NEGATIVE SETS HAVE EULER CHARACTERISTIC AND DIMENSION

Stephen H. Schanuel
Mathematics Department S.U.N.Y. at Buffalo
106 Diefendorf Hall, Buffalo, NY 14214-3093 (U.S.A.)

1. Where are the negative sets?

Though ill-posed, the question is suggestive; a good answer should complete the diagram

$$\begin{array}{ccc} \mathbb{S} & \hookrightarrow & \mathbb{E}\ ? \\ \downarrow & & \downarrow \\ \mathbb{N} & \hookrightarrow & \mathbb{Z} \end{array}$$

where $\mathbb{S}$ is the category of finite sets; we seek an enlargement $\mathbb{E}$, the isomorphism classes of which should give rise to all integers, rather than just natural numbers. Why is this desirable? The utility of the observation that natural numbers are the isomorphism classes of finite sets derives primarily from the fact that sets can carry structure. For instance, with Euler's function $\varphi(n)$ (the number of integers $0 \leq x < n$ relatively prime to n), the equation $\varphi(mn) = \varphi(m)\varphi(n)$ for relatively prime m and n is but a pale reflection of the isomorphism of rings $\mathbb{Z}/mn \cong \mathbb{Z}/m \times \mathbb{Z}/n$. The isomorphism of rings induces an isomorphism of their groups of units, while the equation records only that these groups are isomorphic as sets.

What we seek, then, is a category $\mathbb{E}$ which would allow us to "lift" equations between integers to isomorphisms between objects, because the isomorphism may then preserve some structure relevant to the equation under consideration.

2. A "proof" that there are no negative sets.

We would hope to find $\mathbb{E}$ with finite coproducts and finite products, satisfying at least the distributive laws (that the canonical maps $0 \to A\times 0$ and $A\times B + A\times C \to A\times(B + C)$ are isomorphisms). But already with the coproduct, a difficulty presents itself: $A + B \cong 0$ implies $A \cong B \cong 0$, since to have exactly one map $A + B \to X$ is to have exactly one map $A \to X$ and one map $B \to X$. So the isomorphism classes of objects in a category with coproducts never constitute a nontrivial group under addition. The most we can expect is that the universal map from the set of (isomorphism classes of) objects of $\mathbb{E}$ to a cancellative monoid ($a + b = a + c$ implies $b = c$) will have $\mathbb{Z}$ as codomain.

To clarify our goal, then: $\mathbb{E} \supset \mathbb{S}$ should be a category satisfying distributive laws, and its "rig" of isomorphism classes should have the ring of integers as its reflection into cancellative rigs. A *rig* is a commutative "ring without negatives", that is, having two commutative monoid structures $(0, +)$ and $(1, \times)$ related by the distributive laws $0 = a0$ and $ab + ac = a(b + c)$. Examples abound, *e.g.* $\mathbb{N}$ and $\mathbb{N}/(1 + 1 \sim 1)$, whose modules are commutative monoids and, respectively, sup-semilattices. Other examples include the rig of isomorphism classes of vector bundles on a space, or of finitely generated projective modules over a commutative ring, under direct sum and tensor product. While it is customary to reflect rigs into rings by tensoring with $\mathbb{Z}$, it is by no means always desirable to ignore the extra information contained in the rig. (Steenrod remarked that much of his mathematics came from analyzing the information that others had deliberately discarded by performing such identifications; their "garbage", he called it.) Of most importance for us is the Burnside rig of isomorphism classes of objects in any distributive category (defined below).

3. Euler and counting.

Undeterred by the proof that there are no negative sets, Euler proceeded to find them, in his analysis of the formula $V - E + F = 2$ for the numbers of vertices, edges, and faces in suitable polyhedra. While some later accounts focus on this "Euler characteristic" as a topological invariant, we wish to emphasize instead the irrelevance of topology, and treat the Euler characteristic of a polyhedron rather as a finitely additive measure. Roughly, Betti numbers (ranks of homology groups) depend on how a space is pieced together, but Euler characteristic doesn't; if a space is a disjoint union of two parts, the Euler characteristics add.

Euler's analysis, which demonstrated that in counting suitably "finite" spaces one can get well-defined negative integers, was a revolutionary advance in the idea of cardinal number — perhaps even more important than Cantor's extension to infinite sets, if we judge by the number of areas in mathematics where the impact is pervasive. In any case, it leads us to the desired categories $\mathbb{E}$, which we now describe.

4. Polyhedra and semialgebraic sets.

By a *polyhedron*, (respectively, *semialgebraic set*), we mean a pair n, $P \subset \mathbb{R}^n$, where P is in the boolean algebra generated by subsets of the form $f(x_1, x_2, \ldots, x_n) > 0$, with $f = b + \sum a_i x_i$ (or, respectively, f a polynomial.) These are the objects of a category $\mathbb{P}$ (respectively, $\mathbb{SA}$); a map in the category from $P \subset \mathbb{R}^n$ to $Q \subset \mathbb{R}^m$ is any map of sets whose graph (in $\mathbb{R}^{m+n}$) is a polyhedron (respectively, a semialgebraic set.) We'll treat $\mathbb{P}$ in some detail, and just describe the corresponding facts for $\mathbb{SA}$.

A typical polyhedron in the plane might be the union of the open first quadrant and a line, with finitely many triangles, line segments, and points added or deleted. Any bijective map in $\mathbb{P}$, for example from $(0, 1) \cup \{2\} \cup (3, 4)$ to $(0, 2)$ by $f(t) = t$ for $t \in (0, 1)$, $f(2) = 1$, and $f(t) = t - 2$ for $t \in (3, 4)$, is invertible. Perhaps a better name for $\mathbb{P}$ would be $\mathbb{SL}$, for "semilinear", were it not for the usage requiring semilinear maps to be continuous.

The categories $\mathbb{S}$, $\mathbb{P}$, and $\mathbb{SA}$ are distributive, where $\mathbb{E}$ *distributive* means that $\mathbb{E}$ has finite limits, finite coproducts, and $\mathbb{E}^2 \to \mathbb{E}/(1+1)$ by $(A, B) \longmapsto [(A+B) \to (1+1)]$ is an equivalence. From this it follows that $\mathbb{E}$ satisfies the earlier distributive laws and that $\mathbb{E}/B$ is distributive for any object B in $\mathbb{E}$. (The terminology is not yet standard, with reason: Walters, Cockett, and others have shown that a weaker notion, not requiring all finite limits, is also useful in computer science and elsewhere. The strong notion we use here was suggested by lectures of Lawvere.) In addition these categories are *boolean:* every subobject is a summand; or equivalently, subobjects of P in $\mathbb{E}$ are classified by maps $P \to 1 + 1$. The full subcategory $\mathbb{P}_0 \subset \mathbb{P}$ of *bounded* polyhedra (those which are bounded in $\mathbb{R}^n$) shares all these properties. Our basic task is to calculate the Burnside rigs of these categories and to show their relationship to our original problem. We preface this with some general remarks on Burnside rigs.

5. On Burnside rigs of distributive categories.

The Burnside rig (of isomorphism classes of objects, added by coproduct and multiplied by product) of a distributive category has some special features, the first of which we have already seen.

1) If $a + b = 0$, then $a = b = 0$.

2) If $\sum a_i = \sum b_j$, then there exist c_{ij} such that $\sum_j c_{ij} = a_i$ and $\sum_i c_{ij} = b_j$.

3) If a is connected ($a \neq 0$, and $a = b + c$ implies $b = 0$ or $c = 0$), then a is cancellable ($a + x = a + y$ implies $x = y$).

4) 1 is cancellable (whether it is connected or not; in our examples it is connected).

5) If $ab = 1$, then $a = b = 1$.

Properties (1) and (3) follow from (2), which follows easily from the observation that coproduct decompositions $A = \sum_I A_i$ correspond to maps $A \to 1 + 1 + \ldots + 1$ (I terms). I don't know what additional properties characterize Burnside rigs of distributive categories.

6. The Burnside rig of bounded polyhedra: the open interval as "–1".

To calculate the Burnside rig $\mathcal{B}(\mathbb{P}_0)$ of the category of bounded polyhedra, it turns out that there is only one basic observation needed. The isomorphism class x of the open interval $(0, 1)$ satisfies $x = 2x + 1$, or perhaps better, $x = x + 1 + x$, because

$$(0, 1) = (0, {}^1\!/2) \cup \{{}^1\!/2\} \cup ({}^1\!/2, 1)$$

is a coproduct decomposition. (Recall that maps in our category are not required to be continuous.) Thus while $(0, 1)$ is not "minus one", it comes as close as it can: $0 = x + 1$ is impossible, but $x = 2x + 1$ can be achieved.

Hence the canonical map from the free rig on one generator, $\mathbb{N}[X]$, to $\mathcal{B}(\mathbb{P}_0)$, by $X \longmapsto x$, factors through $\mathbb{N}[X]/(X \sim 2X + 1)$:

$$\mathbb{N}[X]/(X \sim 2X + 1) \longrightarrow \mathcal{B}(\mathbb{P}_0),$$

and I claim this is an isomorphism. Surjectivity is easy, because every bounded polyhedron is a disjoint union of open simplices Δ_n°; and $\Delta_n^\circ \cong (0, 1)^n$. The heart of the matter is the injectivity of our map; and for this we need to introduce two invariants, Euler characteristic and dimension.

For any rig R (recall that all rigs are commutative), define the *Euler characteristic*

$$R \xrightarrow{\chi} E(R)$$

to be the universal map to a rig with additive cancellation. The description of this is well known: $E(R) = R/\sim$, where $r \sim s$ if and only if there is a t with $r + t = s + t$. Similarly, define the *dimension*

$$R \xrightarrow{\dim} D(R)$$

to be the universal map to a rig in which $1 + 1 = 1$ (and hence $x + x = x$). This seems less known: $D(R) = R/\sim$, where $r \sim s$ if and only if $r \leq s$ and $s \leq r$, where $r \leq s$ means that "a finite sum of copies of s can swallow r", *i.e.* there exist a natural number n, and $t \in R$, with $r + t = ns$.

Let $R = \mathbb{N}[X]/(X \sim 2X + 1)$; anticipating a bit, we will call this the *rig of geometric cardinalities*. Now, E(R) and D(R) are easy to calculate; we get $E(R) = \mathbb{Z}$, with $\chi(X) = -1$. Equally simple, if less familiar, is D(R): it is

$$D(R) = \{0 = d^{-\infty},\ 1 = d^0,\ d^1, d^2,\ \ldots\}$$

with $d^i d^j = d^{i+j}$ and $d^i + d^j = d^{\max(i,j)}$. The exponential notation is in keeping with the idea that multiplying polyhedra adds dimensions, while adding gives the maximum of the two dimensions.

To complete the proof that $R = \mathbb{N}[X]/(X \sim 2X + 1) \longrightarrow \mathcal{B}(\mathbb{P}_0)$ is an isomorphism, we need only define a rig homomorphism

$$(\overline{\chi}, \overline{\dim}):\ \mathcal{B}(\mathbb{P}_0) \longrightarrow \mathbb{Z} \times D(R),$$

check that the composite of this with $R \longrightarrow \mathcal{B}(\mathbb{P}_0)$ is $(\chi, \dim)$, and show that $(\chi, \dim)$ is injective. This last is a simple induction, after noting that $(\chi, \dim)(f(X)) = (f(-1), \text{degree } f)$. So the definitions of $\overline{\chi}$ and $\overline{\dim}$ need attention. One defines these, at an object P, by writing P as a disjoint union of atoms A in the boolean algebra given by a hyperplane decomposition of space, $P = \cup_{\text{atoms } A \subset P} A$, and then setting

$$\overline{\chi}(P) = \Sigma(-1)^{\mathrm{gdim}(A)} \qquad \text{and} \qquad \overline{\dim}(P) = d^{\sup \mathrm{gdim}(A)},$$

where gdim(A) is the ordinary geometric dimension of the atom A. It's easy to check that adding a hyperplane leaves these quantities unchanged, and then that they're isomorphism-invariant.

Summing up: a geometric cardinality can be identified with an equvalence class of polynomials with natural number coeficients — two such being equivalent if they have the same degree and the same value at -1 — or with an isomorphism class of bounded polyhedra. As we'll see shortly, it is also an isomorphism class of semialgebraic sets, or of finitely subanalytic sets, and is an equivalence class of constructible sets (the boolean closure of the class of algebraic sets in $\mathbb{C}^n$).

7. The Burnside rig of unbounded polyhedra.

The major themes of this paper can be understood without the corresponding (and somewhat more cumbersome) calculation for $\mathbb{P}$, the category of all polyhedra. Nevertheless, since $\mathbb{P}$ is of interest in connection with linear programming and related matters, we give a sketch of the necessary changes to convert the calculation for $\mathbb{P}_0$ to that for $\mathbb{P}$. There are two generators: $x = (0, 1)$, as before, and $y = (0, \infty)$. We easily get three relations: $x = 2x + 1$, $y = x + 1 + y$ because $(0, \infty) = (0, 1) \cup \{1\} \cup (1, \infty)$, and $y^2 = 2y^2 + y$ because $(0, \infty)^2 = \{(r, s) \mid r < s\} \cup \{(r,s) \mid r = s\} \cup \{(r, s) \mid r > s\}$. So the rig $\overline{R}$ presented by these maps to $\mathcal{B}(\mathbb{P})$:

$$\overline{R} = \mathbb{N}[X, Y]/(X \sim 2X + 1, Y \sim X + 1 + Y, Y^2 \sim 2Y^2 + Y) \longrightarrow \mathcal{B}(\mathbb{P}),$$

and we claim this is an isomorphism. As before, we calculate:

$$E(\overline{\mathrm{R}}) = \mathbb{Z}\times\mathbb{Z}, \text{ with } \chi(X) = (-1,-1) \text{ and } \chi(Y) = (-1,0).$$

(The first relation gives, after cancellation, $0 = X + 1$, so we get $\mathbb{Z}$ as subrig; then the second relation becomes vacuous, while the third gives that $Y + 1$ is idempotent in $E(\overline{R})$.) Again, $D(\overline{R})$ is less familiar. An element of $D(\overline{R})$ is a finitely generated (hence finite) order-ideal in the partially ordered set of monomials X^iY^j, ordered as a monoid with $1 < X < Y$; so $X^iY^j \le X^pY^q$ means $j \le q$ and $i+j \le p+q$. These order ideals are multiplied by multiplying elementwise and down-closing, and are added by union. (Note that $D(\overline{R})$ could have been given by a similar description, using the poset $\{1 < X < X^2 < ....\}$.). It is worth noting that both $D(\overline{R})$ and $D(R)$ have multiplicative cancellation: for $a \neq 0$, $ab = ac$ implies $b = c$. This will show that any (bounded) polyhedron with cancellable (bounded) Euler characteristic is multiplicatively cancellable.

Checking the surjectivity of $\overline{R} \to \mathcal{B}(\mathbb{P})$ by $X \longmapsto (0,1) = x$ and $Y \longmapsto (0,\infty) = y$ is harder than before, but not much. We must show that every polyhedron P is isomorphic to a sum of monomials $(0,1)^i \times (0,\infty)^j$. For this, we show that P can be decomposed into pieces each linearly (rather than just piecewise-linearly) isomorphic to $\Delta_i^\circ \times (0,\infty)^j$. This is done by induction on the geometric dimension of P, and P can be supposed to be an atom in a decomposition of $\mathbb{R}^n$ by hyperplanes; but it is important to first ensure that the family of hyperplanes includes at least n that are independent, *i.e.* the linear functionals f in the equations $f(x) = c$ are linearly independent. By induction, each face of the atom P can be suitably decomposed; and then one decomposes P by choosing any point p in P and taking the open *truncated cones* consisting of the points $tp + (1-t)x$ where $0 < t < 1$ and x ranges over any of the parts into which the faces of P have been decomposed. The truncated cone on $\Delta_i^\circ \times (0,\infty)^j$ is $\Delta_{i+1}^\circ \times (0,\infty)^j$. These cones do not exhaust P, but what's left is an infinite closed cone with vertex p; and decomposing its (bounded) intersection with a suitable hyperplane into open simplices cuts this cone into a sum of powers of y. The whole proof is thus quite parallel to the proof that bounded polyhedra can be decomposed into disjoint open simplices by decomposing the boundary, picking a point inside, and "coning"; the only new ingredient is that in the unbounded case there is still a cone left over, which one proves can also be decomposed as a sum of parts linearly isomorphic to monomials in y.

To map $\mathcal{B}(\mathbb{P})$ to $E(\overline{R}) \times D(\overline{R})$ turns out to be a bit easier than one might expect; the clue is that each atom in a decomposition of $\mathbb{R}^n$ (by at least n independent hyperplanes, as before) is in fact (polyhedrally isomorphic to) a monomial x^iy^j. The sum $i+j$ is just the geometric dimension of the atom, while j is that of the cone in $\mathbb{R}^n$ obtained by intersecting the closed half spaces given by the hyperplanes through the origin parallel to the faces of P. It's easier not to prove this at this stage; for now, it shows how to define the map from $\mathcal{B}(\mathbb{P})$ to $E(\overline{R}) \times D(\overline{R})$. The proof that this map is well-defined, *i.e.* unchanged by adding another hyperplane and isomorphism-invariant, is straightforward; and the rest goes just as before, with just a little more care in the inductive argument to show that $\overline{R} \to E(\overline{R}) \times D(\overline{R})$ is injective.

A geometric description of the "Euler characteristic" $\chi(P) = (m, n)$ and "dimension" $\dim(P) = F \subset \{X^iY^j\}$ is now not difficult. First, m is the "expected" Euler characteristic, since $x = (0, 1)$ and $y = (0, \infty)$ are "alike", except that there is no piecewise-linear isomorphism between them. (They are semi-algebraically isomorphic, by $t \longmapsto t^{-1} - 1$. This will be useful later.) Second, n is the "bounded Euler characteristic" of $P \subset \mathbb{R}^n$, *i.e.* $\chi(P \cap C)$ for any sufficiently large closed cube $C = [-B, B]^n$. The dimension of $P \subset \mathbb{R}^n$ is just the set of monomials $x^i y^j$ which are subobjects of P (linearly, if you want.) For example, for a geometrically 2-dimensional polyhedron, the possible dimensions, in increasing order, are the order ideals generated by x^2, by x^2 and y, by xy, and by y^2, exemplified by, respectively, a 2-simplex, the union of a 2-simplex and a ray, an infinite strip bounded by parallel lines, and the plane.

8. Related examples.

Modulo well-known facts, it is easy to check that the Burnside rig of the category $\mathbb{SA}$ of semialgebraic sets is the same as that for $\mathbb{P}_0$, the rig R of geometric cardinalities. The main ingredient is Hironaka's theorem of semialgebraic triangulability of semialgebraic sets.

Thus the inclusions of distributive categories

$$\mathbb{S} \hookrightarrow \mathbb{P}_0 \hookrightarrow \mathbb{P} \hookrightarrow \mathbb{SA}$$

(the first two of which are full) give, on passage to Burnside rigs,

$$\mathbb{N} \longrightarrow R \longrightarrow \overline{R} \longrightarrow R,$$

exhibiting, geometrically, R as a retract of $\overline{R}$ (by $y \longmapsto x$).

The geometric interpretation of the injectivity, for each of these rigs $A = \mathbb{N}, R, \text{or } \overline{R}$, of $A \to E(A) \times D(A)$, should be clear. For instance, for polyhedra, it says: if P and Q are polyhedra which are "cancellation equivalent" (*i.e.* $P + T \cong Q + T$ for some polyhedron T) and "comparable" (*i.e.* $P \le nQ$ and $Q \le mP$ for some natural numbers m, n), then $P \cong Q$. I do not know a proof of this, in any of the categories $\mathbb{P}_0$, $\mathbb{P}$, $\mathbb{SA}$, which does not use essentially the entire calculation sketched above! A similar remark applies to proving that $2P \cong 2Q$ implies $P \cong Q$. For these, it might be helpful if one could find a simpler characterization of those rigs A for which $A \to E(A) \times D(A)$ is injective; that is, those in which $a + t = b + t$ & $a + s = nb$ & $b + r = ma$ implies $a = b$.

One trivial generalization: everything said about $\mathbb{P}_0$ and $\mathbb{P}$ remains valid if the reals are replaced by any ordered field; for instance, the Burnside rig is unchanged. (One should note that not every order-convex subset of the line is a polyhedron; for instance with $\mathbb{Q}$ as the field, $(0, \sqrt{2})$ is not polyhedral, since it is not defined by finitely many inequalities with rational coefficients.) More interesting is the category $\mathbb{FA}$ of finitely subanalytic sets, which van den Dries has shown shares enough of the properties of $\mathbb{SA}$ that our calculation also gives R, the rig of geometric cardinalities, as $\mathcal{B}(\mathbb{FA})$. All our categories satisfy the *axiom of choice:* every epimorphism has a section; but $\mathbb{FA}$ and $\mathbb{SA}$ satisfy the stronger *generic triviality* theorem: every map $A \to B$ is isomorphic to a coproduct of product projections, $\sum B_i \times F_i \to \sum B_i$. This is false in $\mathbb{P}$, and even in $\mathbb{P}_0$, as the example of the map from the open triangle with vertices (0, 0), (0, 1), and (1, 1) to the interval (0, 1) by projection on the first coordinate demonstrates. However, any map $A \to B$ in $\mathbb{P}_0$ is a coproduct of maps $A_i \to B_i$ in which the isomorphism class of the fiber is constant; and the

Euler characteristic behaves as if such a map were a product: $\chi(A_i) = \chi(B_i)\chi(\text{fiber})$. This observation suggests a reduced Burnside rig in our next example.

Genuinely different is the category $\mathbb{CS}$ of *constructible sets:* an object is a subset of $\mathbb{C}^n$ in the boolean algebra generated by zero-sets of polynomials; a map is a function with a constructible graph. Essentially by construction, this category is distributive and boolean, though without axiom of choice; but its Burnside rig is complicated. A reduced Burnside rig $\mathcal{B}_{red}(\mathbb{CS})$ and rig homomorphism $\mathcal{B}(\mathbb{CS}) \to \mathcal{B}_{red}(\mathbb{CS})$ are defined to be universal for rig homomorphisms π with domain $\mathcal{B}(\mathbb{CS})$ satisfying: whenever $A \to B$ is in $\mathbb{CS}$ and $\pi(\text{fiber})$ is constant, $\pi(A) = \pi(B)\pi(\text{fiber})$. This gives rise to the rig of geometric cardinalities again! The generator X is the twice-punctured complex plane $\mathbb{C}\setminus\{0, 1\}$. To get the desired relation $\pi(X) = 2\pi(X) + 1$, consider $Y = \mathbb{C}\setminus\{0, 1, -1\}$, and note that $X = Y + 1$ while all fibers of the squaring map $Y \to X$ have two points. The proof that X generates is inductive, projecting the constructible set on a coordinate hyperplane; and the injectivity of the map from R to our reduced rig is proved by using the forgetful functor from $\mathbb{CS}$ to $\mathbb{SA}$, viewing $\mathbb{C}^n$ as $\mathbb{R}^{2n}$. (This calculation is related to work of van den Dries, Marker, and Martin, "Definable equivalence relations on algebraically closed fields", *J. Symbolic Logic* 54 (1989) 928-935.) It is suggestive that in this example as in the earlier ones, the "minus 1" object comes from the basic bipointed parameter object for homotopies: $\{0, 1\} \to \mathbb{C}$ for algebraic geometry, respectively $\{0, 1\} \to [0, 1]$ for topology, by deleting the two marked points.

Applications of these ideas to geometry will have to be treated on another occasion; some work by Beifang Chen on curvature measures along these lines will appear in *Advances in Mathematics*. Also postponed are the analysis of colimits and of the relation of a boolean distributive category to its "Gaeta topos", following ideas of Lawvere which have exerted a continuing influence on the shape of the work described here.

I wish to express my gratitude to the organizers of the conference, and particularly to Aurelio Carboni and many colleagues at the University of Milan, for their warm and generous hospitality.

This paper is in final form and will not be published elsewhere.