the integral.

We are not done yet, for we must still take the summation over *k*. This becomes:

$$\sum_{k=1}^{n}\left[\frac{\binom{r+1}{2}k^{r-1} + \binom{r+1}{3}k^{r-2}(-1)^3 \ldots + \binom{r+1}{r}k^1(-1)^r + (-1)^{r+1}}{r+1}\right]$$

which we can separate into terms according to

$$\frac{1}{r+1}\binom{r+1}{2}\sum_{k=1}^{n}k^{r-1}+\frac{1}{r+1}\binom{r+1}{3}\sum_{k=1}^{n}k^{r-2}(-1)^{3}...+\frac{1}{r+1}\binom{r+1}{r}\sum_{k=1}^{n}k^{1}(-1)^{r}+\frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1}$$

Now, we can finally use the inductive step. Since we have assumed that $\sum_{k=1}^{n}k^{r-1}$ is a polynomial of degree *r*, and since we can likewise argue that each of the summations will be of lesser degree, the whole expression must be (at most) degree *r*. Applying the above expression into the equation for $\sum_{k=1}^{n}k^{r}$, we obtain

$$\begin{aligned}\sum_{k=1}^{n}k^{r} &= \int_{0}^{n}x^{r}\,dx+\sum_{k=1}^{n}\left[\int_{k-1}^{k}(k^{r}-x^{r})\,dx\right] \\ &= \frac{n^{r+1}}{r+1}+\left[\begin{aligned}&\frac{1}{r+1}\binom{r+1}{2}\sum_{k=1}^{n}k^{r-1}\ \frac{1}{r+1}\binom{r+1}{3}\sum_{k=1}^{n}k^{r-2}(-1)^{3}+... \\ &+\frac{1}{r+1}\binom{r+1}{r}\sum_{k=1}^{n}k^{1}(-1)^{r}+\frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1}\end{aligned}\right]\end{aligned} \tag{4}$$

From this we see that $\sum_{k=1}^{n}k^{r}$ is a polynomial of degree $r+1$. Thus, we have proven that the assertion is true for ($r\in N$).

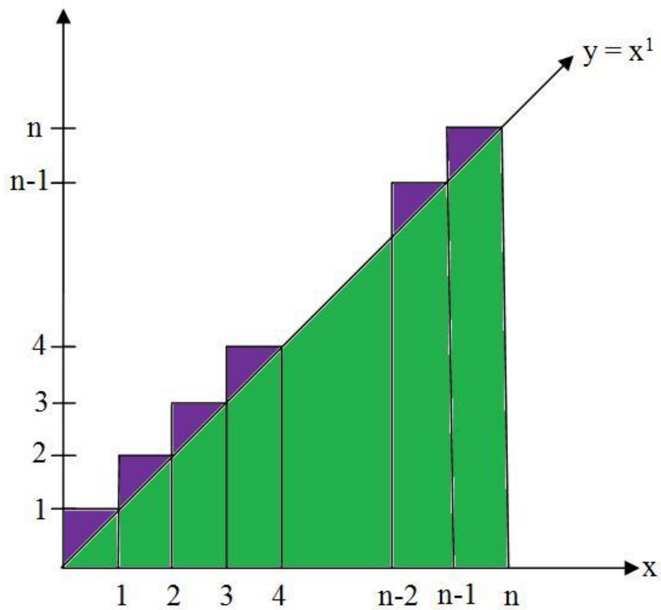

Figure 1. The sum $\sum_{k=1}^{n} k$ is the total colored area. $\int_0^n x\,dx$ is the area in green and $n/2$ is the collective area in purple.

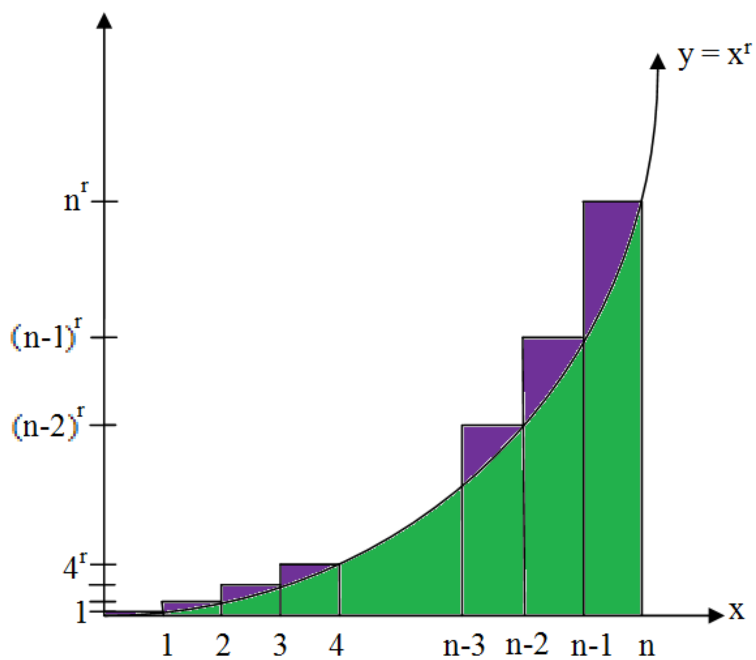

Figure 2. The sum $\sum_{k=1}^{n} k^r$ is the total colored area. $\int_0^n x^r\,dx$ is the area in green. The collective area in purple is $\sum_{k=1}^{n}\left[\int_{k-1}^{k}(k^r - x^r)\,dx\right]$, where each chunk has volume $\int_{k-1}^{k}(k^r - x^r)\,dx$.

Corollary 1: $(r+1)!\sum_{k=1}^{n} k^r$ is a polynomial in *n* with integer coefficients.

Proof: If we multiply Eq. 4 by (*r* + 1)!, we obtain

$$(r+1)!\sum_{k=1}^{n} k^r = n^{r+1} + \left[\begin{array}{l} \binom{r+1}{2}\sum_{k=1}^{n} k^{r-1} + (-1)^3\binom{r+1}{3}\sum_{k=1}^{n} k^{r-2} + \ldots \\ \quad + (-1)^r\binom{r+1}{r}\sum_{k=1}^{n} k^1(-1)^r + (-1)^{r+1}\sum_{k=1}^{n} 1 \end{array}\right].$$

Now, the binomial coefficients defined in Eq. 3 are always natural numbers (if you don't believe this, try a few examples yourself). However, some terms are multiplied by -1, so overall we can only conclude that the coefficient of each term is an integer.

As a quick check, we verify that $2!\sum_{k=1}^{n} k = n(n+1)$ and $3!\sum_{k=1}^{n} k^2 = n(n+1)(2n+1)$ are indeed polynomials in *n* with integer coefficients.

Corollary 2: *n* = 0 is a root of the polynomial $\sum_{k=1}^{n} k^r$ $(r \in N)$.

Proof: We will prove this by induction. For the base case *r* = 1, the polynomial $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ which we obtained from Eq. 1 clearly has a root at *n* = 0. For the case of arbitrary $r \in N$, let's refer to the result of Eq. 4:

$$\sum_{k=1}^{n} k^r = \frac{n^{r+1}}{r+1} + \left[\begin{array}{l} \frac{1}{r+1}\binom{r+1}{2}\sum_{k=1}^{n} k^{r-1} + \frac{1}{r+1}\binom{r+1}{3}\sum_{k=1}^{n} k^{r-2}(-1)^3 + \ldots \\ \quad + \frac{1}{r+1}\binom{r+1}{r}\sum_{k=1}^{n} k^1(-1)^r + \frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1} \end{array}\right]$$

Observe that $\frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1}$ gives $\pm n/(r+1)$, so the first term and the last term have 0 as a root. If we apply the induction hypothesis, then each term in the brackets, except for the last term, has a factor $\sum_{k=1}^{n} k^s$ where *s* < *r*. The induction hypothesis applies to yield that each of these terms has 0 as a root, thus proving the corollary.

Corollary 3: *n* = -1 is a root of the polynomial $\sum_{k=1}^{n} k^r$ $(r \in N)$.

Proof: We will prove this by induction, but in a slightly different way than for Corollary 2. For the base case *r* = 1, the polynomial $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ clearly has a root at *n* = -1. For the case of arbitrary $r \in N$, we again refer to the result of Eq. 4:

$$\sum_{k=1}^{n} k^r = \frac{n^{r+1}}{r+1} + \left[\begin{array}{c} \frac{1}{r+1}\binom{r+1}{2}\sum_{k=1}^{n} k^{r-1} + \frac{1}{r+1}\binom{r+1}{3}\sum_{k=1}^{n} k^{r-2}(-1)^3 + \ldots \\ + \frac{1}{r+1}\binom{r+1}{r}\sum_{k=1}^{n} k^1(-1)^r + \frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1} \end{array} \right]$$

We know that *n* = -1 is a root of the polynomial $\sum_{k=1}^{n} k^r$ if we get 0 after we plug *n* = -1 into the polynomial equation. Our induction hypothesis (*n* = -1 is a root of all polynomials $\sum_{k=1}^{n} k^s$ ($s < r$) allows us to conclude that all of the summations in the square brackets are 0 except $\frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1}$, because this is a summation of k^0. We are left with the equation

$$\sum_{k=1}^{n} k^r = \frac{n^{r+1}}{r+1} + \frac{1}{r+1}\sum_{k=1}^{n}(-1)^{r+1}$$

Plugging *n* = -1 into this equation does not make sense because we would be taking a summation from 1 to -1. However, we can rewrite that term to give the polynomial

$$\sum_{k=1}^{n} k^r = \frac{n^{r+1}}{r+1} + \frac{(-1)^{r+1} n}{r+1},$$

which we can evaluate at *n* = -1 without problem. We get

$$\sum_{k=1}^{n} k^r = \frac{(-1)^{r+1}}{r+1} + \frac{(-1)^{r+1}(-1)}{r+1} = 0$$

Thus, *n* = -1 is a root of the polynomial $\sum_{k=1}^{n} k^r$.