Bijective Principles of Cancellation

(a shortened version of an article
to appear in *Advances in Mathematics*,
as "Producing New Bijections from Old")

March 30, 1992

David Feldman
University of New Hampshire

James Propp
Massachusetts Institute of Technology

ABSTRACT

Combinatorics has many general constructions that produce finite sets from finite sets, such as the Cartesian product and power set constructions. Bijections of the given sets give rise to bijections of the constructed sets in an obvious way. We explore the reverse process, the problem of defining a bijection between the given objects from one between the constructed objects. We formulate a precise interpretation of the question and present a concrete criterion for when such cancellation is or is not possible. We apply the criterion to several standard combinatorial constructions. In particular, we show that a bijection between the mth powers of sets suffices to define a bijection between the sets themselves, but a bijection between power sets generally does not. Our effective cancellation procedures are in some cases computationally infeasible, but we do offer a reasonably explicit procedure, in the spirit of the involution principle of Garsia and Milne, to cancel the construction of forming the Cartesian product of a given finite set with a fixed set that carries a distinguished element. By contrast, our methods allow us to conclude that Cartesian product with a fixed set without extra structure cannot be cancelled.

Key words: Bijection, bijective proof, G-set, Cartesian power, power set

Introduction

Any analytical demonstration of a combinatorial identity compels us to inquire about the existence of a direct demonstration by means of a bijection. While it is conceivable that all true identities will eventually yield to sufficient cleverness, it is also conceivable that this is not the case. To allow for the possibility of settling such questions in the negative, eventually we must formalize the notion of "bijective proof."

Fixing on a good definition seems difficult. Instances of the following sort, however disguised, must surely be excluded: an identity is actually established by analytic means so that the mere existence of a bijection is assured, and by specifying enumerations of the sets on each side of the identity we single one out. Perhaps the bijection itself is to be judged faulty in that it may be inefficient to compute. Perhaps the defect is with the form of the specification of such a bijection, which makes use of the structure of the families of combinatorial objects rather than just the structure of the objects themselves. Or one's complaint may simply be that such a bijection is irrelevant to the proof of the identity.

It is less clear how one should treat instances of the following kind. In the course of showing that two (indexed families of) sets A and B have the same cardinality, a combinatorial construction is performed on each, yielding sets QA and QB and a self-evident bijection between QA and QB is then produced. (We may formalize "combinatorial construction" as a faithful endofunctor on the category of finite sets and bijections. This captures that idea that a construction Q on a set A does not depend on how the elements of A are named.) Assuming it is known that the cardinality of QS determines the cardinality of S, the equicardinality of A and B is proved. This may well be a bijective proof that $|QA| = |QB|$, but is it a bijective proof that $|A| = |B|$?

Sometimes a principle is available that cancels the construction Q, thereby uniformly specifying a bijection between A and B whenever a bijection between QA and QB is given. The involution principle used by Garsia and Milne in their combinatorial proof of the Rogers-Ramanujan identities is a celebrated example, though it must be slightly recast to fall within our rubric. In practice combinatorialists generally accept such proof as "bijective" but the objections raised above still apply. In particular, even when the bijection between QA and QB is computationally efficient the derived bijection between A and B may not be. The specification of the bijection

between A and B has a global character, since one takes account of the entire bijection between QA and QB for its definition. Lastly, the truth the identity is evident as soon as one has the bijection from QA to QB, so the bijection between A and B is not strictly relevant to the proof.

One criterion that separates artificial "combinatorifications" of analytic proofs (which we wish to exclude) from proofs that use the involution principle (which we wish to allow) is that the latter may be described as canonical in that they do not involve arbitrary choices. They are, if you will, symmetrical. Hence symmetry considerations will play a pivotal role in our analysis of combinatorial constructions.

We have no hope of settling the debate here as to what truly constitutes a bijective proof. Sidestepping the thorny issue of the absolute existence of bijective proofs of combinatorial identities, we consider more tractable questions of relative existence. A combinatorial cancellation principle for the construction Q takes a bijection that demonstrates the identity $|QA| = |QB|$ and yields a bijection that demonstrates the identity $|A| = |B|$.

We study questions concerning the existence of bijective cancellation principles. The condition we formulate for the cancellation of a construction F is certainly necessary and sufficient. Our main examples show the construction $QA = A^m$, the m^{th} Cartesian power of A, may be cancelled, but $QA = 2^A$, the power set of A, may not.

Even our negative results may have a positive value for the working combinatorialist. Knowing that the power set construction cannot generally be cancelled suggests the possibility of interesting bijections from 2^A to 2^B even in situations where no bijection from A to B is available.

Our methods come from group theory and are purely mathematical, as opposed to metamathematical. Nevertheless it may be fruitful to interpret our results as statements about combinatorics in various toposes, though we do not pursue this.

The organization of our paper is as follows. In the next section, we give a few preliminaries concerning group actions. Then we dive immediately into considerations that lead us to our Fundamental Lemma and Main Theorem, which give us computable criteria for the existence of cancellation principles. The five sections that follow apply this theorem to obtain various positive and negative results, some new and some old: a positive result for cancellation of disjoint union (old), a negative result for Cartesian products (a folk-observation that, while not in the research literature, seems to surprise few seasoned combinatorists), a positive result for Cartesian products with pointed sets (new), a positive result for Cartesian powers (new), and a generally negative result for power sets (new) which is clarified by the solution an interesting group-theoretical problem. The final section situates our work

in a more abstract, category-theoretic setting.

Preliminaries

A G-set is a set with an action by a group G (on the left). Two G-sets are isomorphic if and only if there is a bijection from one to the other that commutes with the action by G. A G-orbit is a transitive G-set, and a G-set is a disjoint union of G-orbits. The stabilizer Stab x of an element of a G-set is the set of elements g with $gx = x$. (For background on group actions on finite sets, see Chapter 1 in [Jacobson].)

Note that if S is a G-orbit containing an element x, then the action of G on S is isomorphic to the action of G on the left-cosets of Stab x. Furthermore, if x and y are both in the orbit S, Stab x and Stab y are conjugate subgroups. It follows that a G-orbit is determined up to isomorphism by the conjugacy class of the stabilizer of any of its elements.

If a G-orbit contains an element whose stabilizer is H, we say that it is a G-orbit of H-type. Note that for H, H' conjugate in G, an orbit is of H-type if and only if it is of H'-type.

For each group G, the above remarks allow us to formulate a complete numerical isomorphism invariant for G-sets – indeed, several such invariants, each with its own uses. Let $L(G)$ be the lattice of subgroups of G. For any G-set S, we define maps $\sigma_S, \sigma_S^+, \tau_S$ from $L(G)$ to $\mathbf{Z}$ by letting $\sigma_S(H)$ be the number of elements of S whose stabilizer is H, $\sigma_S^+(H)$ be the number of elements of S that are stabilized by H (i.e., whose stabilizer contains H), and $\tau_S(H)$ be the number of G-orbits of S of H-type. Note that all three functions take a single value on each conjugacy class of subgroups of G. For every non-negative integer-valued function $\tau(H)$ that is constant on conjugacy classes of subgroups there exists an associated G-set.

Equivariance Criterion

Our first goal is to fix a way of making questions about bijective cancellation precise. In particular, we must specify carefully that information which is to be available to a cancellation procedure. Input to the procedure will consist solely of:
(i) rosters of the elements of sets A, B, QA and QB;
(ii) the bijection between QA and QB;
(iii) the declaration of any extra structure that the sets A and B might carry, and whatever additional structure QA and QB acquire through the construction Q.
The procedure must not depend on extraneous information implicit in the presentation of the data. It must make no use of the ordering of the elements on the rosters, no use of the specifics of the coding of their names or anything

else, lest the task of producing a bijection between A and B be trivialized. Now, as we shall indicate, these stipulations may be formalized so as to treat the cancellation procedure as a black box. Having no need to specify a particular model of computation, we do not do so.

We will write $\mathrm{Bij}(X,Y)$ for the set of bijections between X and Y. For the moment, let us fix two sets A and B of cardinality n. Let Q be an endofunctor of the category of finite sets and bijective maps, which is to say, a species in the sense of A. Joyal (see [Joyal]). Thus Q takes a finite set A to a finite set QA, takes a bijection $f : A \to B$ to a bijection $Qf : QA \to QB$, and respects composition of composable bijections. We also ask that Q be faithful, so $Qf = Qg$ if and only if $f = g$. Then if a cancellation procedure for Q exists, it will provide a function

$$\mathcal{F}_{A,B} : \mathrm{Bij}(QA, QB) \to \mathrm{Bij}(A, B) .$$

The procedure cannot depend on the names of the elements of A and B, so the function $\mathcal{F}_{A,B}$ must be invariant under relabelings. More exactly: Write $\mathrm{Sym}\ S$ for the symmetric group on a set S. The group $\mathrm{Sym}\ A \times \mathrm{Sym}\ B$ acts on $\mathrm{Bij}(A,B)$ by

$$(\rho_1, \rho_2)f = \rho_2 f \rho_1^{-1}$$

and on $\mathrm{Bij}(QA, QB)$ by

$$(\rho_1, \rho_2)F = \overline{\rho_2} F \overline{\rho_1^{-1}},$$

where $\overline{\rho}$ is shorthand for $Q\rho$. Invariance under relabelings amounts to

$$\mathcal{F}_{A,B}((\rho_1, \rho_2)f) = (\rho_1, \rho_2)\mathcal{F}_{A,B}(f) ,$$

that is, the equivariance of $\mathcal{F}_{A,B}$ with respect to the action of $\mathrm{Sym}\ A \times \mathrm{Sym}\ B$. The existence of such an equivariant map, for each n, is thus a necessary condition for the existence of an effective procedure of the desired kind. (We do not stipulate that $\mathcal{F}_{A,B}$ should be a one-sided inverse to F, but this is an easy consequence of equivariance, by the fundamental lemma below.)

Is it also a sufficient condition? Write $[n]$ for $\{1, \ldots, n\}$. If an equivariant map exists, for any A and B, then certainly there is a map

$$\mathcal{F}_{[n],[n]} : \mathrm{Bij}(Q[n], Q[n]) \to \mathrm{Bij}([n], [n])$$

equivariant under $\mathrm{Sym}\ [n] \times \mathrm{Sym}\ [n]$. (Of course, producing an equivariant $\mathcal{F}_{[n],[n]}$ from $\mathcal{F}_{A,B}$ requires well-ordering A and B; since A and B are finite sets this can certainly be done.) It is easy to define a lexicographic order on the set of *all* maps from $\mathrm{Bij}(Q[n], Q[n])$ to $\mathrm{Bij}([n], [n])$, since $[n]$ comes with a canonical ordering. So we may effectively single out one equivariant

map for each n, namely, the lexicographically earliest one. Let this be done. (Note that in practice this step may be computationally infeasible.)

We are trying to pass from the mere existence of an equivariant $\mathcal{F}_{A,B}$ to an effective implementation by way of $\mathcal{F}_{[n],[n]}$. Suppose we are given a bijection F between QA and QB, where A and B each have cardinality n. Our procedure cannot make use of a particular pair of bijections $a : A \to [n]$ and $b : B \to [n]$, since there is no canonical way choose such. Nevertheless, the procedure may consider the totality of all possible pairs a and b. Each such pair yields bijections $\overline{a} : QA \to Q[n]$ and $\overline{b} : QB \to Q[n]$ from which we obtain the self-bijection of $Q[n]$

$$F' = \overline{b} F \overline{a}^{-1}$$

and the bijection

$$b^{-1}(\mathcal{F}_{[n],[n]}(F'))a$$

between A and B. The equivariance of $\mathcal{F}_{[n],[n]}$ guarantees that the same bijection between A and B will result in every instance. Thus we have a procedure for effectively determining bijection from A to B. We conclude that the existence of an equivariant $\mathcal{F}_{A,B}$ is indeed sufficient.

Some foundational reflection on the last argument may be justified. We are working within the assumptions of classical set theory and logic. Nevertheless, our problem, dealing as it does with sets containing indistinguishable elements, is awkward to formalize in that context. For us, the distinction between a non-empty finite set whose elements are entirely indistinguishable and an empty set is fundamental. The set of well-orderings of a non-empty set of indistinguishable elements may also have no distinguished elements; nevertheless it is non-empty.

Casting the argument somewhat differently may appeal more to some readers. Suppose $\mathcal{F}_{[n],[n]}(F') = f'$. Fix a sufficiently rich first-order language with variables ranging over the elements of $[n]$ (but with no constants for the elements of $[n]$), so that we may characterize the isomorphism type of any F' by a finite predicate. Since F' may have symmetries, some elements of $[n]$ may be indistinguishable by predicates, but equivariance guarantees that we may describe the function f' by predicates nevertheless. Construct such predicates for every pair (F', f'). This family of predicates may now be used to determine when $\mathcal{F}_{A,B}(F) = f$. Thus one never need refer to a well-ordering of A or B. One need only presume a computational model powerful enough to implement the necessary logical procedures.

In view of the above, we adopt the following principle:

Equivariance Criterion: A combinatorial construction Q is bijectively cancellable iff for all A, B there exists a Sym $A \times$ Sym B-equivariant map from Bij(QA, QB) to Bij(A, B).

Main Theorem

We will now concentrate on the question of existence of an equivariant $\mathcal{F}_{A,B}$.

Fundamental Lemma Let G be a (finite) group and S and T be G-sets. (sets with a G action.) Then the following are equivalent:

(i) There is a G-equivariant map h from S to T.

(ii) For every $s \in S$ there is a $t \in T$ such that Stab $s \subseteq$ Stab t.

Proof
(i)⇒(ii): Stab $s \subseteq$ Stab $h(s)$ since $gs = s$ implies $gh(s) = h(gs) = h(s)$.
(ii)⇒(i): Pick a representative $s_{\mathcal{O}}$ of each orbit $\mathcal{O}$ of S. Set $h(s_{\mathcal{O}}) = t_{\mathcal{O}}$ for some element $t_{\mathcal{O}}$ of T satisfying Stab $s_{\mathcal{O}} \subseteq$ Stab $t_{\mathcal{O}}$. Extend h to a map from S to T by sending $h(gs_{\mathcal{O}})$ to $gt_{\mathcal{O}}$. To see that h is well–defined, suppose $g_1 s_{\mathcal{O}} = g_2 s_{\mathcal{O}}$. Then $g_2^{-1} g_1 \in$ Stab $s_{\mathcal{O}}$ and $g_2^{-1} g_1 \in$ Stab $t_{\mathcal{O}}$, that is $g_1 t_{\mathcal{O}} = g_2 t_{\mathcal{O}}$ as desired. To see that h is equivariant, note that for all $g' \in G$, $h(g'(gs_{\mathcal{O}})) = g'gt_{\mathcal{O}} = g'(h(gs_{\mathcal{O}}))$. □

Here is how the Fundamental Lemma is applied.

For $f \in \mathrm{Bij}(A, B)$, Stab f is exactly the group

$$\{ (\rho, f\rho f^{-1}) \mid \rho \in \mathrm{Sym}\ A \} .$$

According to the Fundamental Lemma, there is an equivariant map

$$\mathcal{F}_{A,B} : \mathrm{Bij}(QA, QB) \to \mathrm{Bij}(A, B)$$

if and only if for every $F \in \mathrm{Bij}(QA, QB)$, Stab F is contained in one of the groups Stab f.

Suppose $(\rho_1, \rho_2) \in$ Stab F. Then ρ_2 is determined by ρ_1. Indeed if Stab F also contains (ρ_1, ρ_2') then it contains $(e, \rho_2(\rho_2')^{-1}) \in$ Stab F as well, where e is the identity map, and

$$\overline{\rho_2(\rho_2')^{-1}} = F\overline{e}F^{-1} = \overline{e} .$$

Since Q is faithful, we have $\rho_2(\rho_2')^{-1} = e$, so $\rho_2 = \rho_2'$.

We may regard A and B as Stab F-sets by taking $(\rho_1, \rho_2)(a) = \rho_1(a)$ for $a \in A$ and $(\rho_1, \rho_2)(b) = \rho_2(b)$ for $b \in B$. In particular, the following two statements are equivalent:
(i) There exists $f : A \to B$ such that Stab $F \subset$ Stab f.
(ii)There is an $f : A \to B$ which is an isomorphism of Stab F-sets.

For (i) says that $\rho_2 = f\rho_1 f^{-1}$ and (ii) says that for $(\rho_1, \rho_2) \in \text{Stab } F$ we have $f\rho_1 = \rho_2 f$.

We may also regard QA and QB as (Stab F)-sets by taking $(\rho_1, \rho_2)(\overline{a}) = \overline{\rho_1}(\overline{a})$ for $\overline{a} \in QA$ and $(\rho_1, \rho_2)(\overline{b}) = \overline{\rho_2}(\overline{b})$ for $\overline{b} \in QB$. Certainly $F : QA \to QB$ is an isomorphism of (Stab F)-sets.

In general, if G is a group and A is a G-set, QA is a G-set in a natural way, since Q is a functor. We are now ready for the

Main Theorem The faithful endofunctor Q is bijectively cancellable if and only if for all finite groups G and for all finite sets A and B, QA and QB are isomorphic G-sets if and only if A and B are isomorphic G-sets.

Proof In one direction, the stated condition, along with the recently noted fact that $F : QA \to QB$ is an isomorphism of (Stab F)-sets, implies condition (ii) above. This implies condition (i), which by the Fundamental Lemma implies the existence of an equivariant map.

Conversely, suppose the stated condition fails. Then let A and B be nonisomorphic G-sets with QA and QB isomorphic as G-sets, and let $F : QA \to QB$ be a G-set isomorphism. Then G maps to Stab F, but the equivalence of (i) and (ii) forbids Stab $F \subset$ Stab f for any $f : A \to B$, lest A and B be isomorphic, and the Fundamental Lemma says that no equivariant map exists. □

Cancellation of Disjoint Union

In this and the following sections, we take A, B, and C to be non-empty sets.

First we consider the disjoint union construction $X \mapsto X \mathring{\cup} C$, which we will hereafter write as $X + C$ to suggest the analogy with arithmetic addition. Suppose that one knows a bijection $F : A + C \to B + C$. Given $a \in A$, iterating f sufficiently often on a must eventually produce and element b of B. Thus we define a bijection $\hat{f} : A \to B$. (See [Stanley].) Notice that the only *a priori* bound on the number of iterations required is the cardinality of C. What is crucial for us is that the construction of $F = \hat{f}$ used only the information provided by the bijection f, not the properties or names of the elements of A or B. Underlying this situation is the simple group-theoretical fact that, for any finite group G and G-set C, G-sets A and B are isomorphic if and only if G-sets $A + C$ and $B + C$ are isomorphic. (See [Burnside].) More specifically, note that for a fixed group G, the effect of the operation Q on the each of the G-set invariants σ, σ^+, τ is to add a constant vector,

which is clearly an invertible operation.

In light of the Main Theorem, the reader may wonder why we allow the possibility that C carries a nontrivial G-action. Q here is the construction "disjoint union with the fixed set C." Since we do not insist on C being disjoint from A and B, the relabeling their elements may induce a permutation on C as well. The possibility of C carrying a non-trivial group actions corresponds to the fact that disjoint union of sets with C can be accomplished in a C-equivariant way (i.e., one that does not depend on the labelling of the elements of C).

Noncancellation of Products

Now we consider the Cartesian product construction $X \mapsto X \times C$. Fix a set X, $|X| > 1$, and set both A and C equal to the set of linear orderings of X. Let B be the set of permutations of X. We may define a bijection between between $A \times C$ and $B \times C$ by using the fact that two linear orderings of X induce a permutation of X. There can be no canonical bijection between A and B, however: The set of permutations of X has a distinguished element, the identity permutation, but the set of linear orderings does not.

Here is the general group-theoretic viewpoint. Let G be a finite group. Let G_t be the G-set obtained by having G act on itself by translation and G_c be the G-set obtained by having G act on itself by conjugation (so that g sends g' to $gg'g^{-1}$). For a nontrivial group, G_t and G_c cannot be isomorphic as G_t is transitive but G_c is not. Nevertheless $h : G_t \times G_t \to G_c \times G_t$ defined by $h(g_1, g_2) = (g_1 g_2^{-1}, g_2)$ is always an isomorphism of G-sets.

Cancellation of Pointed Products

The situation is altogether different when we multiply A and B by a set C with a distinguished point.

From the group-theoretic viewpoint, the analog of C is a G-set with a fixed point.

Theorem A and B and G-sets and and C is a G-set with a distinguished point, and $A \times C$ is isomorphic to $B \times C$, then A is isomorphic to B.

Proof A G-set A is determined by the invariant, σ_A, the integer-function on the set of subgroups of G which records how many times each subgroup of H of G occurs as a stabilizer. Given a point $a \in A$ and a point $c \in C$, the stabilizer of (a, c) in $A \times C$ is $\mathrm{Stab}(a) \cap \mathrm{Stab}(c)$. It follows immediately that $\sigma_{A \times C}$ is linear in σ_A. Represent the linear transformation by the matrix M.

We show that M is nonsingular. Order the subgroups of G so that their cardinalities are non-decreasing. Then M is upper triangular. The fixed point of C guarantees that the diagonal entries of M don't vanish. □

Notice the contrast with the previous situation. The G-set obtained by having G act on itself by translation had no fixed point and this led to an upper triangular matrix with some zeros on the diagonal.

We may also exhibit this cancellation principle more explicitly by the following construction. Now A and B are finite sets, and C is a finite pointed set, with a distinguished element called $*$. Let $f : A \times C \to B \times C$ be a bijection. We define a map $f_* : A \to B$ as follows. $f_*(a)$ is the projection of $f((a, *))$ onto B. A map $f_*^{-1} (= (f^{-1})_*) : B \to A$ is defined similarly. Iteration of the map $f_* \cup f_*^{-1} : A \cup B \to A \cup B$ produces some cycles. We then use f_* to pair any element a of A that occurs in a cycle with an element $f(a)$. This gives us a nontrivial partial bijection between A and B, say from $\tilde{A}$ to $\tilde{B}$. Multiplying by C we get a nontrivial partial from $A \times C$ to $B \times C$ taking $\tilde{A} \times C$ to $\tilde{B} \times C$. This induces a bijection from $(A \setminus \tilde{A}) \times C$ to $(B \setminus \tilde{B}) \times C$ (see the section "Cancellation of Disjoint Union") and now we may iterate the process until we get a bijection from A to B. It should be noted that this construction was only discovered after the group-theoretical approach suggested its feasibility.

Cancellation of Powers

Given a bijection between between the m-fold Cartesian powers of two finite sets, we will use the Main Theorem to define a bijection between the sets themselves.

Theorem Fix a finite group G. If the m-fold powers r^m and r'^m of the permutation representations r and r' of G are isomorphic, then r and r' are isomorphic.

Proof Without loss of generality, assume r and r' both represent G in Sym S for some finite set S. Recall that $L(G)$ is the lattice of subgroups of G, and that the permutation representation r determines a function $\sigma_r : L(G) \to \mathbf{Z}$ which counts the number of times each subgroup of G occurs as the stabilizer of an element in S. The stabilizer of an m-tuple $(s_1, \ldots, s_m)$ is just $\bigcap \text{Stab } s_i$. This observation makes it easy to compute σ_{r^m} from σ_r: for any subgroup G' of G,

$$\sigma_{r^m}(G') = \sum_{\substack{H_1, \ldots, H_m \\ G' = \bigcap H_i}} \sigma_r(H_1) \cdots \sigma_r(H_m).$$

In particular, $\sigma_{r^m}(G) = \sigma_r(G)^m$ since an intersection of subgroups that gives G cannot involve groups other than G itself. We will prove that σ_r can be recovered from σ_{r^m} by induction on the length of the longest chain from a subgroup G' to the top of the lattice. We have already seen that $\sigma_r(G) = (\sigma_{r^m}(G))^{1/m}$. Let G' be a proper subgroup of G, and assume that we have calculated $\sigma_r(H)$ for all groups H strictly containing G'. Then

$$\sigma_{r^m}(G') = \sum_{\substack{H_1,\ldots,H_m \supset G' \\ G' = \bigcap H_i}} \sigma_r(H_1)\cdots\sigma_r(H_m) + \sum_{j=1}^{m} \binom{m}{j} \left(\sum_{H_1,\ldots,H_{m-j} \supset G'} \sigma_r(H_1)\cdots\sigma_r(H_{m-j}) \right) \sigma_r(G')^j$$

where the H_i are subgroups of G that properly contain G'. Since the right side is a polynomial in $\sigma_r(G')$ with positive coefficients, the equation can have at most one non-negative solution. Thus σ_{r^m} determines σ_r. Since $\sigma_{r^m} = \sigma_{r'^m}$, it follows that $\sigma_r = \sigma_{r'}$, which implies that r and r' are isomorphic. □

The Main Theorem now implies that the m-fold Cartesian power functor is cancellable. However, we do not have a satisfying, concrete description of such a cancellation principle.

Noncancellation of Power Sets

The power set of S will be denoted 2^S. We show that a bijection between 2^{S_1} and 2^{S_2} does not generally induce a canonical bijection between S_1 and S_2 by finding a finite group G and a pair of G-sets S_1 and S_2 which are not isomorphic even though 2_1^S and 2_2^S are isomorphic. The counterexample is then the isomorphism between 2_1^S and 2_2^S, considered simply as a bijection between power sets. In this light, the G-action appears as relabelling-symmetries. If there were a canonical induced bijection between S_1 and S_2 it would also possess the same relabelling-symmetries. This is impossible, since G has a different action on the two sets.

Since the stabilizer of an element is H if it is stable under H but under no strictly larger subgroup of G, we have

$$\sigma_S^+(H) = \sum_{K \supseteq H} \sigma_S(K)$$

and

$$\sigma_S(H) = \sigma_S^+(H) - \sum_{K \supset H} \sigma_S(K).$$

The function $\sigma_S(H)$ can be determined from $\sigma_S^+(H)$ working inductively down the lattice of subgroups. Moreover, each orbit of H-type contains $[N_G(H):H]$ elements with stabilizer H, so

$$\sigma_S(H) = [N_G(H):H]\tau_S(H)$$

where $[N_G(H):H]$ is the index of H in its normalizer. The isomorphism type of a G-set S is clearly determined by the function $\tau_S(H)$, and so by either of the functions $\sigma_S^+(H)$, $\sigma_S(H)$.

Let $\mathcal{O}$ be a G-orbit of H-type, with $s \in \mathcal{O}$ having stabilizer H. Pairing elements gs of $\mathcal{O}$ with cosets gH of H gives a G-set isomorphism between $\mathcal{O}$ and G/H. If K is a subgroup of G, the following are clearly all equivalent:
(i) gs is in the same K-orbit as $g's$
(ii) there exist $k \in K$ such that $gs = kg's$
(iii) there exist $k \in K$ such that $gH = kg'H$
(iv) there exist $k \in K$, $h \in H$ such that $g = kg'h$
(v) $KgH = Kg'H$.
Let $o(K,H)$ be the number of K-orbits in G/H. The equivalence of (i) and (v) implies that $o(K,H)$ is also the number of double cosets of the form KgH. For any H' conjugate to H, $\mathcal{O}$ is also of H' type, so $o(K,H') = o(K,H)$. As the number of double cosets KgH equals the number of double cosets HgK (note $(HgK)^{-1} = Kg^{-1}H$), $o(K,H) = o(H,K)$. In particular, $o(K,H)$ also only depends on the conjugacy class of K.

The key formula is

$$\sigma_{2^S}^+(K) = 2^{\sum_H o(H,K)\tau_S(H)},$$

where H ranges over representatives of the conjugacy classes of G. The left side is the number of subsets of S which are stabilized by K. A subset of S is stabilized by K exactly if it is the union of K-orbits of S. The number of K-orbits of S is $\sum_{H \subseteq G} o(H,K)\tau_S(H)$ since each G-orbit of H-type decomposes into $o(H,K)$ K-orbits.

Example 1 Take $G = Z/2Z \times Z/2Z$, the Klein 4-group. Since G is Abelian the issue of conjugacy of subgroups is moot. The subgroups of G are G itself, the three 2-element subgroups R_1, R_2 and R_3, and $\{e\}$. Taking the subgroups in this order, the matrix

$$[o(H,K)] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 \\ 1 & 1 & 2 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 2 & 4 \end{bmatrix}$$

is singular, with vector $[-2\ 1\ 1\ 1\ -1]$ in the kernel. The linear transformation $[o(H,K)]$ takes the same values on $[2\ 0\ 0\ 0\ 1]$ and $[0\ 1\ 1\ 1\ 0]$. Let S_1 be a G-set with two orbits of type G and one of type $\{e\}$. Let S_2 be a G-set with one orbit of type R_i, $i = 1,2,3$. Then 2^{S_1} and 2^{S_2} are isomorphic G-sets.

Example 2 Take $G = S_3$, the symmetric group on three letters. The conjugacy classes of subgroups of G are represented G itself, the 3-element cyclic subgroup C_3, any of the 2-element subgroups R_i generated by a reflection, and $\{e\}$. Taking the subgroups in this order, the matrix

$$[o(H,K)] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 6 \end{bmatrix}$$

is also singular, with vector $[-2\ 1\ 2\ -1]$ in the kernel, so the linear transformation $[o(H,K)]$ takes the same values at $[2\ 0\ 0\ 1]$ and $[0\ 1\ 2\ 0]$. Let S_1 be a G-set with two orbits of type G and one of type $\{e\}$. Let S_2 be a G-set with one orbit of type C_3, and two of type R_i. Then 2^{S_1} and 2^{S_2} are isomorphic G-sets.

The preceding examples are far from rare, as we will now show.

We thank Goetz Pfeiffer for a demonstration of the following

Theorem The matrix $[o(H,K)]$ has nonzero determinant precisely when G is cyclic.

We have modified Pfeiffer's approach to avoid quoting results from representation theory. First we will need a

Lemma Let G be a finite group with subgroups H and K. Then the number of double cosets KgH is the scalar product of the permutation characters of H and K.

Proof The number of double cosets KgH is the number of K-orbits in G/H. By the orbit counting formula, this is

$$\begin{aligned} \frac{1}{|K|}\sum_{k\in K}|\mathrm{Fix}_{G/H}k| &= \frac{1}{|K|}\sum_{k\in K}\frac{1}{|H|}\sum_{g\in G}\delta(kgH = gH) \\ &= \frac{1}{|K|}\frac{1}{|H|}\sum_{k\in K}\sum_{g\in G}\delta(g^{-1}kg \in H) \end{aligned}$$

Here $\delta(\mathcal{P})$ is 1 if $\mathcal{P}$ is true and 0 if not. Since each pair (k,g) satisfying $k \in K, g^{-1}kg \in H$ gives rise to $|G|$ triples $(x,y,z) = (g'kg'^{-1}, g', g'g)$ satis-

determines the cardinality of A. The categorical view says Q is cancellable if there is a functor from $\mathbf{Fin}_Q$ back to $\mathbf{Fin}$. In general one cannot expect such functors. Indeed there is a paucity of interesting group homomorphisms

$$\mathrm{Hom}_{\mathbf{Fin}_Q}(A,A) = \mathrm{Hom}_{\mathbf{Fin}}(QA,QA) \to \mathrm{Hom}_{\mathbf{Fin}}(A,A)\ .$$

The question arises then, if our equivariant maps

$$\mathcal{F}_{A,B} : \mathrm{Bij}(QA,QB) \to \mathrm{Bij}(A,B)$$

do not define functors, how close do they come? Indeed we have functor-like gadgets $\mathcal{F}$, mapping objects to objects and arrows to arrows from $\mathbf{Fin}_Q$ to $\mathbf{Fin}$. Nevertheless, for composable arrows x and y we only insist that

$$\mathcal{F}(xy) = \mathcal{F}(x)\mathcal{F}(y)$$

when at least one of x or y is in the image of the functor Q. We call such an $\mathcal{F}$ a semifunctor. In general we can consider semifunctors in any category with a distinguished subcategory. Note that a single equivariant bijection $\mathcal{F}_{A,B}$ induces a semifunctor on a component of the category; there are no further compatability constraints. If we take $A = B$, we are just looking at maps from $\mathrm{Bij}(QA,QA)$ to $\mathrm{Bij}(A,A)$ as two sided $\mathrm{Bij}(A,A)$-sets, so we are back to group theory. Since these are both symmetric groups, we frame the following

Problem Classify pairs (G,H), with G and H symmetric groups satisfying $H \subset G$, according to the existence of a map $G \to H$ that respects G and H as two-sided H sets.

REFERENCES

[Burnside] W. Burnside, *Theory of Groups of Finite Order.* Dover Publications, 1955.

[Jacobson] Nathan Jacobson, *Basic Algebra I.* W. H. Freeman, 1974.

[Joyal] André Joyal, *Une théorie combinatoire des séries formelles,* Adv. in Math. **42**, Academic Press (1981), 1-82.

[Stanley] Richard Stanley, *Enumerative Combinatorics I.* Wadsworth, 1986.

That G be cyclic is also sufficient for D to have full rank. Let $\tilde{C} = [\tilde{c}_{ij}]$ be the $s \times s$ matrix, with rows indexed by the subgroups of G, columns indexed by a generator for each subgroup and $\tilde{c}_{ij}$ equal to the number of elements fixed when g_j acts on O_i times the number generators of $\langle g_j \rangle$. Clearly $\tilde{C}\tilde{C}^T = CC^T = D$, since $\tilde{C}$ amounts to "collecting" the identical rows of C. On the other hand $\tilde{C}$ is invertible since its columns are nonzero multiples of the columns in the table of marks. This concludes the proof.

Comment The previous argument gives a formula for the determinant of $D = D(n)$ when $G = \mathbf{Z}/\langle n \rangle$:

$$\mathrm{Det}(D) = \prod_{d|n} d\phi(d) .$$

Alternatively, the determinant of a tensor product is given by the formula

$$\mathrm{Det}(A \otimes B) = \mathrm{Det}(A)^{\dim(\mathrm{B})} \cdot \mathrm{Det}(B)^{\dim(\mathrm{A})} .$$

If $n = pq$ with p and q relatively prime, then essentially

$$D(n) = D(p) \otimes D(q) .$$

If p is prime, then

$$\mathrm{Det}(D(p^n)) = p^{p-1}(p-1) \cdot \mathrm{Det}(D(p^{n-1}))$$

and $\mathrm{Det}(D(p)) = p - 1$.

The theorem guarantees a rich supply of counterexamples to the cancellation of the power set construction, but in a small way it has a positive aspect as well, for it shows that bijection between powers sets which have only cyclic symmetry may in fact be cancelled.

Categorical Viewpoint

Till now, we have regarded constructions Q as endofunctors of the category **Fin** of finite sets and bijections. It is better here to consider the functor $\widehat{Q}$ from **Fin** to $\mathbf{Fin}_Q$, the full image of Q. The objects of $\mathbf{Fin}_Q$ are the objects of **Fin**, but

$$\mathrm{Hom}_{\mathbf{Fin}_Q}(A, B) = \mathrm{Hom}_{\mathbf{Fin}}(QA, QB) .$$

On objects $\widehat{Q}$ is the identity map (this avoids the technical annoyance that $QA = QB$ does not generally imply $A = B$) and on maps $\widehat{Q}$ coincides with Q. Note that $\widehat{Q}$ is surjective on objects where Q is not.

Our notion of a canonical cancellation for a construction Q lies between two extremes. The classical view says Q is cancellable if the cardinality of QA

fying $y^{-1}xy \in K, z^{-1}xz \in H$, we may rewrite what we had as

$$\begin{aligned}
& \frac{1}{|G|}\frac{1}{|K|}\frac{1}{|H|}\sum_{x\in G}\sum_{y\in G}\sum_{z\in G}\delta(y^{-1}xy \in K, z^{-1}xz \in H) \\
= & \frac{1}{|G|}\sum_{x\in G}\frac{1}{|K|}\sum_{y\in G}\delta(y^{-1}xy \in K)\frac{1}{|H|}\sum_{z\in G}\delta(z^{-1}xz \in H) \\
= & \frac{1}{|G|}\sum_{x\in G}\frac{1}{|K|}\sum_{y\in G}\delta(xyK = yK)\frac{1}{|H|}\sum_{z\in G}\delta(xzH = zH) \\
= & \frac{1}{|G|}\sum_{x\in G}\mathrm{Fix}_{G/K}x\mathrm{Fix}_{G/H}x
\end{aligned}$$

as desired. □

W. Burnside's classic *Theory of Groups of Finite Order* introduces a matrix called there the "table of marks" associated to any finite group G. Let

$$G_1 = \{e\}, G_2, \ldots, G_s = G$$

be a sequence of representatives for conjugacy classes of subgroups of G ordered so that

$$|G_1| \leq |G_2| \leq \cdots \leq |G_s|,$$

and let O_i be a G-orbit of G_i-type, e.g., G/G_i. The "table of marks" is then the $s \times s$ matrix $B = [m_{ij}]$ where m_{ij} is just the number of elements of O_i fixed by G_j. The stabilizers of elements of O_i are conjugates of G_i, so $m_{ij} = 0$ unless G_j is contained in some conjugate of G_i; in particular $m_{ij} = 0$ if $i < j$. On the other hand m_{ii} is always positive since O_i certainly contains at least one G_i-stable element. Thus $[m_{ij}]$ is a lower triangular matrix with non-zero entries on the diagonal, and the determinant of $[m_{ij}]$ does not vanish.

Now let $C = [c_{in}]$ be the $s \times |G|$ matrix, with rows indexed by the conjugacy classes of subgroups of G, columns indexed by the elements of G, and c_{in} equal to the number of elements fixed when g_n acts on O_i. The rows of C are by definition the permutations characters of G. The n^{th} column of C coincides with the column of the table of marks corresponding to the (conjugacy class of) the subgroup $\langle g_n \rangle$ of G generated by g_n, since $\langle g_n \rangle$ and g_n leave the same elements fixed. The rank of C is thus the number of conjugacy classes of cyclic subgroups of G.

By the lemma, CC^T is the $s \times s$ matrix $D = [d_{ij}]$, with rows and columns indexed by conjugacy classes of subgroups of G and $d_{ij} = |G|o(G_i, G_j)$. The rank of D cannot be more the rank of C, so for D to have full rank, every subgroup of G, including G itself, must necessarily be cyclic.