What is the "polynomial identity trick"?

<u>A reminder on polynomials.</u> Polynomials are NOT functions.

Informally, a polynomial (with rational coefficients, in 1 variable X) is a "formal expression" of the form

$$\alpha X^a + \beta X^b + \gamma X^c + \cdots + \omega X^z \quad \text{with } \alpha, \beta, \ldots, \omega \in \mathbb{Q}$$

and $a, b, \ldots, z \in \mathbb{N}$.

These expressions obey rules: $\varphi X^n + \psi X^n = (\varphi + \psi) X^n$ ("combining like terms"),

and terms $0X^a$ can be removed,

and terms can be swapped.

and X^1 is written as X,

X^0 is written as 1,

subtraction is defined ~~as~~ by $(\alpha X^a + \beta X^b) - (\gamma X^c + \delta X^d)$
$= \alpha X^a + \beta X^b + (-\gamma) X^c + (-\delta) X^d$

etc.;

multiplication is defined by distributivity and $(\alpha X^a)(\beta X^b) = \alpha\beta X^{a+b}$;

the <u>degree</u> of a polynomial is the largest exponent appearing in it with coefficient $\neq 0$ (e.g., the degree of $2+7X^5+3X^2$ is 5).

<u>Substituting</u> a number or matrix or polynomial x into a polynomial $P = \alpha X^a + \beta X^b + \gamma X^c + \dots$ yields $\alpha x^a + \beta x^b + \gamma x^c + \dots$. This result is called $P(x)$.

A number x ($\in \mathbb{Q}$ or $\in \mathbb{R}$ or $\in \mathbb{C}$) is a <u>root</u> of a polynomial P if $P(x) = 0$.

For a formal definition of polynomials, see [detnotes, §1.5], or [Loehr] (most recommended!), or most good algebra texts.

<u>Thm. 3.20 ("polynomial identity trick").</u>

(a) A polynomial (with rational coefficients, in 1 variable X) of degree $\leq n$ (for a given $n \in \mathbb{N}$) has $\leq n$ roots (in $\mathbb{Q}$, in $\mathbb{R}$ or in $\mathbb{C}$), unless it is the 0 polynomial (i.e., its coefficients are all 0).

(To me, the 0 polynomial has degree $-\infty$, which is $<$ to any integer.)

(b) If a polynomial has infinitely many roots, then it is the 0 polynomial.

(c) Let P and Q be two polynomials. If

$$P(x) = Q(x) \quad \forall\, x \in \mathbb{N},$$

then $P = Q$.

Proof. E.g., see Goodman, "Algebra: Abstract & Concrete", Cor. 1.8.24. $\square$

Salvaging our 1st proof of Theorem 3.18. Fix $y \in \mathbb{Q}$, $n \in \mathbb{N}$.

We've already proven

$$(1) \qquad \binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k}\binom{y}{n-k}$$

for all $x \in \mathbb{N}$. We want to prove it for $x \in \mathbb{Q}$.

Define two polynomials P and Q by

$$P = \binom{X+y}{n} \quad \text{and} \quad Q = \sum_{k=0}^{n} \binom{X}{k}\binom{y}{n-k}.$$

These are well-defined polynomials, since

$$P = \binom{X+y}{n} = \frac{(X+y)(X+y-1)\cdots(X+y-n+1)}{n!}$$

and

$$Q = \sum_{k=0}^{n} \binom{X}{k}\binom{y}{n-k} = \sum_{k=0}^{n} \frac{X(X-1)\cdots(X-k+1)}{k!}\binom{y}{n-k}.$$

Then, $P(x) = Q(x)$ $\forall x \in \mathbb{N}$ (since we've showed (1) for all $x \in \mathbb{N}$). Thus, Thm. 3.20 (c) says: $P = Q$. Hence, $P(x) = Q(x)$ $\forall x \in \mathbb{Q}$. In other words, (1) holds for all $x \in \mathbb{Q}$. This completes 1st proof of Theorem 3.18. $\square$

<u>Salvaging our 2nd proof of Theorem 3.18.</u> Same idea, but we need to do it twice:

<u>1st step:</u> Fix $y \in \mathbb{N}$, $n \in \mathbb{N}$, and use the same argument as above to prove that (1) holds for all $x \in \mathbb{Q}$.

2nd step: Fix $x \in \mathbb{Q}, n \in \mathbb{N}$. Use an analogous argument (using y instead of x) to prove that (1) holds for all $y \in \mathbb{Q}$. $\square$

Rmk. Theorem 3.20 (c) can be applied to some other identities:

- Proposition 3.2 (Trinomial revision) says
$$\binom{n}{a}\binom{a}{b} = \binom{n}{b}\binom{n-b}{a-b} \quad \forall n, a, b \in \mathbb{R}.$$
We gave a bijective proof for the case $n, a, b \in \mathbb{N}$.
Using Thm. 3.20 (c), we can extend this to $n \in \mathbb{R}$, but a, b still need to stay $\in \mathbb{N}$ (we cannot replace a by X, since "$\binom{n}{X}$" doesn't make sense).

- Cor. 3.3 said $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = [n=0] \quad \forall n \in \mathbb{N}$.
This cannot be generalized to $n \in \mathbb{Q}$ or $n \in \mathbb{R}$, since n appears as an upper bound of the sum.

- Thm. 3.15 said $k^m = \sum_{i=0}^{m} \operatorname{sur}(m,i) \binom{k}{i} \quad \forall k \in \mathbb{N}\ \forall m \in \mathbb{N}$.
Thm. 3.20 (c) yields that this also holds $\forall k \in \mathbb{R}\ \forall m \in \mathbb{N}$.

But n must remain $\in \mathbb{N}$.

Thm. 3.18 is called the <u>(Chu-) Vandermonde (convolution) identity</u>. It has several "mutated" versions.

The following two identities can be used for the "mutation":

- UpNeg (upper negation $=$ Prop. 1.11 b): $\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}$ $\forall n \in \mathbb{R}$, $k \in \mathbb{Z}$.
- Symm (symmetry $=$ Thm. 1.13): $\binom{n}{k} = \binom{n}{n-k}$ $\forall n \in \mathbb{N}$, $k \in \mathbb{Z}$.

One example of a "mutation" of Chu-Vandermonde is:

<u>Prop. 3.21 ("Upside-down Vandermonde")</u>: Let $n, x, y \in \mathbb{N}$.

Then:
$$\binom{n+1}{x+y+1} = \sum_{k=0}^{n} \binom{k}{x} \binom{n-k}{y}.$$

~~Proof~~ <u>Remark.</u> This cannot be generalized using Thm. 3.20 (c), since the RHS is not a polynomial function in any of n, x, y. And indeed, the equality is false for $n = 3$, $x = 1/2$, $y = 1/2$.

1st proof of Prop. 3.21. (from [detnotes, §2.3])

If $n < x+y$, then both sides are 0 (in fact, each addend on the RHS is 0).

(in fact, let $k \in \{0,1,\ldots,n\}$; if $\binom{k}{x}\binom{n-k}{y} \neq 0$, then $k \geq x$ & $n-k \geq y$ $\Longrightarrow n = k+(n-k) \geq x+y$.)

So we WLOG assume $n \geq x+y$.

Then,
$$\sum_{k=0}^{n} \binom{k}{x}\binom{n-k}{y} \overset{0}{=\!=} \sum_{k=x}^{n-y} \underbrace{\binom{k}{x}}_{(1)} \underbrace{\binom{n-k}{y}}_{(2)}$$

$$(1):\quad \binom{k}{x} \overset{\text{Symm}}{=\!=} \binom{k}{k-x} \overset{\text{UpNeg}}{=\!=} (-1)^{k-x} \binom{-k+(k-x)-1}{k-x} = (-1)^{k-x}\binom{-x-1}{k-x}$$

$$(2):\quad \binom{n-k}{y} \overset{\text{Symm}}{=\!=} \binom{n-k}{n-k-y} \overset{\text{UpNeg}}{=\!=} (-1)^{n-k-y} \binom{-(n-k)+(n-k-y)-1}{n-k-y} = (-1)^{n-k-y}\binom{-y-1}{n-k-y}$$

$$= \sum_{k=x}^{n-y} (-1)^{k-x} \binom{-x-1}{k-x} (-1)^{n-k-y} \binom{-y-1}{n-k-y}$$

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

(here, we substituted $k+x$ for k)

$$= (-1)^{n-x-y} \underbrace{\sum_{k=0}^{n-x-y} \binom{-x-1}{k} \binom{-y-1}{n-x-y-k}}_{= \binom{(-x-1)+(-y-1)}{n-x-y}}$$

(by Thm. 3.18, applied to $n-x-y$, $-x-1$ and $-y-1$ instead of n, x and y)

$$= (-1)^{n-x-y} \binom{(-x-1)+(-y-1)}{n-x-y}$$ ###

$$\overset{\text{UpNeg}}{=} \underbrace{(-1)^{n-x-y}\,(-1)^{n-x-y}}_{=1} \underbrace{\binom{(n-x-y)-((-x-1)+(-y-1))-1}{n-x-y}}_{=\binom{n+1}{n-x-y}}$$

$$= \binom{n+1}{n-x-y} \overset{\text{Symm}}{=} \binom{n+1}{(n+1)-(n-x-y)} = \binom{n+1}{x+y+1}. \qquad \square$$

2nd proof of Prop. 3.21. Double-count the number of ~~...~~ $(x+y+1)$-elt. subsets of $[n+1]$:

1st answer: $\binom{n+1}{x+y+1}$.

2nd answer: Here is a way of constructing these subsets:

- Choose the $(x+1)$-th smallest element of this subset. Call it $k+1$ (so $k \in \{0,1,\dots,n\}$).
- Choose the x ~~elements of the~~ smallest elements of this subset. There are $\binom{k}{x}$ options for this (since we're choosing them from the k-elt. set $\{1,2,\dots,k\}$).
- Choose the remaining y elements of this subset. There are $\binom{n-k}{y}$ options for this (since we're choosing them from the $(n-k)$-elt. set $\{k+2, k+3, \dots, n+1\}$).

$\Rightarrow$ The ~~...~~ answer is $\sum_{k=0}^{n} \binom{k}{x}\binom{n-k}{y}$.

Compare $\Rightarrow \binom{n+1}{x+y+1} = \sum_{k=0}^{n} \binom{k}{x}\binom{n-k}{y}$. $\square$

3.5. COUNTING SUBSETS AGAIN

Recall Theorem 1.15: It says that if S is an n-elt. subset (for some $n \in \mathbb{N}$), and if $k \in \mathbb{Z}$, then

$$\binom{n}{k} = (\#\text{ of } k\text{-elt. subsets of } S).$$

We proved this by induction. We'll now reprove it by "multijection".

<u>Def.</u> Let S be a set. Let $k \in \mathbb{N}$. A k-tuple $(s_1, s_2, \ldots, s_k) \in S^k$ is called <u>injective</u> if $s_1, s_2, \ldots, s_k$ are distinct.

Let S^k_{dist} be the set of all ~~k~~ injective k-tuples in S^k.

(<u>Ex</u>: $(3,2,5)$ is injective 3-tuple, but $(4,2,4)$ is not.)

(Injective k-tuples = k-samples without replacement.)

<u>Prop. 3.22.</u> Let S be a set. Let $k \in \mathbb{N}$. Then,

$$|S^k_{\text{dist}}| = |S| \cdot (|S|-1) \cdot (|S|-2) \cdot \cdots \cdot (|S|-k+1).$$

Proof. The injective k-tuples are in 1-to-1 corresp. with the injective maps from $[k]$ to S.

Rigorously: There is a bijection

$$\{\text{inj. maps from } [k] \text{ to } S\} \to S^k_{\text{dist}},$$
$$f \mapsto (f(1), f(2), \ldots, f(k)).$$

Thus,
$$|S^k_{\text{dist}}| = |\{\text{inj. maps from } [k] \text{ to } S\}|$$
$$= (\text{\# of inj. maps from } [k] \text{ to } S)$$
$$= |S| \cdot (|S|-1) \cdot (|S|-2) \cdot \ldots \cdot (|S|-k+1)$$

(by Thm. 3.5, applied to $A = [k]$, $B = S$, $m = k$, $n = |S|$). $\square$

2nd proof of Thm. 1.15. WLOG assume $k \geq 0$ (else we just claim $0 = 0$). Then, $|S| = n$. Hence, Prop. 3.22 yields

(1) $$|S^k_{\text{dist}}| = n(n-1)(n-2) \cdots (n-k+1).$$

On the other hand,

$$|S^k_{\text{dist}}| = (\#\text{ of inj. } k\text{-tuples } \vec{s} \in S^k)$$

$$= \sum_{W \subseteq S;\ |W|=k} \underbrace{(\#\text{ of inj. } k\text{-tuples } \vec{s} \in S^k \text{ such that the set of the entries of } \vec{s} \text{ is } W)}_{= (\#\text{ of inj. } k\text{-tuples } \vec{s} \in W^k)}$$

(because if $\vec{s} \in W^k$ is an inj. k-tuple, then the pigeonhole principle shows that all elements of W must appear in $\vec{s}$, and thus the ~~en~~ set of the entries of $\vec{s}$ is W)

$$= \sum_{W \subseteq S;\ |W|=k} \underbrace{(\#\text{ of inj. } k\text{-tuples } \vec{s} \in W^k)}_{= |W^k_{\text{dist}}| = k(k-1)(k-2)\cdots(k-k+1)}$$

(again by Prop 3.22)

$= k!$.

$$= \sum_{W \subseteq S;\ |W| = k} k!$$

$$= k! \cdot (\text{\# of } k\text{-element subsets } W \text{ of } S).$$

Comparing this with (1), we get

$$k! \cdot (\text{\# of } k\text{-element subsets } W \text{ of } S) = n(n-1)(n-2)\cdots(n-k+1).$$

Hence,

$$(\text{\# of } k\text{-element subsets } W \text{ of } S) = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

$$= \binom{n}{k}.$$

$\square$

So Theorem 1.15 again.

(This was an example of "multijective proof" or "shepherd's principle".)

4th proof of Thm. 3.18 in the case when $x \in \mathbb{N}$ and $y \in \mathbb{N}$.

Rename x and y as a and b.

So we must show:

$$(1) \qquad \binom{a+b}{n} = \sum_{k=0}^{n} \binom{a}{k} \binom{b}{n-k}.$$

Now, consider the polynomial $(1+X)^{a+b}$. Compare

$$(1+X)^{a+b} = \sum_{m=0}^{a+b} \binom{a+b}{m} X^m \quad \text{(by the binomial formula)}$$

$$\overset{\circ}{=} \sum_{m} \binom{a+b}{m} X^m$$

with

$$(1+X)^{a+b} = \underbrace{(1+X)^a}_{\substack{= \sum_i \binom{a}{i} X^i \\ \text{(by the binom. fml.)}}} \underbrace{(1+X)^b}_{\substack{= \sum_j \binom{b}{j} X^j \\ \text{(by the binom. fml.)}}} = \left(\sum_i \binom{a}{i} X^i \right) \left(\sum_j \binom{b}{j} X^j \right)$$

$$= \sum_{(i,j)\in\mathbb{N}^2} \binom{a}{i} X^i \binom{b}{j} X^j$$

$$= \sum_{i,j} \binom{a}{i}\binom{b}{j} X^{i+j} = \sum_m \Bigg(\underbrace{\sum_{\substack{i,j \\ \text{such that} \\ i+j=m}} \binom{a}{i}\binom{b}{j}}_{= \sum_{i=0}^{m} \binom{a}{i}\binom{b}{m-i}} \Bigg) X^m$$

$$= \sum_m \left(\sum_{i=0}^{m} \binom{a}{i}\binom{b}{m-i} \right) X^m,$$

We get $\displaystyle\sum_m \binom{a+b}{m} X^m = \sum_m \left(\sum_{i=0}^{m} \binom{a}{i}\binom{b}{m-i} \right) X^m.$

These are equal as polynomials, i.e., corresponding coefficients are equal. In other words, $\forall\, m \in \mathbb{N}$, we have

$$\binom{a+b}{m} = \sum_{i=0}^{m} \binom{a}{i}\binom{b}{m-i} = \sum_{k=0}^{m} \binom{a}{k}\binom{b}{m-k}.$$

Apply this to $m=n$, and get (1). $\square$

Rmk. A similar argument can prove the following identity:

(2) $$\sum_{i=0}^{m} (-1)^i \binom{n}{i}\binom{n}{m-i} = \begin{cases} (-1)^{m/2}\binom{n}{m/2} & \text{if } m \text{ is even;} \\ 0 & \text{if } m \text{ is odd} \end{cases}$$

for any $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

Indeed, you get (2) by expanding

$$(1-X)^n (1+X)^n = (1-X^2)^n$$

(using the binom. formula again)

and comparing coefficients.

In particular, if $m=n$, then (2) simplifies to

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i}^2 = \begin{cases} (-1)^{n/2}\binom{n}{n/2} & \text{if } n \text{ is even;} \\ 0 & \text{if } n \text{ is odd} \end{cases}.$$