Kathy O'Hara's Constructive Proof of the Unimodality of the Gaussian Polynomials

DORON ZEILBERGER*, *Drexel University*

DORON ZEILBERGER received his Ph.D. degree from the Weizmann Institute of Science in 1976, under the direction of Harry Dym. He has been at Drexel University since 1983. His interest in combinatorics began when he tried to find combinatorial applications for his results on partial difference equations. He then became enamored with combinatorics for its own sake. Since 1978 he has attempted to find a constructive proof of the unimodality of the Gaussian polynomials.

1. Introduction. Kathy O'Hara has recently [7], [8] solved a long-standing open problem in combinatorics: to find a *direct combinatorial proof* of the unimodality of the coefficients of the Gaussian polynomials. A unimodal sequence $c_0, c_1, \ldots, c_n$ is one that increases up to a point and then decreases: i.e. there is an integer r such that $c_0 \leq c_1 \leq \cdots \leq c_r \geqslant c_{r+1} \geq \cdots \geq c_n$.

Combinatorialists love to prove that counting sequences are unimodal. For instance, the prototype of all unimodal sequences is the sequence of binomial coefficients

$$\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n-1}, \binom{n}{n}. \tag{1}$$

Many famous combinatorial sequences are unimodal: the Stirling numbers of each kind, for instance. But how can we prove that a sequence is unimodal?

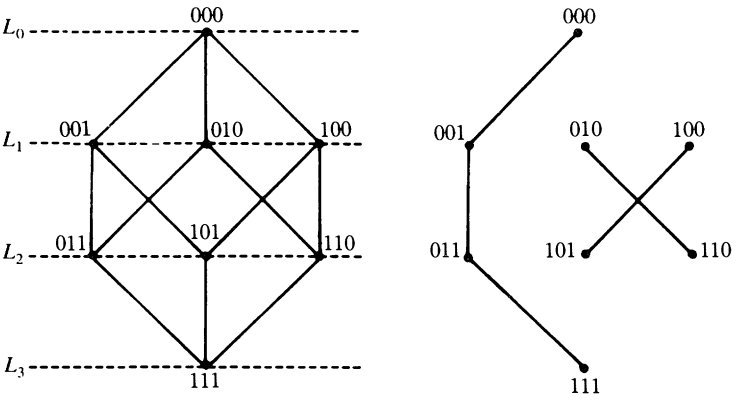
One way is to use analytical methods. If a polynomial has only negative real zeros then its coefficient sequence is unimodal, and even a little more, it is logarithmically convex (see, e.g., [5]). Since the binomial coefficients belong to the polynomial $(1 + x)^n$, which evidently has only negative real zeroes, those coefficients must be unimodal. Since the Stirling numbers of the first kind are the coefficients of the polynomial $(x + 1)(x + 2) \cdots (x + n - 1)$, those numbers must be unimodal too.

Combinatorialists prefer to use combinatorial methods when dealing with such problems, rather than the constructs of analysis. It's often harder, but it's somehow 'purer' too. How might we prove that the binomial coefficients are unimodal in a purely combinatorial way?

Here's one method. Since the coefficients are obviously symmetric about the middle, it will be (necessary and) sufficient to prove that $\left\{\binom{n}{k}\right\}$ increases for $0 \leq k \leq n/2$. To do that, a fine idea would be to produce, for each fixed $k < n/2$, an injection of the k-subsets of $[1, 2, \ldots, n]$ into the $k + 1$-subsets. The reader is invited to try to invent such a family of injections, and will then appreciate how hard these proofs can be.

*Supported in part by the National Science Foundation.

Here's another method, that is in fact the one used in O'Hara's proof, the main subject of this paper. Think of the Boolean lattice B^n of all 2^n of the subsets of $[1, 2, \ldots, n]$, partially ordered by inclusion. Picture the lattice drawn with all $\binom{n}{k}$ of the k-subsets laid out on the same horizontal layer, for each $k = 0, 1, 2, \ldots, n$, as shown in Fig. 1(a).



(a) B^3 (b) An SCD for B^3

FIG. 1

Now the binomial coefficients count the number of subsets in each of the layers. Hence what we have to prove is that as we go upstairs in the building, *the layers get fatter and then thinner.*

Well that's a plan, anyway, but how shall we carry it out? By a *chain* we mean a *linearly* ordered collection of these subsets, e.g.,

$$S_1 \subset S_2 \subset \cdots \subset S_r. \tag{2}$$

A chain (2) is *maximal symmetric* if, as we look along the chain from left to right, the cardinalities of the sets increase by exactly 1, and if $\text{card}(S_1) + \text{card}(S_r) = n$.

In Fig. 1(b), maximal symmetric chains are shown as jagged lines that meet every layer from layer q to layer $n - q$, for some q.

Suppose we can *cover* the Boolean lattice with pairwise disjoint maximal symmetric chains (such a covering of B^3 is shown in Fig. 1(b)). Then we would have proved the unimodality of the coefficient sequence. Indeed, we claim that we would then have a family of injections from the k-subsets to the $k + 1$ subsets, for each $k < n/2$. For, given a k-subset S: map it to the $k + 1$-subset S' that is its successor in the unique maximal chain to which S belongs.

That leaves us with 'merely' the problem of finding a covering of the Boolean lattice with symmetric chains. The reader is invited to think about that one too, and thereby to re-appreciate how hard these proofs can be.

The proof of O'Hara that I want to describe to you is one that results in proving that a certain famous sequence, one that is much more delicate than the binomial coefficients, is unimodal. Hers is the first combinatorial proof of this result, although various non-elementary (i.e., analytic or algebraic) proofs have been

known for some time. The sequence in question (to be described shortly) is touchy: it stays constant for a while, then increases by 1, then stays constant for a while, etc. It is, so to speak, just barely unimodal.

What she did was first to invent a partially ordered set in which the members of the given sequence count the numbers of inhabitants in the layers, and then to invent a maximal symmetric chain decomposition of that partially ordered set. The reader who tries both of our suggested exercises above will appreciate that that was no mean feat.

Specifically, she proved that the polynomial in q ('Gaussian polynomial')

$$G(b, a) = \frac{(1 - q^{a+b})(1 - q^{a+b-1}) \cdots (1 - q^{a+1})}{(1 - q^{b})(1 - q^{b-1}) \cdots (1 - q)} \tag{3}$$

is unimodal, i.e., its coefficient sequence is unimodal.

For example, $G(2,2) = 1 + q + 2q^2 + q^3 + q^4$, is clearly unimodal.

The best way to learn a topic is by teaching it. Similarly the best way to understand a new proof is by writing an expository paper about it. That was the original motivation of this paper. Once written, I thought it would be nice if the readers of the Monthly had the opportunity to savor the elegant proof. All the central ideas and constructions are O'Hara's, but I have made a few improvements and shortcuts that I believe will make the argument clearer. I would like to thank Kathy O'Hara for stimulating conversations and correspondence.

The unimodality of the Gaussian polynomials had previously received several 'fancy' proofs, the first one by Sylvester [**15**]. The most elementary proof before O'Hara's was that of Proctor [**10**] who used only linear algebra. The reader is urged to look up Proctor's beautiful paper [**10**] for the history and the significance of the problem. I should also mention White's [**17**] elegant proof that uses Pólya theory.

The coefficients of the Gaussian polynomials $G(b, a)$ have well known combinatorial interpretations, and the polynomials themselves have a number of interesting properties. What follows is a list of some of these. The reader who has not encountered these before could do no better than to consult Pólya, Szegö [**9**], and follow the beautiful exercises 60.1 to 60.11 of Chapter 1, and their solutions, after which the properties listed below will all have been proved.

(a) The coefficient of q^k in $G(b, a)$, let us call it $c_k(b, a)$, is the number of 'zigzag walks' in the plane (i.e., each step increases either x or y by one unit), from $(0, 0)$ to (b, a), such that the area under the walk (i.e., between the path, the x-axis, and the line $x = a$) is exactly k.

(b) $c_k(b, a)$ is the number of 'p-vectors' in the set

$$U_k(b, a) = \{(p_1, p_2, \ldots, p_a) : 0 \le p_1 \le p_2 \le \cdots \le p_a \le b,\ p_1 + \cdots + p_a = k\} \tag{4}$$

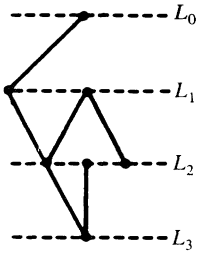
(c) $G(b, a)$ is a polynomial in q of degree ab, whose coefficients are symmetric about the middle, i.e., $c_k = c_{ab-k}$.

In order to prove that $G(b, a)$ is unimodal it will be enough to show that $|U_k(b, a)| \le |U_{k+1}(b, a)|$ for $0 \le k < ab/2$. A *direct combinatorial proof* will consist of exhibiting an explicit injection of $U_k(b, a)$ to $U_{k+1}(b, a)$ for those values of k.

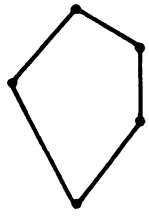
2. Posets. We will need some standard definitions and elementary results from the theory of partially ordered sets (*posets*). All that we will need will be presented here, but the reader who wishes to know more is referred to chapter 2 of Stanton and White's excellent text [**14**], and to Greene and Kleitman's survey [**4**].

A poset $(P, \leq)$ is a set P together with a partial order. We say that $a < b$ if $a \leq b$ and $a \neq b$, and that b *covers* a if $a < b$ and there is no c such that $a < c < b$. If b covers a we will write $a \to b$. A poset is *ranked* if each element $a \in P$ has been assigned a positive integer, $\operatorname{rank}(a)$, such that $a \to b$ implies that $\operatorname{rank}(b) = \operatorname{rank}(a) + 1$. The set $L_i = \{a \in P: \operatorname{rank}(a) = i\}$ is called the *set of rank* i or the ith *level set*.

A good way to visualize a poset P is in terms of its *Hasse diagram*. The vertices of the Hasse diagram are the elements of P and its edges are the covering relations of P. The Hasse diagram of a poset uniquely determines that poset, by transitivity of the order relation. Figs. 2(a) and 2(b) show examples of Hasse diagrams of a ranked and an unranked poset respectively.



(a) a ranked poset



(b) an unranked poset

FIG. 2

A finite ranked poset is *rank unimodal* if the numbers $w_i = |L_i|$ are unimodal. P is *rank symmetric* if there is an integer m such that $w_i = 0$ for $i > m$ and $w_i = w_{m-i}$ for all i. We will call that integer m the rank of P.

A *maximal symmetric chain* in a rank symmetric poset is a sequence of elements of P, $a_1 \to a_2 \cdots \to a_r$ in which $\operatorname{rank}(a_1) + \operatorname{rank}(a_r) = \operatorname{rank}(P)$. A *maximal symmetric chain decomposition* (SCD) of a rank symmetric poset P is a covering of P by pairwise disjoint maximal symmetric chains.

If a poset has an SCD then it is obviously rank unimodal, and so here is the strategy of the proof.

Strategy. Define the poset $U(b, a)$ as follows. Its elements are the p-vectors in the set

$$U(b, a) \overset{\text{def}}{=} \bigcup_k U_k(b, a)$$
$$= \{(p_1, \ldots, p_a): 0 \leq p_1 \leq \cdots \leq p_a \leq b\}.$$

The partial order relation is simply that $p' < p''$ means that the sum of the entries of p' is less than the sum of the entries of p''. We wish to construct an SCD for this poset.

To do this we will need the following three simple facts.

Fact 1. If P is a rank-symmetric poset of rank m and if $\overline{P}$ is the same poset but with the rank defined by

$$\overline{\text{rank}}(p) = \text{rank}(p) + \alpha,$$

for some integer α, then $\overline{P}$ is a rank-symmetric poset of rank $m + 2\alpha$.

Indeed, $w_i = \overline{w}_{\alpha+i}$ and $w_i = w_{m-i}$ imply that $\overline{w}_{\alpha+i} = \overline{w}_{m-i+\alpha}$.

We say that $\overline{P}$ is the *shift* of P by α.

Fact 2. Let P and Q be ranked posets and let π: $P \to Q$ be a rank preserving, order preserving bijection from P to Q. If P has an SCD then so does Q: its chains are the images, under π, of the chains of the SCD of P.

If P and Q are posets, then their *Cartesian product* $(P \times Q, \leqslant)$ is the poset whose elements are those of the set $P \times Q$, and in which the order relation is defined by $(p, q) \leq (p', q')$ iff $p \leq p'$ and $q \leq q'$. If P and Q are ranked, then $P \times Q$ can be ranked by defining $\text{rank}(p, q) = \text{rank}(p) + \text{rank}(q)$.

Fact 3. ([3]) If P and Q are rank-symmetric ranked posets, of ranks m, m' respectively, and if P and Q have SCD's, then $P \times Q$ also has an SCD, and it can be constructed quite explicitly from the two given SCD's. Furthermore, $\text{rank}(P \times Q) = \text{rank}(P) + \text{rank}(Q)$.

The best way to think about the proof of Fact 3 is first to picture the cartesian product of two chains C and D in rectangular form

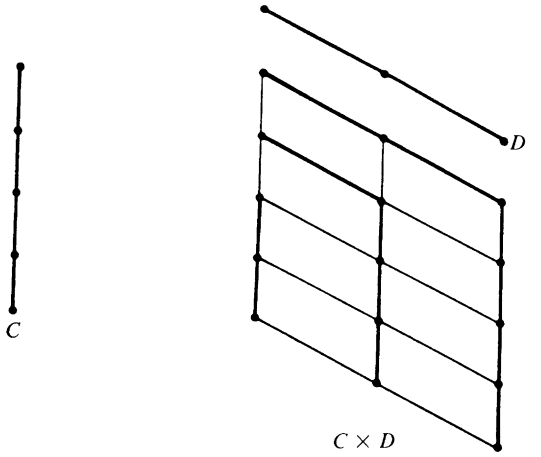
$$\begin{array}{cccc} (p_1, q_1) & (p_2, q_1) & \dots & (p_r, q_1) \\ (p_1, q_2) & (p_2, q_2) & \dots & (p_r, q_2) \\ \dots & \dots & \dots & \dots \\ (p_1, q_s) & (p_2, q_s) & \dots & (p_r, q_s). \end{array}$$

The chains in the SCD of $P \times Q$ that are contributed by just this one chain C from the SCD of P and the one chain D from the SCD of Q are the ones that are obtained by successively 'peeling off' the chain that is obtained by going from left to right along the top row and continuing all the way down the rightmost column, as shown in Fig. 3.

3. The art of constructing symmetric chain decompositions. How does one go about constructing SCD's for a family of posets such as $U(b, a)$? One possibility might have been to express $U(b, a)$ as the cartesian product of smaller $U(b', a')$. For example, for the Boolean lattice B^n, we have $B^n = \{0,1\} \times B^{n-1}$, and by applying Fact 3 recursively we obtain an SCD for B^n ([3]). This does not seem possible for $U(b, a)$.

O'Hara's ingenuity consisted in *dividing and conquering*: she found a refinement of $U(b, a)$, a certain doubly indexed family of subposets, $U(b, a; m, d)$, that appeared to be also rank unimodal, rank symmetric, and *of the same rank*, ab, as $U(b, a)$. She then discovered a structure theorem for these smaller subposets, that expressed each of them in terms of the operations 'union' and 'Cartesian product' of posets $U(b', a'; m', d')$ for $a' < a$.

It was therefore possible to use the three Facts in order recursively to construct an SCD for each of these $U(b, a; m, d)$ in terms of those of smaller a. The base cases $a = -1$, $a = 0$ and $a = 1$ being trivial, this showed that each of the



Decomposition of a product of two chains into chains.

FIG. 3

$U(b, a; m, d)$ had an SCD, and by taking the union over all m and d, one had an SCD for $U(b, a)$ itself. We will now present the details.

4. O'Hara's construction. Recall that

$$U(b, a) = \{ p = (p_1, \ldots, p_a); 0 \le p_1 \le \cdots \le p_a \le b \}.$$

Let us make the convention henceforth that $p_i = 0$ for $i \le 0$ and $p_i = b$ for $i > a$. For each p-vector in $U(b, a)$ we define its *spread* and its *degree*, as follows. First,

$$spread(p) = \max\{ p_i - p_{i-2} : 2 \le i \le a + 1 \}.$$

Next we define $M(p)$ to be the set of indices on which the spread is attained,

$$M(p) = \{ 2 \le i \le a + 1 : p_i - p_{i-2} = spread(p) \}.$$

Partition $M(p)$ into $\cup D_j$, where the D_j are maximal *intervals* (of consecutive integers), and define the degree of p by

$$deg(p) = \sum_j \left\lfloor \frac{|D_j| + 1}{2} \right\rfloor.$$

For example, for $a = 13$, $b = 10$ and $p = (1, 2, 2, 3, 4, 5, 6, 6, 6, 7, 8, 8, 10)$, we find that $spread(p) = 2$, $M(p) = \{2, 5, 6, 7, 11, 13, 14\}$, $D_1 = \{2\}$, $D_2 = \{5, 6, 7\}$, $D_3 = \{11\}$, $D_4 = \{13, 14\}$, and $deg(p) = 5$.

Next we define the following subposets:

$$U(b, a; m, d) = \{ p \in U(b, a) : spread(p) = m \text{ and } deg(p) = d \},$$

and

$$\overline{U}(b, a; m) = \{ p \in U(b, a) : spread(p) \le m \}.$$

We make the reasonable convention that $U(b, 0)$ contains a single p-vector, namely the empty vector. If a is negative then $U(b, a) = \varnothing$. Further, we make the

convention that a Cartesian product $\varnothing \times P$, of the empty set with a poset P, is isomorphic to P, and its elements will be denoted by $(-, p)$.

Everything would follow from

THE O'HARA STRUCTURE THEOREM. *Let a, b, m, d be positive integers. The mapping*

$$\sigma\colon \overline{U}(b - md, a - 2d; m - 1) \times U(ma + 2m - 2b, d) \to U(b, a; m, d),$$

to be defined below, is an order preserving bijection such that for every $q \in \overline{U}(b - md, a - 2d; m - 1)$ and $r \in U(ma + 2m - 2b, d)$, we have

$$rank(\sigma(q, r)) = rank(q, r) + 2bd - md(d + 1). \tag{5}$$

Before we go ahead and define σ, let us show that indeed 'everything' would follow from the theorem.

We prove, by induction on $a = 1, 2, \ldots$, the following proposition: 'For all b, m, d, the poset $U(b, a; m, d)$ has an SCD and has rank ba.' The base case $a = 1$ being trivial to verify, let $a > 1$ be fixed, and let us assume the truth of the proposition for all $1 \le a' < a$.

We claim first that for every $p \in U(b, a)$ we have $deg(p) \le \lfloor (a + 1)/2 \rfloor$. Indeed, for every $S \subset \{2, \ldots, a + 1\}$, if we write $S = \cup D_j$, where the D_j are intervals, it is easy to check, by induction on a, that

$$\sum_j \lfloor (|D_j| + 1)/2 \rfloor \le \lfloor (a + 1)/2 \rfloor,$$

which proves the claim.

Now, if $U(b, a; m, d) \neq \varnothing$ we must have $d \le \lfloor (a + 1)/2 \rfloor$, and since $a > 1$, we have $d < a$. Since $U(ma + 2m - 2b, d)$ is a union of certain $U(ma + 2m - 2b, d; m', d')$, each of which, inductively, has rank $(ma + 2m - 2b)d$ and we know how to construct an SCD for it, it follows that $U(ma + 2m - 2b, d)$ has rank $(ma + 2m - 2b)d$ and we know how to construct an SCD for it. Likewise, $\overline{U}(b - md, a - 2d; m - 1)$ has rank $(b - md)(a - 2d)$ and we know how to construct an SCD for it.

Hence, by Fact 3 we can construct an SCD for the product $\overline{U}(b - md, a - 2d; m - 1) \times U(ma + 2m2b, d)$, and its rank is $(b - md)(a - 2d) + (ma + 2m - 2b)d$. The O'Hara Structure Theorem tells us that $U(b, a; m, d)$ is the image of that product poset under an order preserving bijection, which is actually rank preserving if $U(ma + 2m - 2b, d)$ is shifted by $2bd - md(d + 1)$. By Fact 2 we know how to construct an SCD for $U(b, a; m, d)$, and by Fact 3 its rank is $(b - md)(a - 2d) + (ma + 2m - 2b)d + 2(2bd - md(d + 1)) = ab$.

In order to prove the structure theorem we have to

(i) define the mapping σ
(ii) define its alleged inverse π
(iii) prove that σ is well defined
(iv) prove that π is well defined
(v) prove that the composition $\pi\sigma$ is the identity
(vi) prove that $\sigma\pi$ is the identity
(vii) prove that (5) is true
(viii) prove that σ is order preserving.

In order to make this exposition entirely self-contained, we have included complete proofs of these eight propositions in the Appendix. The proof of the unimodality of the Gaussian coefficients is now finished.

5. Discussion. The partial ordering that we put on the set of p-vectors was the right one to get this proof done, but another partial ordering on the same set of objects has been known for some time, and it is even less well understood. In the set $U(b, a)$ of all p-vectors of $\leq a$ components each of which is $\leq b$, a natural order to consider is the one in which $p \leq q$ means that $\forall i$: $p_i \leq q_i$. Under this order $U(b,a)$ is called Young's lattice. More generally, given a partition λ, the Young lattice Y_λ is defined to be the set of all p-vectors such that $p \leq \lambda$ with the same definition of $\leq$ and rank as before. The $U(b, a)$ are the special cases in which $\lambda = b^a$ (b repeated a times). $U(b, a)$ is rank unimodal; but what about Y_λ for general λ? Of course Y_λ is no longer rank-symmetric, but it appears to be rank-unimodal for small λ. For example, in Y_{12}, the numbers of elements of each rank are $1, 1, 2, 1$.

It was conjectured for some time that the Y_λ, while not rank-symmetric, are nevertheless rank unimodal. It came as a great surprise when Dennis Stanton [**13**], assisted by a computer, found a counterexample: for $\lambda = (4, 4, 8, 8)$, Y_λ is *not* rank unimodal. Furthermore Stanton did something that no computer can do by itself: he found *infinite* families of counterexamples, the simplest one being $(4, 4, 2k, 2k)$, for $k \geq 4$.

A poset that possesses an SCD automatically enjoys the 'Sperner property,' which is to say that its largest possible independent set is obtained by taking the largest level set.

Even the *existence* of an SCD for $U(b, a)$, with Young's lattice order, is still an open problem. However to deduce the Sperner property, as well as rank unimodality, you don't quite need a *symmetric* chain decomposition. It is enough to have a maximal chain decomposition all of whose chains pass through the largest level set. Stanley [**12**], using the 'hard Lefschetz theorem' from algebraic geometry, proved the *existence* of such a chain decomposition for the poset $U(b, a)$ as well as for some other posets, among them the poset $M(n)$ of all partitions of integers into distinct parts $\leq n$, with the same rank and order relation as for $U(b, a)$.

Stanley's proofs for $U(b, a)$ and $M(n)$ were subsequently simplified by Proctor [**10**]. Incidentally, the fact that $M(n)$ has the Sperner property, that Stanley discovered almost as an afterthought, trying to apply the 'hard Lefschetz' hammer to as many nails as possible, turned out to be equivalent to a longstanding conjecture of Erdös and Moser. Stanley found out about it quite by accident, from Larry Harper, during a phone call whose original purpose was to discuss a house sublet during a Sabbatical leave! See Proctor [**10**] for more about this story.

However, both Stanley's proof and Proctor's simplifications were nonconstructive existence proofs. It is still an open problem to find an explicit construction of an SCD for $U(b, a)$ (or even that of a nonsymmetric decomposition as above). To date such a construction is known only when a or b are either 3 or 4 (see Lindstrom [**6**], West [**16**], Riess [**10**]).

I am sure that O'Hara's breakthrough will lead to further work in this area. It would be very nice to find another decomposition of $U(b, a)$ in which the analogue of σ would be also order preserving in the Young's lattice order. It would also be

interesting to find an O'Hara-style constructive proof of Stanley's result that the posets $M(n)$ are rank unimodal. This fact is equivalent to the unimodality of the polynomial $(1+t)(1+t^2)\cdots(1+t^n)$.

More generally, Almkvist [1] conjectured that the polynomials

$$\prod_{\nu=1}^{n} \frac{1-t^{r\nu}}{1-t^{\nu}}$$

are unimodal when r is even and $n \geq 1$ and when r is odd and $n \geq 11$. He developed an interesting analytical method that is capable, at least in principle, of proving this conjecture for every *specific* r. Assisted by computer, he proved his conjecture for $3 \leq r \leq 20$ and $r = 100, 101$. The conjecture is still open for general r. It is equivalent to the rank unimodality of the poset of partitions in which each part can appear at most $r-1$ times, and whose parts are all $\leq n$. An O'Hara-style proof would be particularly gratifying because it would demonstrate that purely bijective methods are capable not only of duplicating results that have been found by other branches of mathematics, but of proving *new* ones.

Appendix. The fine print (proofs of (i)–(viii)). It is easiest to start with the definition of π. So we will perform the eight tasks in the following order: (ii), (i), (iv), (iii), (v), (vi), (vii), (viii).

(*ii*) *Definition of* π.

$$U(b, a; m, d) \to \bar{U}(b - md, a - 2d; m-1) \times U(ma + 2m - 2b, d)$$

Let $p = (p_1, \ldots, p_a) \in U(b, a; m, d)$. We will define $\pi(p) = (q(p), r(p)) = (q, r)$. Let $t+1 = \max M(p) = \max\{2 \leq i \leq a+1;\ p_i - p_{i-2} = m\}$, and define a new p-vector $p' = (p'_1, \ldots, p'_{a-2})$ by putting $p'_i = p_i$ if $1 \leq i \leq t-2$ and $p'_i = p_{i+2} - m$ else. We will show that $p' \in U(b-m, a-2; m, d-1)$ when $d > 1$ and $p' \in \bar{U}(b-m, a-2; m-1)$ when $d = 1$.

Define $q(p) = p'$, if $d = 1$, and, recursively, $q(p) = q(p')$, if $d > 1$. Also, let $r' = \varnothing$ if $d = 1$, and, recursively, $r' = r(p')$ if $d > 1$ (we will show that r' has $d-1$ nonvanishing components). Finally, put $r_1 = p_t + p_{t+1} + m(a-t-1) - 2b$, and $r_i = r'_{i-1}$ for $2 \leq i \leq d$.

(i) *Definition of* σ. $\bar{U}(b - md, a - 2d; m-1) \times U(ma + 2m - 2b, d) \to U(b, a; m, d)$

Let $q \in \bar{U}(b - md, a - 2d; m-1)$ and $r \in U(ma + 2m - 2b, d)$; we will define $\sigma(q, r) = p$, say.

Let r' be the p-vector with $d-1$ components obtained from r by deleting the first component. Obviously $r' \in U(ma + 2m - 2b, d-1)$. Next let $p' = q$ if $d = 1$, and $p' = \sigma(q, r')$ if $d > 1$ (we will show that p' has $a-2$ nonvanishing entries).

We will now define a certain integer t, $1 \leq t \leq a$. If $r_1 = 0$, let $t = a$. Else, let t be the unique integer that satisfies

$$p'_{t-1} + p'_t - mt < r_1 - m(a+2) + 2b \leq p'_{t-2} + p'_{t-1} - m(t-1). \quad \text{(A1)}$$

We define $p = \sigma(q, r)$ by

$$p_i = \begin{cases} p'_i, & 1 \le i \le t-1 \\ r_1 - p'_{t-1} - m(a-t+2) + 2b, & i = t \\ p'_{i-2} + m, & t+1 \le i \le a. \end{cases}$$

(iv) π is well defined.

p' is a genuine p-vector since $p'_{t-1} = p_{t+1} - m = p_{t-1} \ge p_{t-2} = p'_{t-2}$. Here we have used the fact that $t+1 \in M(p)$, and thus $p_{t+1} - m = p_{t-1}$. Now it is readily seen that $M(p') = M(p) \setminus \{t+1\}$ if $t \notin M(p)$, and $M(p') = M(p) \setminus \{t, t+1\}$ if $t \in M(p)$. In either case, $deg(p') = deg(p) - 1$ and $spread(p') = m$, if $d > 1$, while $spread(p') < m$ if $d = 1$.

Hence $p' \in U(b-m, a-2; m, d-1)$ if $d > 1$ and $p' \in \overline{U}(b-m, a-2; m-1)$ if $d = 1$. If $d = 1$, clearly $q \in \overline{U}(b - md, a - 2d; m-1)$, and if $d > 1$, then inductively, q belongs to $\overline{U}(b - m - m(d-1), a - 2 - 2(d-1); m-1) = \overline{U}(b - md, a - 2d; m-1)$. Let $t' + 1 = \max M(p')$. Of course, $t' \le t - 2$. It remains to show that r is a genuine p-vector in $U(m(a+2) - 2b, d)$.

By the inductive hypothesis r' belongs to $U(m(a-2+2) - 2(b-m), d-1) = U(m(a+2) - 2b, d-1)$, and so to show that r is a bona fide element of $U(m(a+2) - 2b, d)$ we will have to show that when $d > 1$ we have $0 \le r_1 \le r_2$, i.e., $0 \le r_1 \le r'_1$, i.e., that

$$0 \le p_t + p_{t+1} + m(a-t+1) - 2b \le p'_{t'} + p'_{t'+1} + m((a-2) - t' + 1) - 2(b-m), \tag{A2}$$

and when $d = 1$ we have to show that $r_1 \le m(a+2) - 2b$, i.e.,

$$0 \le p_t + p_{t+1} + m(a-t+1) - 2b \le m(a+2) - 2b.$$

If we make the convention that $t' = -1$ when $d = 1$, then both of the above are included if we prove (A2) for all $d \ge 1$. Now the left side of (A2) is equivalent to $(b-m) + b - p_t - p_{t+1} \le m(a-t)$, while the right side of (A2) is equivalent to $p_t + p_{t+1} - p'_{t'} - p'_{t'+1} \le m(t-t')$. Both of these follow from the proposition below: the first by taking $A = a$, $B = t$, and noting that $p_{a+1} = b$ and $p_a \le b - m$; the second by taking $A = t$, $B = t'$, and noting that since $t' \le t - 2$, we have $p'_{t'} = p_{t'}$ and $p'_{t'+1} = p_{t'+1}$.

PROPOSITION. *Let $A \ge B$. Then $p_A + p_{A+1} - p_B - p_{B+1} \le m(A-B)$.*

Proof of proposition. If $A - B$ is even then

$$\begin{aligned} p_A + p_{A+1} - p_B - p_{B+1} &= [p_A - p_B] + [p_{A+1} - p_{B+1}] \\ &= [(p_A - p_{A-2}) + \cdots + (p_{B+2} - p_B)] \\ &\quad + [(p_{A+1} - p_{A-1}) + \cdots + (p_{B+3} - p_{B+1}) \\ &\le m(A-B)/2 + m(A-B)/2 = m(A-B). \end{aligned}$$

Similarly, if $A - B$ is odd,

$$\begin{aligned} p_A + p_{A+1} - p_B - p_{B+1} &= [p_A - p_{B+1}] + [p_{A+1} - p_B] \\ &= [(p_A - p_{A-2}) + \cdots + (p_{B+3} - p_{B+1})] \\ &\quad + [(p_{A+1} - p_{A-1}) + \cdots + (p_{B+2} - p_B)] \\ &\le m(A - B - 1)/2 + m(A + 1 - B)/2 = m(A - B). \end{aligned}$$

(iii) σ is well defined.

Let $q \in \overline{U}(b - md, a - 2d; m - 1)$ and $r \in U(m(a + 2) - 2b, d)$. We will prove that $p = \sigma(q, r)$ is a p-vector in $U(b, a; m, d)$.

Since $r' \in U(m(a + 2) - 2b, d - 1) = U(ma - 2(b - m), d - 1)$ and $q \in \overline{U}(b - md, a - 2d; m - 1) = \overline{U}((b - m) - m(d - 1), (a - 2) - 2(d - 1); m - 1)$, it follows by the induction hypothesis, for $d > 1$, that $p' \in U(b - m, a - 2; m, d - 1)$. If $d = 1$, then of course $p' \in \overline{U}(b - m, a - 2; m - 1)$.

Now we claim that $p'_{a-1} + p'_a - ma \le r_1 - m(a + 2) + 2b \le p'_{-1} + p'_0 - m(1 - 1)$, which, since $p'_{-1} = p'_0 = 0$ and $p'_{a-1} = p'_a = b - m$, is equivalent to $2b - m(a + 2) \le r_1 - m(a + 2) + 2b \le 0$. That is equivalent to $0 \le r_1 \le m(a + 2) - 2b$, which is obvious.

Now it is easy to see that the closed (discrete) interval $[2b - m(a + 2), 0]$ can be partitioned as follows into a union of a single point and half open intervals:

$$\begin{aligned} [2b - m(a + 2), 0] &= \{2b - m(a + 2)\} \\ &\quad \cup \bigcup_{t=a}^{t=1} [p'_{t-1} + p'_t - mt, p'_{t-2} + p'_{t-1} - m(t - 1)] \end{aligned}$$

Hence the t that was defined in the definition of σ was in fact well defined.

We must also show that $p_t \ge p_{t-1}$ and $p_{t+1} \ge p_t$, both of which follow easily from (A1). Furthermore, $p_a = p'_{a-2} + m \le b - m + m = b$, so $p \in U(b, a)$. Also $t + 1 \in M(p)$.

Now since $r_1 - m(a + 2) + 2b \le r_2 - m(a - 2 + 2) + 2(b - m)$, the 'previous t,' let us call it t', that was obtained in the previous step out of r_2, satisfies $t' \le t - 2$ (if $d = 1$ then $t' = -1$). Thus $M(p) = M(p') \cup \{t + 1\}$ or $M(p) = M(p') \cup \{t, t + 1\}$ (if $d = 1$, $M(p') = \varnothing$). If $d > 1$, then $deg(p) = deg(p') + 1 = d - 1 + 1 = d$. Of course, $spread(p) = m$, and if $d = 1$, $deg(p) = 1$, since $M(p)$ consists of either $\{t\}$ or $\{t, t + 1\}$, so in either case $p \in U(b, a; m, d)$.

(vi) $\pi\sigma$ is the identity.

Let $p = \sigma(q, r)$. We have to show that $\pi(p) = (q, r)$. We have just seen that $M(p)$ is obtained from $M(p')$ by adjoining either t alone, or else both t and $t + 1$. In either case $t + 1 = \max M(p)$, so the 't obtained in doing $\pi(p)$' is the same as 'the t obtained in doing $\sigma(q, r)$.' By the inductive hypothesis, if $p' = \sigma(q, r')$ then $\pi(p') = (q, r')$. Finally, 'the r_1 obtained from doing $\pi(p)$' is

$$\begin{aligned} &p_t + p_{t+1} + m(a - t + 1) - 2b \\ &\quad = r_1 - p'_{t-1} - m(a - t + 2) + 2b + p'_{t-1} + m + m(a - t + 1) - 2b \\ &\quad = r_1, \end{aligned}$$

as it should.

(v) $\sigma\pi$ is the identity.

Let $(q, r) = \pi(p)$. We have to show that $\sigma(q, r) = p$.

We have $p_t + p_{t+1} = r_1 - m(a - t + 1) + 2b$. But $p_{t+1} = p_{t-1} + m$ since $t + 1 \in M(p)$. So $p_t = r_1 - m(a - t + 2) + 2b - p_{t-1}$. Now t may or may not be in $M(p)$, thus $p_t - p_{t-2} \leq m$, which yields the right side of (A2), and $t + 2 \notin M(p)$, so $p_{t+2} - p_t < m$, which yields the left side of (A2). Thus 'the t obtained in doing $\sigma(\pi(p))$' is the same as 'the t obtained in doing $\pi(p)$.' The rest follows from the inductive hypothesis.

(vii) *Proof* that $\operatorname{rank} \sigma(q, r) = \operatorname{rank}(q, r) + 2bd - md - md^2$.

We have

$$\begin{aligned}\operatorname{rank} \sigma(q, r) &= \operatorname{rank}(p) = p_1 + \cdots + p_a \\ &= p'_1 + \cdots + p'_{t-1} + (r_1 - p'_{t-1} - m(a - t + 2) + 2b) \\ &\quad + (p'_{t-1} + m) + \cdots + (p'_{a-2} + m) \\ &= \operatorname{rank}(p') + r_1 + 2(b - m).\end{aligned}$$

First, if $d = 1$ then $p' = q$ and $r = r_1$, so $\operatorname{rank} \sigma(q, r) = \operatorname{rank}(q) + \operatorname{rank}(r) + 2(b - m)$, which was to be shown in this case. Next, if $d > 1$, the inductive hypothesis gives

$$\begin{aligned}\operatorname{rank} \sigma(q, r) &= \operatorname{rank} \sigma(q, r') + r_1 + 2(b - m) \\ &= \operatorname{rank}(q) + \operatorname{rank}(r') + 2(b - m)(d - 1) \\ &\quad - m(d - 1)d + r_1 + 2(b - m) \\ &= \operatorname{rank}(q) + \operatorname{rank}(r) + 2bd - md(d + 1) \\ &= \operatorname{rank}(q, r) + 2bd - md(d + 1).\end{aligned}$$

(viii) σ is order preserving.

This is immediate from part (vii) since the order on $U(b, a)$ is the trivial one in which $p < q$ iff $\operatorname{rank}(p) < \operatorname{rank}(q)$.

REFERENCES

1. Gert Almkvist, Proof of a conjecture about unimodal polynomials, *J. of Number Theory*, to appear.
2. George Andrews, The Theory of Partitions, Addison Wesley, Reading, MA, 1976.
3. N. G. de Bruijn, C. van E. Tenbergen, and D. Kruyswijk, On the set of divisors of a number, *Nieuw. Arch. Wisk.*, (2) 23 (1952) 191–193.
4. Curtis Greene and Daniel Kleitman, Proof techniques in the theory of finite sets, in *Studies in Combinatorics*, G. C. Rota, ed., Mathematical Association of America, 1978, pp. 22–79.
5. G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, Cambridge University Press, 1952.
6. B. Lindstrom, A partition of $L(3, n)$ into saturated chains, *European J. Comb.*, 1 (1981) 60–63.
7. Kathleen M. O'Hara, Unimodality of Gaussian coefficients: a constructive proof, research announcement, preprint, University of Iowa.
8. _____, Unimodality of Gaussian coefficients: a constructive proof, *J. Comb. Theory*, Ser. A, to appear.
9. G. Pólya and G. Szegö, Problems and Theorems in Analysis, Springer-Verlag, New York, 1972.
10. Robert Proctor, Solution of two difficult combinatorial problems with linear algebra, this MONTHLY 89 (1982) 721–734.
11. W. Riess, Zwei optimizierungprobleme auf Ordnungen, *Arbeitsberichte des Instituts für mathematische Maschinen und datenverarbeitungen*, 11 (1978) no. 5 Erlangen.

12. Richard P. Stanley, Weyl groups, the hard Lefschetz theorem, and the Sperner property, *SIAM J. Alg. Discr. Meth.*, 1 (1980) 164–184.
13. Dennis Stanton, Unimodality and Young's lattice, *J. Comb. Th.*, to appear.
14. Dennis Stanton and Dennis White, Constructive Combinatorics, Springer-Verlag, New York, 1986.
15. James J. Sylvester, Proof of the hitherto undemonstrated fundamental theorem of invariants, Collected Math. Papers, vol. 3, Chelsea, New York, 1973, pp. 117–126.
16. Douglas B. West, A symmetric chain decomposition of $L(4, n)$, *European J. Comb.*, 1 (1980) 379–383.
17. Dennis E. White, Monotonicity and unimodality of the pattern inventory, *Adv. Math.*, 38 (1980) 101–108.

LETTERS TO THE EDITOR

Editor:

While I agree with Joseph Fulda (Material implication revisited, this MONTHLY, March, 1989, 247–250) that there is no need to revise the truth table for the conditional connective of propositional logic in order to explain most uses of the words "if" and "then," or "implies," I would like to suggest an alternate explanation for the widespread confusion on this point. Many sentences containing the words "if" and "then" should properly be interpreted as containing not only a conditional, but also an unstated universal quantifier. The belief that the truth table for the conditional cannot be used to explain the meaning of such sentences is, I believe, a result of failing to recognize the presence of this quantifier, and does *not* indicate any defect in the truth table for the conditional. If the truth table for the conditional is properly combined with the meaning of the quantifier, the resulting interpretation of the sentence is in complete agreement with intuition.

For example, we all know that for sequences of real numbers, monotone and bounded implies convergent, although neither monotone nor bounded alone is sufficient to imply convergent. But according to the conditional truth table, the statements $(M \wedge B) \rightarrow C$ and $(M \rightarrow C) \vee (B \rightarrow C)$ are equivalent. (If you don't believe it, make the truth tables!) Does this mean that the conditional cannot be used to interpret the word "implies" in this case? The problem is not with the conditional, but with a missing quantifier. When we say, "Monotone and bounded implies convergent," what we really mean is, "For all sequences s, if s is monotone and bounded then s is convergent." Inserting the missing quantifiers, we find that we should have been comparing the statements $\forall s[(M(s) \wedge B(s)) \rightarrow C(s)]$ and $\forall s(M(s) \rightarrow C(s)) \vee \forall s(B(s) \rightarrow C(s))$. These are not equivalent; in fact, interpreting "M," "B," and "C" as meaning "monotone," "bounded," and "convergent," respectively, on the universe of sequences of real numbers gives a counterexample to the equivalence. I do not see how one can explain the trouble in this example, as Fulda suggests, by claiming that there is a confusion between "if" and "iff." One need not think that all convergent sequences are monotone and bounded to be puzzled by this example.

Even some nonmathematical "paradoxes" involving the conditional can be explained as missing quantifiers. For example, consider the sentence, "If the sun is shining then it must be between 2:00 and 3:00 PM." The conditional truth table