Lawvere Archive Posthumous Publication – February 9, 2025

The Role of History in Toposes in Geometry and Analysis*

F. William Lawvere

Talk for 'Journée Houzel', Institut Henri Poincaré, Paris

November 23, 2007

It is a very great honor to speak for Professor Houzel. Even though I had not seen him for 26 years, I often think of him and I am happy to see him again. It was a pleasure to follow the last lecture, since it was its subject, that had originally attracted me to consult with Professor Christian Houzel. I was impressed by the first talk today to change my topic a little. Of course, I knew that Professor Houzel is a very respectable historian of mathematics, and I think it is important to speak more about the role of the history of mathematics within mathematics itself. I have spoken before about the role of pedagogy in the progress of mathematics. For example, Bourbaki was primarily a pedagogical project, one of historical dimensions, which led to an incredible amount of new and very valuable mathematical research that partly grew out of the struggle to explain mathematics, as it existed, to wider circles of people. Similarly, there is a role for history of mathematics. It is the fact that there are latent ideas. They are invented by the collective, and at a certain point the collective is coming closer to it, and it is made explicit by somebody who is the discoverer. This is the frequent form of conceptual progress.

*Delivered at the celebration of Christian Houzel's 70th Birthday; edited transcript by Bill. See also the different prepared lecture 'Toposes in Geometry and Analysis'

Vito Volterra

Trying to discern, for example, the historical origins of the notion of cartesian closed category, started from a polemic launched by Jean Dieudonné against (of all people) Vito Volterra. He claimed that Volterra could not possibly have done any functional analysis because he did not phrase it in terms of topological vector spaces. Looking into Volterra's earliest publications, I discovered answers to many questions. Volterra gave me the tools to explain the so-called paradox of points in algebraic geometry, and many other areas. I will come to that later. Also, it was Volterra who formulated and proved the so-called Poincaré lemma (Sorry...) long before Poincaré. Thus, looking into history because of current questions and plucking out part of it, permits finding something that nobody else noticed, and put that into mathematical exposition.

Leonhard Euler

Consider Euler. There is the dogma that many histories of mathematics still repeat that 'Euler was not rigorous', 'and therefore', continue these current philosophies, 'we should not be rigorous either, we should go for beauty and not rigor'. 'Rigor is a bad thing, rigor destroys creativity, etc. etc.'

To the contrary, I claim that Euler was rigorous, or at least that should be the presumption in undertaking the study of someone like Euler. Of course, I have not read all the many volumes, but Umberto Bottazzini showed me where to find the particular thing that I suspected would be relevant to my current concerns: the particular claim made by Euler that real numbers are ratios of infinitesimals. Everyone has been taught to say: 'Of course, that is nonsense.' But in fact, it is a good definition of real numbers, because it leads to direct proofs of some of the properties of real numbers that don't follow very easily from Dedekind's definition, for example. (I just wanted to post it here, I will explain later.)

Georg Cantor

I was led to look into Cantor, because of something that occurred in the present. I was reading

Cantor while riding in a train, but noticed something I had never seen before, even though I had heard many lectures by set theorists and glanced through many books by historians about the person Cantor. Something I had never heard of caught my eye. The train was going to Zurich, so I immediately went to my friend Ernst Specker, a very substantial set theorist, and pointed this out to him. He did not know of it either. We went to the ETH library and sure enough, the same point was in the book available there. The point was Cantor's attribution of his idea of equivalence of sets to the geometer Jakob Steiner. No set theorists had ever mentioned that, but Cantor himself did quite explicitly. Specker turned out to be an expert on Steiner, because Steiner was a Swiss mathematician (who worked in Berlin). Specker commonly used to give popular lectures on Steiner, but he had not noticed this either. Was it the lack – in the prevailing dogmas – of any mention of such a connection that rendered it invisible? This led me to investigate more closely what Cantor considered to be Steiner's discovery. That discovery was something that the set theorists have not analyzed or developed but that in my view should lead to a reasonable answer to their so-called problem about the continuum hypothesis.

There are a couple of other examples of this sort: starting from the present, being led to look into the past, and then finding something quite surprising that is not part of the common knowledge, but nonetheless is relevant to current mathematical research.

Alexander Grothendieck

I was led to the 1960 Cartan Seminar because of a commonplace repeated in books and papers (for example, in Wikipedia in an entry called Origin and Genesis of Topos Theory); it says that topos theory came out of the idea of generalized space, that toposes are generalized spaces. Of course, Grothendieck had the petit étale topos and that was a really significant example of a generalized space, but it was not the only origin of topos theory and you can see in the 1960 seminar that a completely different kind of topos was discussed there, which was not a generalized space in any real sense. It was not yet called a topos, but in fact it was.

Étendues and adequately separable toposes

I have arrived at a tentative account of the distinction. A generalized space involves a $\mathcal{U}$-topos that has a site with no idempotents. Of course, the ordinary classical topological spaces, when viewed as toposes, have posets as sites. Posets have no idempotents, because in fact they have no endomorphism at all, but immediately important along with the classical topological spaces, was the topos of G-sets where G is a group. From the algebraic topology of the 30's we see that clearly groups and spaces ought to belong in the same category, since there is a map from a space to its fundamental group whose kernel is the universal covering space; there is a diagram which is taking place in some category in which both groups and spaces are on equal footing. Grothendieck's first attempt to explain what that category is, isolated his special toposes known as étendues; these are locally topological spaces and this includes groups. As it turned out they included quite a bit more, because étendues have a site consisting only of monomorphisms, and conversely.

Peter Johnstone

On the other hand, there is another special class of toposes in which the separable objects are adequate. Peter Johnstone first studied them explicitly, (unless someone tells me otherwise) under the name of QD, quotient of decidables. A separable object is one in which the diagonal has a complement. In any topos it has a Heyting complement, but that is not usually a real Boolean complement in the sense that the union is not the whole square. Separability requires that the square be the sum of the diagonal and its complement. If all objects in a topos were separable, then it would be Boolean and very special. But often those special objects are adequate in the sense that every object in the topos is an inductive limit of these. That is, the QD class, interesting for us in this regard, because the petit étale topos of a scheme is like that. If among the connected étale objects you have two maps between two of them, which are equal somewhere, even on a very small part, then they are equal. That says that the map from the small part to the domain is an epimorphism; thus indeed all maps are epimorphisms. As it turns out the adequately separable $\mathcal{U}$-toposes have sites consisting entirely of epimorphisms, and conversely. In the site just as a category in itself, every map is epic; of course, when we embed the site in the topos, not everything is epi, but just within its own little universe one has the universal cancellation property. Both of these are good classes of categories to use as sites, since we can exponentiate them arbitrarily, and any category has such a category as reflection, etc. There is a common generalization that is pretty obvious, a bi-cancellation property, meaning that we assume two things in order to conclude an equality. This suggested my still more general property that there exists a site in which there are no idempotents. That is a very reasonable property, but it has not been very much studied. It seems to be a notion of generalized space that is still not a wholesale generalization, but does capture the relevant examples.

By contrast, Grothendieck's implicit toposes in 1960 that later became much more explicit in the form of the gros Zariski and gros étale toposes, have an opposite property, mainly that they have to have degeneracies.

I think Grothendieck's calling the figures 'points' was too general, because points are intuitively more special. I use the term 'figure' which everybody sort of understands.

Volterra used the term 'elements' which was very good at the time, because the idea that elements are something irreducible is certainly not part of ordinary language: to say that a window pane is an element of the window is an element of the room, is certainly a usual way of speaking of such structural elements. Volterra's use of this term was quite correct then, and would be even now, except that set theory has made us accustomed to the irreducibility presumption. A figure of shape A is just a map with domain A; if A is special, the figure might deserve the name of point. A map from one figure to another is an arrow making a commutative triangle. Let's call this an incidence relation. An incidence relation could say, for example, that a certain curve lies on a certain surface. There is thus the slice category of figures and incidence relations in each given space. I normally use it when the shapes of the figures are limited, for example, to objects of a specific site. This simple-minded construction refutes a very common rumor, namely that in category theory you cannot get inside the objects. The objects are said to be 'opaque'. But it is actually the best theory of how we can get inside, because the geometry of figures and incidence relations is the inside;

we should perhaps even officially use the word ‘inside’ to mean this slice category. If we have a way of picturing the A’s themselves, then we have a way of picturing $\mathcal{X}$ as consisting of crisscrossing and interlocking A’s. It may or may not be an adequate picture. [Schapira agrees and says there is also the ‘outside’]. Again, if we have a special class of objects, we may want to give a special name to maps whose codomains are in the class, but with arbitrary given domain. Of course, they should be called functions. Functions have always meant something a little more special than maps or morphisms, even though sometimes we identify the terms. But, on the other hand, function theory uses morphisms to study special morphisms. In the same spirit we can take the ‘outside’ of a geometrical object to be its algebra of functions.

The slice category $\mathcal{X}/\mathcal{A}$ is the algebra of functions on $\mathcal{X}$ because the morphisms are algebraic operations. Note that if $\mathcal{A}$ had products, these operations would include addition and multiplication and so forth, since A could be, for example, the square of another A'. The simple definition of the inside immediately implies that, given a general morphism, it will induce a functor, actually a morphism of discrete fibrations, (the shape of a figure is given by the fibration, and similarly, the type of a function is given by an op-fibration). The induced functor is always continuous, in the sense that it maps figures into figures without tearing the incidence relations.

(What else could continuity mean?) Sequential continuity in topological spaces is a specific example of this, because one could take as figure type the well-known generic convergent sequence, along with its single limit point. In the category of spaces generated by that, continuity means transforming convergent sequences into convergent sequences, considered as figures into other figures without tearing the limit away from the rest of the sequence. The inclusion of the point as the limit point is the crucial incidence relation. Dually, of course, we have a homomorphism in the opposite direction between function algebras. (All these theorems are cases of associativity).

Adequacy, as a general concept, was defined by Isbell (that is, adequacy of a choice of subcategory of figure shapes) to mean that every morphism of discrete fibrations (or every natural transformation of presheaves) comes from an actual map, in other words, the embedding of the ambient category into the presheaves on the adequate part is full and faithful.

Exponentials and Infinitesimals

Within this general framework, why is it that real numbers are ratios of infinitesimals? For that one needs a very important further ingredient, namely the idea of an exponential functor; categories that have this are often called cartesian-closed. This is perhaps the most fundamental ingredient of mathematical content going back 300 years, the idea that given two objects, the morphisms between them in some sense also form an object, that they also have a geometrical structure with their own figures and incidence relations, and their own algebra of functions, now called functionals. They were used for a long time without a name (that gave Frege the license to claim to have invented them in the context of truth valued functions); but the real valued functionals are much more profound. Volterra made this concept precise, then his good friend Hadamard gave it the name 'functionals' as we know. Fréchet and Hadamard pursued this development in a way that did not depend on topological vector spaces. The defining property of the exponential is that for every pair X, Y of objects we have another object, such that for every third object A the morphisms from A to that, correspond naturally to maps from $A \times X$ to Y.

Daniel M. Kan

The exponential is an instance of the right adjoint. Kan's notion of adjoint functors appeared over 50 years ago; in particular, Kan pointed out this example himself; he knew about it from the context of simplicial sets. Specifically, the functor $(\ )^X$ is right adjoint to $X \times (\ )$. Many properties of the map spaces come from adjointness. The fact is, that it was in the early stages of the calculus of variations. Right from the start the needed idea was to explain what is a figure of shape 'an interval' in a map space. Why do we care about that? One was talking about problems of least descent, etc. in which the variable under discussion was infinite-dimensional, and yet one wanted to vary it. So what is a variation? A variation is precisely a path in the map space, where the domain of the path is an interval, that is, a variation is a figure of one-dimensional shape inside a map space.

How can one deal with that? One can deal with it because it is the same thing as a figure of one more variable. This was the technique of the calculation in the calculus of variations from the beginning, and Volterra made this into the definition of analytic functionals. (In the category of complex analytic spaces we might replace the interval by a disk.) Thus one has the possibility of talking about functionals being smooth or being analytic.

What is a smooth functional? It's a morphism of the associated geometries. If I have a functional that is analytic, I take an analytic map like $A \to Y^X$ (actually, given by $A \times X \to Y$, functions of several complex variables are presumed defined); if I substitute this into my functional, I should get again something equally analytic. That is the condition. In some way it is all far simpler than the whole dogma of topological vector spaces would suggest. You can make different definitions of open sets inside the same category, with the automatic property that all morphisms are continuous, by having, for example, some representing object, like Sierpinski space (it could be something more sophisticated in the analytic context). The inverse image under any map could be called open. Then, of course, every map f in the category is continuous in the sense that the further inverse image along f is again of that form.

But it is complicated to get at the open set structure of a map space, knowing the open set structure of the domain and codomain spaces. The covariant structure given by the geometry of figure is taken as the basic measure of the cohesion of a space, rather than the contravariant structure of an algebra of functions, or of an algebra of open sets.

Ratios of Infinitesimals

I want to tell you why real numbers are ratios of infinitesimals, as I promised. In teaching calculus we make the mistake of talking about difference quotients, thus pretending that quotients exist in the same sense as do addition and multiplication. Actually, it is a deeper matter. If we interpret every statement about quotients as a statement by saying that there exists an x such that $ax = b$, i.e. referring back to a question about multiplication, we get something that can be computed correctly. This eliminates automatically a lot of confusion, for example, in the proof of the chain rule. That dividing, or even inverting, is a non-trivial process is seen, in fact, that the theory of localization of rings requires a whole technology, concerning how to take a ring and invert some elements, pass to open subsets, and so on. Dividing is non-trivial, especially dividing two things, not just multiplying something by the inverse of something else, which is the best kind of dividing if you can get it; in general, we can't even get that. Basically, a ratio is a process by which you transform something into something else. Of course, a process with special properties, but still it's a morphism in its own right. Thus the claim is that reals are transformations of infinitesimals into infinitesimals, that is, they form part of the map space D^D, where D is a space of infinitesimals. In other words, I am going to identify the reals with a subspace of a map space. Then these ratios have the natural intrinsic multiplication, which is just composition of transformations. This is sometimes called synthetic differential geometry, because we start with nothing but a pointed object, and produce the algebra out of the mere category operations. (In fact, the infinitesimals D have only one point, so even that could be considered a property, rather than a structure.) But relativizing to arbitrary figure shapes, we still have to put the condition that zero is preserved. The reals R form the submonoid of D^D that is the kernel of the evaluation map. In a suitable ambient category, every monoid has a universal commutative monoid associated to it. I take the commutative reflection of the monoid D^D and put the condition that the composite (with its inclusion of R) is an isomorphism. This implies that in R itself the multiplication is commutative. Note that endomorphism space is always non-commutative, unless the object is one point. This is a general fact, (not something particular about D.) Moreover, following the projection by the inverse, we see that R is actually a retract as a monoid. This object D is supposed to represent the tangent bundle in the sense that X^D is the tangent bundle of X and evaluating at zero is the bundle map. Thus R is the tangent space at zero of D itself. Taking induced maps amounts to taking derivatives of arbitrary maps. Of course, in all the specific examples we know, D has the concrete nature of the spectrum of the dual numbers. However, one can approach it axiomatically by purely synthetic properties of D, some of which I have written down, and deduce in particular that R has not only multiplication, but also addition, which makes R into a ring object. I can take R as a codomain for function algebras, or I can take R as a figure shape and thus speak of paths.

Further, I can form the Lie algebras of group objects and deduce the extensive nature of distributions. Any such category, i.e. a topos equipped with a suitable D, has a natural notion of 'distribution with compact support', the Hom space carved out by equations from a double map space. The point about R being additive is that the distributions on a disjoint sum of two spaces is uniquely given by a pair of distributions.

A standard definition of 'open' is $\mathrm{Aut}(D)$ included in D^D.

Thus Functional analysis has a meaning in any such topos.

Thank you for your patience. (Applause)

Pierre Schapira: Thank you for the categorical approach to analysis, which I appreciate. Are there any questions? I'm not a historian, but something that is very important now is 'correspondence'. We don't consider a function from one space to another one, but kernels. Maybe Volterra was the first to use it systematically? ;)

Chargois: I would like to say that Grothendieck in 1960 discussed not only toposes, but also stacks.

Schapira: so, maybe we take the transition to the next talk [by Chargois].
Deux minutes!

We thank Professor Francisco Marmolejo with his green thumb that makes his work flourish. We appreciate your effort in getting the formatting just right.

The Archive Family