

GROUP FUNCTORS IN PLAIN ENGLISH

18 July 2008, (latest correction 23 September 2008)

Dedicated to my typist, Laurence Barker, Bilkent University.

My purpose is to indicate the notion of group functors in a straightforward way, accessible to anyone with the aptitude and inclination to follow, say, a popular science-fiction novel or film. The simile is helpful: if the space-ship is being sucked into a hyper-space worm-hole and the engineer is screaming that the quantum morphic resonator is jammed, then some of us do glean the gist of the situation.

I would like to thank my typist for surrendering to me. Sometimes, I did wish that he could have helped me by checking up on some sources or perhaps providing some references, maybe even some quotations. But then I would have been on my way towards evolving into an academic paper, with all the tedious responsibilities which that would entail. I know not what I am, nor whether or not there may be others like me, but the one thing I do know for sure is that I wrote myself, *ergo sum*.

Possibly it was Poincaré who suggested, quixotically, that one has not understood something unless one can explain it to just anyone picked off the street. My typist refuses to check the attribution, let alone to chase up the quote.

It (certainly) was Lakatos who offered the following very succinct description of modern pure mathematics: A research programme becomes recognized as interesting when it has been used to resolve a conjecture in another research programme already recognized as interesting. Towers of abstraction rise up. That is to say, sequences of research programmes emerge, P_0 , P_1 , P_2 , ..., where P_{n+1} resolves at least one conjecture in P_n .

This is an idealization, and let us treat it as such. Starting from solid ground, we shall ascend a little tower of abstractions, arriving eventually at an indication as to the theory of group functors. Oddly enough, at this scale of simplification, the levels of our little tower will have a positive correlation with the chronological order of historical development.

One major omission from above description of modern pure mathematics is that, actually, mathematics is a vast city of inter-connected towers. When the architects standing at the top of their tower make assessments as to how to proceed, they often to look across towards neighbouring towers. They gain much of their motivation through recognition of fashionable themes. Kuhn had something to say about this, but let us not pause to rehearse it. The mathematical literature reveals that mathematicians are unable to identify and express the themes (the only alternative is the stranger hypothesis that they are unwilling to communicate about such matters, even to each other). So how might a mathematician explain her research to a layman? If she cannot discuss her research in a self-reflective way with her colleagues, nor present it coherently to her students, nor explain it to a layman, then perhaps she really does not understand what she is doing. So I congratulate my typist, for being sufficiently open-minded to hand the task over to me.

Level 0: Set-wise structure. Let us agree that, initially and ultimately, we are interested in symmetry. But it is an elusive notion, and we have to nail it down somehow. As everyone knows, the abstract essence of symmetry is expressed in the notion of a group. So, to study symmetry is to study group theory.

Immediately, we have been compromised. Already, we have aligned ourselves with a particular approach to mathematics, the *structure-based approach*. It is not the only conceivable approach. Another one, which might be called the *formula-based approach*, has the upper hand in many areas of the mathematics of natural scientists and engineers. Nevertheless, for the

last seven decades or more, the structure-based approach has had complete and uncontested dominion in pure mathematics.

Since the word *structure* is really a mantra, it does not have a definite meaning. It tends to be used in two different senses, both of them quite nebulous. One of them is what we shall call *set-wise structure*. This was the great idea that caused a revolution in mathematics during the last three decades of the 19th century. Another three decades were needed before the idea became sufficiently distilled for mathematicians to actually put it down in words. To hazard a one-sentence summary of the prolific output of Bourbaki: mathematical objects are understood to be sets equipped with structure.

Roughly, a set is a collection of mathematical objects. When we define a new object, we collect together some objects which have already been defined, and then we impose something extra, a mysterious spirit called structure. Some minor logical troubles arise from the circularity — mathematical objects as collections of mathematical objects — but, in practise, those troubles can always be evaded by means of suitable manoeuvring. Of much greater importance: what is structure? Let us give two some examples, both of which will be central to our subsequent discussions.

The familiar rules of arithmetic crop up again and again in other contexts, so we abstract the essence: a ring is a set which acquires structure through being equipped with two operations which behave like the addition and multiplication of numbers; intuitively, we think of the elements of the ring as generalized numbers, even if they do not really have any kind of numerical interpretation.

Symmetry, too, crops up again and again. We can understand the symmetries of an object to be the structure-preserving permutations of the object. Abstracting the essence: a group is a set which acquires structure through being equipped with an operation which behaves like composition of symmetries; intuitively, we think of the elements of the group as generalized structure-preserving permutations, even if they do not actually permute anything.

Where is the structure, here? Well, it really a figment of fantasy, but some sense of structure comes about in the following rather indirect way. When a function from one ring to another preserves the additive and multiplicative operations, we say that the function is structure-preserving and we call the function a homomorphism. If the function is invertible then its inverse must also preserve the two operations. In that case, we say that the function transports structure; we call the function an isomorphism; we say that the two rings have the same structure; we say that the two rings are isomorphic to each other. The same holds for group homomorphisms and group isomorphisms. What we have defined, here, are the terms *structure-preserving* and *transport of structure* and *same structure*. It is these terms which are genuinely meaningful. The fantasy notion of *structure* is a kind of mystical essence which we imagine lies behind those terms.

Another of the great repercussions of this change was that mathematicians now had two different notions of what it is for two objects to be identical or indistinguishable. There was still the traditional notion of identity as equality, but there was now also the new notion of identity as isomorphism. The jargon “up to isomorphism” is often used as a signal to emphasize that isomorphic objects are not to be distinguished from each other.

One of the uses of this idea is that, very often, an object can be defined or constructed in many different ways which are in some sense all equivalent to each other. It may be desirable to retain several different ways of regarding the object, but it may also be desirable to understand that all the versions of the object are in some sense one and the same. This compromise can be achieved by admitting the versions of the object as isomorphic to each other but not equal

to each other. We shall be returning to this idea below, when we discuss natural equivalence.

Since the notion of isomorphism is arguably the single most important notion in all of modern pure mathematics, let us mention a technical example. The group of rotational symmetries of a cube is isomorphic to the group of rigid symmetries of a regular tetrahedron. (This can be seen by identifying the four diagonals of the cube with the four vertices of the tetrahedron.) The two symmetry groups, isomorphic but not equal, are both versions of a group called S_4 ; they are said to be isomorphic copies of S_4 .

One of the principles of the structure-based approach is that mathematics is deemed to be the study of those properties, features, constructions that are isomorphism invariant. For instance, when considering an object X , the only properties of X that are of legitimate concern to us are those properties which also hold for all objects isomorphic to X . But what all those objects have in common is the essence of X , the structure of X . We arrive at the conclusion that, according to this view, mathematics must be the study of that elusive thing, scarcely even a figment of our imagination since we can neither define it nor directly perceive it: structure.

Level 1: Representations, in general: Let us bring in a more dynamic language. René Thom suggested that, in addition to the Kantian intuitions of time and space and number, mathematics involves some other intuitions derived from physical sensation, for instance, a kinetic sense of motion. Mathematical objects do not just sit there passively possessing properties. They move. They do things. When a group expresses some of the symmetries of an object, we say that the group acts on the object. This language accurately reflects our perception of the relationship between the group and the object. The group is the grammatical subject and the object is, indeed, the grammatical object. This kind of relationship between an active group and a passive object is called a *representation* of the group.

In applications of representation theory, the passive object is given and we introduce a group which acts upon it. But, in the abstract study of groups themselves, it may be the group that is given in the first place. Then we can put the matter backwards, introducing an object and letting the given group act upon it. For example, the passive object may start off as an abstract set, its structure entirely banal. But then it acquires further structure as a set upon which the given group acts as permutations. Equipped with this further structure, we call it a permutation set for the group. The relationship between a group and a permutation set is called a *permutation representation* of the group. As another example, the object may start off with a certain kind of additive structure and then it acquires further structure from the group action. The relationship here is called a *linear representation* of the group.

The same thing can be done with rings instead of groups. Rings act on sets which start off with a certain kind of additive structure. Through the action of the ring, these sets acquire further structure. Equipped with this further structure, the sets are called the *modules* of the ring. This relationship between a ring and a module is called a *representation* of the ring.

These ideas can be combined in a useful way. Given a group, we can construct a ring, called the group ring, such that the linear representations of the group are precisely the representations of the group ring. Most of the applications of group theory to natural science and engineering arise from this kind of representation.

Again, we have been speaking freely of that ghost in the machine, the structure. It makes sense in the same way as before. For permutation representations, linear representations, representations of rings, indeed, for representations in general, we can understand the homomorphisms to be the functions which preserve the action and which also preserve all the other equipment. In particular, there is again a notion of isomorphism, same structure.

It was only as a nuance of commentary that we have been speaking of representations as

relationships while speaking of permutation sets and modules as passive objects. Much mental effort can be saved by disregarding such subtle distinctions. A permutation representation is essentially the same thing as a permutation set; the two terms are just two different labels to refer to the same scenario. Similarly, a representation of a ring is essentially the same thing as a module of the ring. And a linear representation of a group is essentially the same thing as a module of the group ring.

Level 2: Fiction. Prior to the 19th century, there were perhaps only three major examples of mathematical ideas with a distinctive air of creation or invention, as opposed to observation and discovery. It is worth saying a few words about those three examples because, even today, they still serve as paradigms. Two of them, from 16th century Italy, are the negative numbers and the imaginary numbers. The negative numbers — which Cardan originally called the *numera ficti* — enabled mathematicians to work freely with the subtraction operation. The imaginary numbers eventually enabled mathematicians to work freely with roots to polynomial equations. Generally, it is often helpful to extend a mathematical object by augmenting it with some extra fictitious material, thus creating a new object, larger than original object but with nicer properties.

The earliest of the three examples, but also the most profound, is the Greek invention of ratios in the 4th century BC. Recognizing that their notion of number was inadequate for comparisons of magnitude, the Greeks introduced ratios of lengths, ratios of volumes, ratios of intervals of time, and so on. Some clues in Aristotle indicate that a few decades were needed for the Greeks to hit upon the notion of a ratio in abstract, independently of any particular kind of magnitude such as length or volume or time.

For the sake of an observation below, let us also point out that the Greek notion of ratio is also an example of what might be called a *sophomoronism*. The notion is nowadays subsumed in the more recent notion of a real number. With hindsight, we find that, if only the Greeks had been willing to view ratios in a more numerical way, then the theory of ratios could have been made less subtle, less complicated and yet more powerful.

The significance of the notion of set-wise structure is that it gave mathematicians the licence and also the technology to invent rather freely. There is only one constraint, but it is severe: our lack of imagination. For inspiration, we can look back to the 16th century and the 4th century BC.

As a trick for eliding over technicalities, let us enclose academic hedging in parentheses. Given a (finite) group, then the set of isomorphism classes of (finite) permutation sets is not quite a ring. But we can extend it by adapting the notion of the negative numbers — introducing fictitious permutation sets called virtual permutation sets — and thus we obtain a ring, called the Burnside ring. The power of this definition comes from the fact that it allows us to play with the isomorphism classes of permutation sets — and virtual permutation sets — all together, as a composite entity which has a kind of arithmetic structure of its own.

But some more fiction needs to be introduced before we can get very far with this. Just for the purposes of this essay, let us understand an *algebra* to be a ring which has been extended so as to have some nice properties. (Just for these purposes, we shall understand an algebra to have a field as the coefficient ring.) The Burnside algebra has a very rich structure, and a great many techniques can be brought to bear on it. For instance, the Burnside algebra has a wealth of elements — as it were, isomorphism classes of fractional virtual permutation sets — which satisfy the very useful idempotent equation $e^2 = e$. Working with (finite-dimensional) modules instead of permutation sets, we obtain some other useful algebras, notably the representation algebras of the group and the character algebras of the group.

It should now become apparent how the set-wise notion of structure not only allows but actually encourages those towers of abstraction that we mentioned earlier. New objects continually being constructed as collections of old objects, there is an inevitable tendency to be forever ascending to higher levels of abstraction. Indeed, it is very easy to generate new mathematics in this way. The obvious idea for the next level of abstraction, surely, would be to examine the modules of algebras such as the Burnside algebra and the character algebra. But productive ideas tend to be those which address problems. To explain one of the fundamental difficulties with group representation theory, we must now discuss another notion of structure.

Level 3: Composition-structure. The term *structuralism*, as understood by present-day philosophers of mathematics, derives from the notion of set-wise structure. But the origin of the term is illuminating. Finite group theory was invented by Galois in his study of the symmetries of polynomial equations. The theory of Lie groups, initiated fifty-odd years later, was intended to as a kind of Galois theory for partial differential equations. Lie introduced the term *Zusammensetzung* — roughly, together-putting — in reference to the way a Lie group can be made up of smaller Lie groups. In 1889, two students of his translated the term into French as *structure*. To defend the two students against the charge of mistranslation, we mention that, in the 1890s, Killing objected to Lie’s term, and suggested *Gestaltung* — literally, shape — as a more appropriate term. (References neglected but, for this delightful historical twist, acknowledgement must be made to Trevor Hawkins.) To avoid ambiguity, let us use the term composition-structure to refer to structure in this sense.

It is not hard to imagine how, over the course of decades, the notion of composition-structure might have mutated into the notion of set-wise structure. Nevertheless, the two notions are quite distinct. When algebraists speak of group structure or module structure, they are always referring to the composition-structure.

Composition-structure is most easily illustrated in the context of modules. One way to investigate a module is to consider the modules it contains, called the submodules. Given a submodule N of a module M , we can obtain another module, the quotient module M/N , by squashing N down to a single point. In this way, the module M has been broken down into two smaller components, the submodule N and the quotient module M/N . Repeating the process, we can sometimes break down a module into simple modules, we mean to say, modules that cannot be broken down any further. If the simple modules of an algebra have been classified — up to isomorphism, of course — then that may be of use towards studying the modules of that algebra in general. When an algebraist says that she has determined the structure of a module, she usually means that she knows the simple subquotients, she knows which of them fit on top of which, but she usually does not have the much more subtle information as to quite how each simple subquotient is fitted on top of the lower ones.

For groups, the situation is a little more tricky. Quotient groups can only be formed for a special kind of subgroup called a normal subgroup. Because of its intricacy, the classification of the simple finite groups is of only limited use. The most successful reduction techniques in finite group theory tend to reduce to particular kinds of subquotients, for instance, subquotient groups whose order is a power of a prime.

But now we come to the problem. As we have said, the most successful methods of group theory tend to reduce to certain kinds of subquotients. Thus, they involve the composition-structure of the group. Similarly, the methods of module theory tend to involve the composition-structure of modules. So, in group representation theory, we are dealing with two kinds of composition-structure: that of the group and that of the modules of the group ring. Relationships between these two compositions-structures exist but they are mysterious. We find

ourselves with two reduction techniques, but we lack a clear understanding of how to use the two techniques in tandem. The conjectures, most notably Alperin's Conjecture, indicate that we are missing some key idea.

Permutation representations and linear representations admit five kinds of operation, called the *elemental operations*, which are related to group structure. Let us list them, because they will come to be of vital importance below. One of them is *restriction* from the representations of a group to the representations of a subgroup. Another is *induction* from the subgroup to the group. Another is *deflation* from a group to a quotient group. Another is *inflation* from the quotient to the group. The fifth, called *isogation*, expresses transport of structure through isomorphism.

Level 4: Category Theory. During the 1940s, just as Bourbaki was getting into full flow on turning the notion of set-wise structure into a dogma, Eilenberg and MacLane introduced some new ideas that undermined it. But the reprieve was short-lived. The dogma achieved a synthesis, and it persisted. A good name for the new form of the dogma might be: neostructuralism.

In the following discussion, we shall speak freely of sets. Some logical troubles arise, but we shall not complicate our discussion with the awkward manoeuvres that are needed to circumvent those troubles.

As we already indicated, mathematical objects are often regarded as sets with structure, and a particular kind of object is often studied by making use of the homomorphisms, structure-preserving functions between objects of that kind. Always, the homomorphisms have certain characteristic features. Eilenberg and MacLane abstracted those features. A category consists of objects and morphisms such that the morphisms have the characteristic features of homomorphisms. Examples include the category of groups, the category of rings, the category of permutation representations of a given group, the category of modules of a ring. In all of those examples, the morphisms are homomorphisms. But the definition contains an evil and zestful liberation, because morphisms are no longer required to be structure-preserving functions, and indeed, objects are no longer required to be sets with structure.

But it is worth taking a look at what Eilenberg and MacLane were aiming to do. For them, the notion of a category was just a stepping-stone to something else. There were many notions in algebraic topology (for example, the homology of a space) which could be defined in several different ways, yet all those versions, in some sense, were equivalent to each other. Eilenberg and MacLane wished to assign a precise meaning to the term *natural equivalence*.

As we noted earlier, this kind of compromise is sometimes achieved by working with two notions of identity: equality and isomorphism. This is, indeed, how Eilenberg and MacLane achieved it. But we need to go into a little more detail. First note that we can consider categories of categories, where the objects are categories and the morphisms, called functors, are the homomorphisms in the usual sense; structure-preserving functions. A category of categories is just another set-wise category, and there is nothing very deep here. (Quibbling logicians may rocket out of their socks here; we hope they land on a planet to their liking.) But we can also consider functor categories, where the objects are functors and the morphisms, called natural transformations, are indeed things that arise naturally, though they are quite unlike any kind of structure-preserving function. The punch-line is: we define a natural equivalence to be an isomorphism in a functor category.

The notion of a group functor is an application of those constructions. From the point of view of group theory, with its reductions to subgroups and quotient groups, the Burnside ring of a group ought to be studied by comparing it with the Burnside rings of the subgroups and

the Burnside rings of the quotients. This can be done using the five elemental operations: restriction, induction, deflation, inflation and isogation. The Burnside rings can be regarded as objects in a certain category and the elemental maps can be regarded as homomorphisms in that category. (Actually, it is the category of abelian groups, not the category of rings.) Similarly, representation rings and character rings admit elemental maps which again, can be regarded as homomorphisms.

The idea that gradually emerged — though the work of Green in the 1970s, Bouc in the 1990s, and many others in-between — was that the essence of the elemental maps can be abstracted. As we noted above, it did take a few decades for the classical Greeks to get used to speaking of ratios in abstract, rather than ratios of lengths, ratios of volumes and so on. Likewise, it took a few decades for modern representation theorists to learn how to discuss, for instance, induction and restriction in abstract, as opposed to the induction and restriction operations on the Burnside ring, the induction and restriction operations on the character ring, and so on.

The device behind the idea was to introduce an abstract category, called the biset category, which has the finite groups as its objects, but the morphisms are something quite different from group homomorphisms, indeed, they are not structure-preserving functions at all. They include the elemental morphisms, namely, the induction morphisms, the restriction morphisms, the inflation morphisms, the deflation morphisms, the isogation morphisms. It might be said that the abstract essences of the elemental operations are captured in the notion of the elemental morphisms.

A biset functor is a functor from the biset category to some other appropriate category. For instance, the Burnside functor carries the abstract structure of the biset category down to a category where the Burnside rings are objects. The elemental morphisms are carried down to become the elemental operations on the Burnside ring. Similarly, a character functor carries the abstract structure down to a category where character rings are objects. The same applies to representation functors and representation rings.

Biset functors are just one kind of group functor. Other kinds of group functor can be defined in similar ways, by removing some of the morphisms from the biset category. For instance, some kinds of linear representation do not admit a deflation operation on the associated character rings. So the group functor here must be of a kind that does not have deflation operations.

The usefulness of this idea is that, passing from rings to algebras, the simple biset functors have been classified, and likewise the simple functors for every other kind of group functor. So now, in addition to the composition-structure of a group and the composition-structure of a linear representation, we have a third kind of composition-structure to bring into play, namely that of a group functor. Furthermore, some peculiar features of the classification of the simple functors indicates that this third kind of composition-structure may be closely related to the other two. A further comment on this admittedly rather vague speculation appears in a comment at the end of this essay.

Level 5: The alchemic algebra. This final section is not an ascent to a higher level of abstraction. It is more a descent to the mesosphere.

For simplicity of discussion, we shall stick to biset functors, though variants of the material apply also to other kinds of group functors. In our account of biset functors, we mentioned that the abstraction here is rather similar to the abstraction behind the notion of ratios: just as the Greeks moved towards abstract ratios in place of ratios of particular kinds of magnitude, we have abstract induction morphisms in place of induction operations in particular instances.

But we also mentioned, earlier, that the Greek notion of ratio is a sophomoronism: the material was later captured in an easier and more powerful way. As a kind of conceit, let us argue for similar conclusion pertaining to biset functors.

The category of biset functors is a particular kind of category called an abelian category, and there is a theorem (the Freyd–Mitchell Embedding Theorem) which asserts than any abelian category is a certain kind of subcategory of the module category of a ring. To put this more simply: there exists a ring Γ such that any biset functor can be regarded as a Γ -module. Fortunately, it not hard to construct such a ring Γ . Called the *alchemic ring*, it is generated by those five kinds of operation: the two upwards operations of induction and inflation, the two downward operations of restriction and deflation, and the more circular operation of isogation. (The name derives from the alchemic understanding that the elements fire and air have an upward motion, the elements water and earth have a downward motion and the fifth element, the quintessence, has the circular motion observed in the planets.) Passing from rings to algebras, the simple biset functors are the simple modules of the alchemic algebra. We have returned to module theory.

This completes our description of the notion of a group functor; regarded either as a functor from an abstract category called the biset category, or else as a module of an abstract ring called the alchemic ring. But is the notion really of interest? Well, the group functor scenario still carries all the objects of traditional concern: permutation representations and linear representations up to isomorphism. The Burnside algebra, the character algebra and the representation algebra are retained too, now regarded as functors or as modules of certain algebras. Our new advantage is that we have now extended to a larger realm with enough “*functi ficti*” to decompose the functors of primary concern into simple functors.

A comment by the typist: First of all, I must express my admiration for this essay, which had the vitality to write itself without any help from me, except for my services in typing it out. Nevertheless, I do feel that the essay has rendered my main research area much in the same way that a Martian might give an account of some area of peculiarly human activity, say, the game of cricket. I can find nothing that is downright wrong with the observations, but still, I expect that any expert will have to exercise some imagination — as I had to — in order to recognize what the essay has been describing.

One point that I found especially difficult to follow was the essay’s inference that the so-called “composition-structure” of group functors might serve as some kind of intermediary between the “composition-structure” of groups and the “composition-structure” of modules. I have struggled at trying to guess what is meant. Perhaps it is something along the following lines: The character functor captures some “composition-structure” of modules because the character algebra of a group has a basis which can be identified with the simple modules of the group. But the character functor also has “composition-structure” as a group functor. It breaks up into simple functors. For any given kind of group functor, the simple functors can be labeled with pairs (H, V) where H is a group and V is a simple module of the group algebra of the outer automorphism group of H (obtained from the group of isomorphisms from H to itself, but quotiented out by the those isomorphisms which come from the so-called conjugation action — read masturbation action — of H acting as symmetries of itself). When considering the character algebra of a given group G , the groups H which appear are the subquotients of G . In this way, I suppose we do obtain a link to the “composition-structure” of G .

Well, that is all very nice, as a piece of sophistry to popularize the topic. But a more realistic and reasonable position to take is just to observe that group functors are a new tool to analyze some structures that have long been recognized as interesting.