A geometric proof of the binomial identity

FRANTIŠEK MARKO AND SEMYON LITVINOV

We give a geometric proof of the binomial identity

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

for all natural n and real a, b. This work was inspired by the book [1], where the binomial identity for $n = 3$ and $a, b > 0$ is proved by breaking a cube C of size $(a + b) \times (a + b) \times (a + b)$ into eight rectangular boxes and counting their volumes as follows.

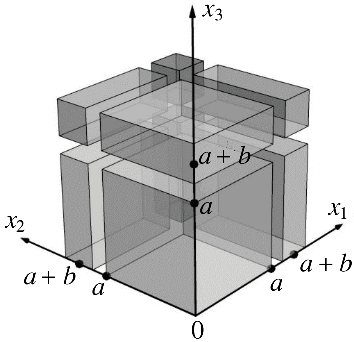

FIGURE 1: Breakdown of cube C ($n = 3$)

Looking at Figure 1, the reader can see that the volume of the cube C is $(a + b)^3$, in which there is one box of volume a^3, three boxes of volume a^2b, three boxes of volume ab^2, and one box of volume b^3. It follows that $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.

It is also mentioned in [1] that the same type of colouring could be carried out in higher dimensions, yielding the identity

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i \qquad \text{for } a, b > 0, \tag{1}$$

where $\binom{n}{i} = \dfrac{n!}{i!\,(n - i)!}$ is the binomial coefficient.

In this Article we present a geometric argument – one that does not involve induction – to establish the binomial identity for all natural n and all real a, b.

In doing so, for a given natural n, we first expand the above argument to prove identity (1) and then provide a geometric counting argument to show that

$$(a - b)^n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} a^{n-i} b^i \qquad \text{for } a > b > 0. \tag{2}$$

https://doi.org/10.1017/mag.2024.111 Published online by Cambridge University Press

The desired binomial identity

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$$

for any real a, b then easily follows; this argument is provided at the end of the Article.

We now proceed to verifying identities (1) and (2). Let us call two objects in n-dimensional Euclidean space *weakly disjoint* if the volume of their intersection is zero. If this volume is positive, we will say that they *intersect strongly*.

Identity (1): Consider the n-dimensional cube C given by the inequalities $0 \leqslant x_j \leqslant a + b$ for $1 \leqslant j \leqslant n$. Divide C by the hyperplanes $x_j = a$, for $1 \leqslant j \leqslant n$, into 2^n rectangular boxes. Each edge of these rectangular boxes is given by either $0 \leqslant x_j \leqslant a$ or $a \leqslant x_j \leqslant a + b$. If a box has exactly i edges given by $a \leqslant x_j \leqslant a + b$, then its volume is $a^{n-i}b^i$. The number of such boxes equals the number of subsets of the set $\{1, \ldots, n\}$ consisting of i elements, which is known to be $\binom{n}{i}$. Since different boxes are weakly disjoint, and the sum of volumes of all boxes equals the volume of the cube C, which is $(a + b)^n$, the argument is complete.

Identity (2): Consider the n-dimensional cube C given by $0 \leqslant x_z \leqslant a$ for $1 \leqslant j \leqslant n$. Divide C by the hyperplanes $x_j = b$, for $1 \leqslant j \leqslant n$, into 2^n rectangular boxes of the first kind, where each edge of a box is given by either $0 \leqslant x_j \leqslant b$ or $b \leqslant x_j \leqslant a$. To a box of the first kind we associate the set $K \subseteq \{1, \ldots, n\}$ such that this box is given by $b \leqslant x_j \leqslant a$ for $k \in K$ and by $0 \leqslant x_j \leqslant b$ for $k \notin K$. Denote this box by B_K^1. For $n = 3$, the breakdown of the cube C into boxes B_K^1 of the first kind is shown in Figure 2. In addition, individually, boxes B_K^1 are displayed in Figure 3.

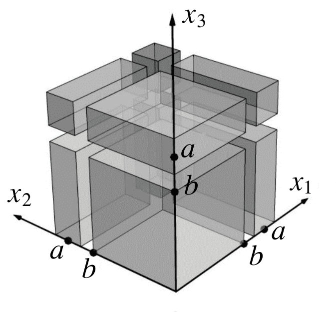

FIGURE 2: Breakdown of cube C into boxes of the first kind ($n = 3$)

https://doi.org/10.1017/mag.2024.111 Published online by Cambridge University Press

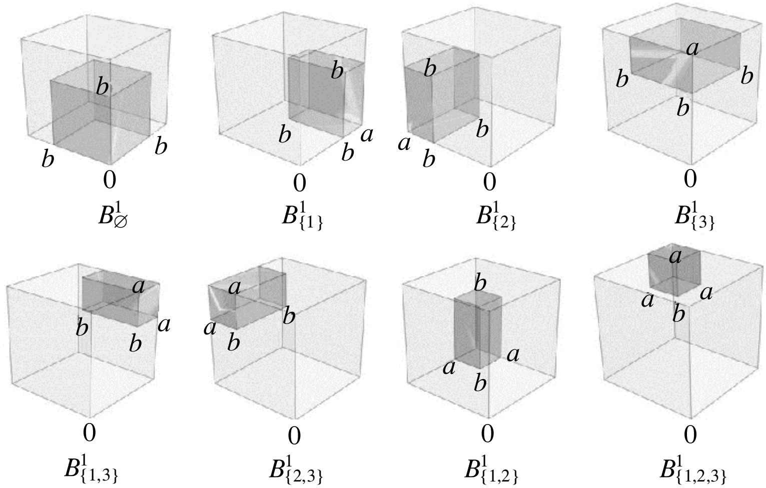

FIGURE 3: Boxes B^1_L of the first kind ($n = 3$)

Also consider 2^n rectangular boxes of the second kind, where each edge is given by either $0 \leqslant x_j \leqslant b$ or $0 \leqslant x_j \leqslant a$. To a box of the second kind we associate the set $L \subseteq \{1, \ldots, n\}$ such that this box is given by $0 \leqslant x_l \leqslant a$ for $l \in L$ and by $0 \leqslant x_l \leqslant b$ for $l \notin L$. Denote this box by B^2_L. For $n = 3$, the boxes B^2_L of the second kind are given in Figure 4.

Note that the boxes of the first kind are pairwise weakly disjoint, while the boxes of the second kind, each of which is a union of some boxes of the first kind, are not. A critical relationship between the boxes of the first and second kind is given in the following statement. Since it is used several times in the sequel, we state it as a lemma.

Lemma: The boxes B^1_K and B^2_L intersect strongly if, and only if, $K \subseteq L$ which happens if, and only if, $B^1_K \subseteq B^2_L$.

Proof: If $K \subseteq L$, then $B^1_K \subseteq B^2_L$, so B^1_K and B^2_L intersect strongly. If $K \not\subset L$, then there exists an index $k_0 \in K \setminus L$. The coordinate x_{k_0} of each point in the box $B^1_K \cap B^2_L$ satisfies the inequalities $b \leqslant x_{k_0} \leqslant a$ and $b \leqslant x_{k_0} \leqslant b$, implying that $x_{k_0} = b$. Therefore the projection of the box $B^1_K \cap B^2_L$ on the k_0-th coordinate axis consists of a single point, implying that the volume of $B^1_K \cap B^2_L$ is zero, so B^1_K and B^2_L are weakly disjoint and $B^1_K \not\subset B^2_L$.

https://doi.org/10.1017/mag.2024.111 Published online by Cambridge University Press

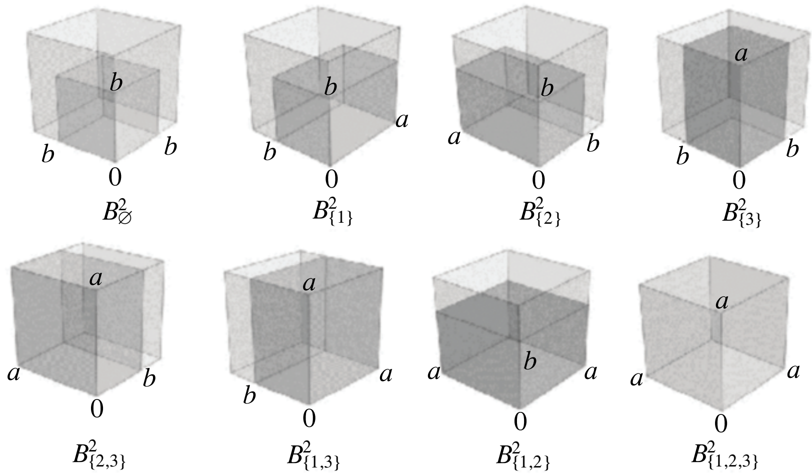

FIGURE 4: Boxes B^2_L of the second kind ($n = 3$)

Denote the volume of the box B^1_K by V^1_K and the volume of the box B^2_L by V^2_L.

The case $n = 3$: For $n = 3$, we use Figure 4 and apply the inclusion-exclusion principle, see [2], to evaluate the volume $V^1_{\{1,2,3\}}$ as follows:

$$V^1_{\{1,2,3\}} = V^2_{\{1,2,3\}} - V^2_{\{1,2\}} - V^2_{\{1,3\}} - V^2_{\{2,3\}} + V^2_{\{1\}} + V^2_{\{2\}} + V^2_{\{3\}} - V^2_{\varnothing}. \quad (3)$$

Calculating the volumes in the above equation yields identity (2) for $n = 3$:

$$(a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3 = \sum_{i=0}^{3} (-1)^i \binom{3}{i} a^{3-i} b^i.$$

Let us now use a slightly different approach – that can readily be employed in higher dimensions – to establish decomposition (3).

Counting contributions of boxes of the first kind approach: Since each box of the second kind is a weakly disjoint union of some boxes of the first kind, we can rewrite the right-hand side

$$E = V^2_{\{1,2,3\}} - V^2_{\{1,2\}} - V^2_{\{1,3\}} - V^2_{\{2,3\}} + V^2_{\{1\}} + V^2_{\{2\}} + V^2_{\{3\}} - V^2_{\varnothing}$$

of (3) in terms of the volumes of boxes of first kind.

Using the Lemma, we observe that in E, the volume of the box $B^1_{\{1,2,3\}}$ is counted once, the volumes of the boxes $B^1_{\{1,2\}}$, $B^1_{\{1,3\}}$ and $B^1_{\{2,3\}}$ are counted $1 - 1 = 0$ times, the volumes of the boxes $B^1_{\{1\}}$, $B^1_{\{2\}}$ and $B^1_{\{3\}}$ are counted $1 - 2 + 1 = 0$ times, and the volume of the box $B^1_{\varnothing}$ is counted $1 - 3 + 3 - 1 = 0$ times. Therefore $E = V^1_{\{1,2,3\}}$, and (3) follows.

https://doi.org/10.1017/mag.2024.111 Published online by Cambridge University Press

Let us illustrate the above argument by an example:

Example: For the $B^1_{\{1,3\}}$ with volume $V^1_{\{1,3\}}$, we have

$$\begin{array}{ccccccccccccccc} E & = & V^2_{\{1,2,3\}} & - & V^2_{\{1,2\}} & - & V^2_{\{1,3\}} & - & V^2_{\{2,3\}} & + & V^2_{\{1\}} & + & V^2_{\{2\}} & + & V^2_{\{3\}} & - & V^2_{\varnothing} \\ & & \cup & & & & \cup & & & & & & & & & & \\ & & V^1_{\{1,3\}} & & & & V^1_{\{1,3\}} & & & & & & & & & & \end{array}$$

using the Lemma, and hence the volume of the box $B^1_{\{1,3\}}$ is counted $1 - 1 = 0$ times.

Similarly, for the box $B^1_{\{2\}}$:

$$\begin{array}{ccccccccccccccc} E & = & V^2_{\{1,2,3\}} & - & V^2_{\{1,2\}} & - & V^2_{\{1,3\}} & - & V^2_{\{2,3\}} & + & V^2_{\{1\}} & + & V^2_{\{2\}} & + & V^2_{\{3\}} & - & V^2_{\varnothing} \\ & & \cup & & \cup & & & & \cup & & & & \cup & & & & \\ & & V^1_{\{2\}} & & V^1_{\{2\}} & & & & V^1_{\{2\}} & & & & V^1_{\{2\}} & & & & \end{array}$$

hence its volume is counted $1 - 2 + 1 = 0$ times.

The general case: Now we will employ the above counting approach to establish (2). We need the following identity:

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0 \qquad \text{for every } n \in \mathbb{N}. \tag{4}$$

The statement is obvious for $n = 1$. If $n \geqslant 2$, we have

$$\begin{aligned} \sum_{i=0}^{n} (-1)^i \binom{n}{i} &= \binom{n}{0} + \sum_{i=1}^{n-1} (-1)^i \binom{n}{i} + (-1)^n \binom{n}{n} \\ &= \binom{n}{0} + \sum_{i=1}^{n-1} (-1)^i \left(\binom{n-1}{i-1} + \binom{n-1}{i} \right) + (-1)^n \binom{n}{n} \\ &= \binom{n}{0} - \sum_{i=0}^{n-2} (-1)^i \binom{n-1}{i-1} + \sum_{i=1}^{n-1} \binom{n-1}{i} + (-1)^n \binom{n}{n} \\ &= 0, \end{aligned}$$

due to telescoping.

Let us establish the main result of the Article, identity (2).

Theorem

$$(a - b)^n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} a^{n-i} b^i \qquad \text{for } a > b > 0.$$

Proof: Denote by $|S|$ the cardinality of a set S. We say that a box B^2_L of the second kind has the *depth* $|L|$. We will show now that, in general,

$$(-1)^n V^1_{\{1,\ldots,n\}} = \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} V^2_S, \tag{5}$$

from which (2) will easily follow.

https://doi.org/10.1017/mag.2024.111 Published online by Cambridge University Press

Each box of the second kind is a union of boxes of first kind: $B_L^2 = \bigcup_{K \subseteq L} B_K^1$. Indeed, if $K \subseteq L$, then $B_K^1 \subseteq B_L^2$, showing $\bigcup_{K \subseteq L} B_K^1 \subseteq B_L^2$. For the opposite inclusion, take $x \in B_L^2$ and define $K \subseteq L$ to be the set of indices $j \in L$ for which $x_j > b$. Then $x \in B_K^1$. Besides, as noted earlier, the union $\bigcup_{K \subseteq L} B_K^1$ is weakly disjoint.

Thus we can count contributions of the volumes of the boxes of the first kind V_K^1 in

$$E = \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} V_S^2,$$

as previously illustrated in the case $n = 3$. We will show that if the cardinality of $K \in \{1, \ldots n\}$ is $k < n$, then the contribution of V_K^1 in the expression E is zero.

By the Lemma, a box B_K^1 is included in a box B_S^2 if, and only if, $K \subseteq S$ and if $K \not\subseteq S$, then B_K^1 and B_S^2 are weakly disjoint. Therefore, *V_K^1 contributes to V_S^2 if, and only if, $K \subseteq S$.* We count the contributions of V_K^1 to E as follows. V_K^1 does not contribute to any V_S^2 when $|S| < k$. It contributes to V_S^2, where $|S| = k$, if and only if $S = K$. If $|S| = k + i$, where $1 \leqslant i \leqslant n - k$, there are $\binom{n-k}{i}$ boxes of the second kind B_S^2 corresponding to sets $S \supset K$ of cardinality $k + i$. Thus V_K^1 contributes $\binom{n-k}{i}$ times towards the sum $\sum_{|S|=k+i} V_S^2$. By the inclusion-exclusion principle, V_K^1 is counted $(-1)^k \sum_{i=0}^{n-k} (-1)^i \binom{n-k}{i}$ times in the sum E. If $k < n$, then (4) implies that its contribution is zero.

If $k = n$, the contribution to E is $(-1)^n V_{\{1,\ldots,n\}}^1$. Thus, (5) follows.

Now let us count the right-hand side of the expression in (5) in a straightforward way. The volume of each box of the second type of depth j is $a^j b^{n-j}$, the number of boxes of the second kind of depth j is $\binom{n}{j}$, the sum of the volumes of the boxes of the second kind of depth j equals $\binom{n}{j} a^j b^{n-j}$. Therefore, the right-hand side of (5) is

$$\begin{aligned}\sum_{j=0}^{n} (-1)^j \binom{n}{j} a^j b^{n-j} &= (-1)^n \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{n-j} a^j b^{n-j} \\ &= (-1)^n \sum_{j=0}^{n} (-1)^i \binom{n}{i} a^{n-i} b^i,\end{aligned}$$

where we used the property $\binom{n}{j} = \binom{n}{n-j}$ and set $i = n - j$. Since $V_{\{1,\ldots n\}}^1 = (a - b)^n$, the result follows from (5).

https://doi.org/10.1017/mag.2024.111 Published online by Cambridge University Press

Finally, let us show how the binomial identity

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{j} a^{n-i} b^i$$

for any real a, b follows from (1), (2) and (4).

First, it is clear that one can assume without loss of generality that $a, b \neq 0$, and also, due to symmetry, that $a \geqslant b$.

Next, to verify the identity for $a + b \geqslant 0$, it is sufficient to consider these three cases:

1. $a, b > 0$;
2. $a > 0 > b > -a$;
3. $a > 0, b = -a$.

In Case 1, the identity is (1). In Case 2, the identity follows from Theorem (identity (2)) if we replace b with $-b$. And, in Case 3, the identity follows from (4).

At last, if $a + b < 0$, we have $-a + (-b) > 0$, so one can write

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

Acknowledgement

The authors would like to thank the anonymous referee for the valuable comments and suggestions that helped us to improve the presentation of the paper.

References

1. Thomas F. Banchoff, *Beyond the third dimension*, Scientific American Library (1996).
2. Inclusion–exclusion principle, *Wikipedia*, accessed March 2024 at https://en.wikipedia.org/wiki/Inclusion-exclusion principle

10.1017/mag.2024.111
Published by
Cambridge University Press
on behalf of
The Mathematical Association

FRANTIŠEK MARKO
SEMYON LITVINOV
The Pennsylvania State University,
76 University Drive, Hazleton,
PA 18202, USA
e-mails: *fxm13@psu.edu,snl2@psu.edu*