The Tannakian formalism and the motivic Galois group

Avi Zeff

1. Introduction

Consider a nice group G (in practice an affine algebraic group over a field k). The category of k-representations $\mathrm{Rep}_k(G)$ has various good properties: it is symmetric monoidal under the tensor product of representations, abelian, semisimple for G nice enough, and satisfies some other good properties making it "rigid." All of these properties are lifted from the "underlying" category $\mathbf{Vect}_k$ along the forgetful functor $\omega : \mathrm{Rep}_k(G) \to \mathbf{Vect}_k$. It turns out that the data of the abstract category $\mathrm{Rep}_k(G)$ with the various structure associated to it as above (symmetric monoidal, abelian, rigid) together with the faithful and exact functor $\omega : \mathrm{Rep}_k(G) \to \mathbf{Vect}_k$ is enough to recover the group G; this is the Tannakian formalism.

One could then ask: suppose we have some arbitrary category C with suitable properties/structure, and a suitable functor $\omega : C \to \mathbf{Vect}_k$. If we apply the reconstruction formalism to the pair (C, ω), if they satisfy the necessary formal properties we can hope to "recover" a group G, even though C did not necessarily originally arise as a representation category; then $C \simeq \mathrm{Rep}_k(G)$, by doing the same process to G. When this is possible for some ω we say that C is Tannakian.

In particular, we know that if we choose the numerical equivalence relation, then the category of motives $\mathbf{Mot}_k$ over k is semisimple abelian, and so we might hope that in fact it is Tannakian, corresponding to some interesting group GMot_k. We could then hope to use this to study motives as representations of this group, similar to how $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is used in number theory; indeed GMot_k is called the motivic Galois group.

2. Rigid symmetric monoidal categories

I'll skip much discussion of symmetric monoidal categories on the theory that these are reasonably familiar by now; if not, the theory mostly consists of writing down obvious statements in complicated ways with big diagrams, have fun. Let's say a little bit about hom objects, though.

Given a symmetric monoidal category, the contravariant functor

$$T \mapsto \mathrm{Hom}(T \otimes X, Y)$$

may be representable, in which case we call its representing object $\underline{\mathrm{Hom}}(X, Y)$. By definition, this is right adjoint to the tensor product, since

$$\mathrm{Hom}(T, \underline{\mathrm{Hom}}(X, Y)) = \mathrm{Hom}(T \otimes X, Y).$$

There is therefore an evaluation map

$$\mathrm{ev}_{X,Y} : \underline{\mathrm{Hom}}(X, Y) \otimes X \to Y$$

adjoint to the identity $\underline{\mathrm{Hom}}(X, Y) \to \underline{\mathrm{Hom}}(X, Y)$. If $\underline{\mathrm{Hom}}$ exists for (X, Y), (Y, Z), and (X, Z), then there is a composition morphism

$$\underline{\mathrm{Hom}}(X, Y) \otimes \underline{\mathrm{Hom}}(Y, Z) \to \underline{\mathrm{Hom}}(X, Z)$$

adjoint to the composition

$$\underline{\mathrm{Hom}}(X,Y)\otimes\underline{\mathrm{Hom}}(Y,Z)\otimes X\xrightarrow{\mathrm{ev}_{X,Y}}\underline{\mathrm{Hom}}(Y,Z)\otimes Y\xrightarrow{\mathrm{ev}_{Y,Z}}Z.$$

Observe that there is an isomorphism

$$\underline{\mathrm{Hom}}(Z,\underline{\mathrm{Hom}}(X,Y))\xrightarrow{\sim}\underline{\mathrm{Hom}}(Z\otimes X,Y)$$

adjoint to the composition

$$\underline{\mathrm{Hom}}(Z,\underline{\mathrm{Hom}}(X,Y))\otimes Z\otimes X\xrightarrow{\mathrm{ev}_{Z,\underline{\mathrm{Hom}}(X,Y)}}\underline{\mathrm{Hom}}(X,Y)\otimes X\xrightarrow{\mathrm{ev}_{X,Y}}Y,$$

since the corresponding functors are

$$T\mapsto\mathrm{Hom}(T\otimes Z,\underline{\mathrm{Hom}}(X,Y))=\mathrm{Hom}(T\otimes Z\otimes X,Y)$$

and

$$T\mapsto\mathrm{Hom}(T\otimes Z\otimes X,Y).$$

If $\mathbf{1}$ is the (an) identity object, then

$$\mathrm{Hom}(\mathbf{1},\underline{\mathrm{Hom}}(X,Y))=\mathrm{Hom}(\mathbf{1}\otimes X,Y)=\mathrm{Hom}(X,Y).$$

Define $X^\vee=\underline{\mathrm{Hom}}(X,\mathbf{1})$ when it exists. There is then an evaluation map $\mathrm{ev}_X: X^\vee\otimes X\to\mathbf{1}$. This gives a contravariant functor: if $X^\vee$ and $Y^\vee$ exist and $f:X\to Y$ is a morphism, there is a unique morphism f^t making the diagram

$$\begin{array}{ccc} Y^\vee\otimes X & \xrightarrow{f^t\otimes\mathrm{id}} & X^\vee\otimes X \\ \downarrow{\scriptstyle \mathrm{id}\otimes f} & & \downarrow{\scriptstyle \mathrm{ev}_X} \\ Y^\vee\otimes Y & \xrightarrow{\mathrm{ev}_Y} & \mathbf{1} \end{array}$$

commute.

Whenever $X^\vee$ and $X^{\vee\vee}=(X^\vee)^\vee$ exist, there is a morphism $X\to X^{\vee\vee}$ adjoint to $X\otimes X^\vee\xrightarrow{\sim}X^\vee\otimes X\xrightarrow{\mathrm{ev}_X}\mathbf{1}$. When this morphism is an isomorphism, we say that X is reflexive.

There is a morphism

$$\underline{\mathrm{Hom}}(X_1,Y_1)\otimes\underline{\mathrm{Hom}}(X_2,Y_2)\to\underline{\mathrm{Hom}}(X_1\otimes X_2,Y_1\otimes Y_2)$$

adjoint to the morphism

$$\underline{\mathrm{Hom}}(X_1,Y_1)\otimes\underline{\mathrm{Hom}}(X_2,Y_2)\otimes X_1\otimes X_2\to_1\otimes Y_2$$

given by the tensor of the evaluation maps, whenever all of the hom objects exist. In the setting of nice symmetric monoidal categories such as $\mathbf{Vect}_k$ or indeed $\mathrm{Rep}_k(G)$, this turns out to be an isomorphism; for example, if $Y_1=X_2=\mathbf{1}$, this is the morphism $X_1^\vee\otimes Y_2\to\mathrm{Hom}(X_1,Y_2)$, which we would certainly like to be an isomorphism. Thus if we hope that our category looks like the representation category of some group we need this isomorphism to always hold, as well as the above two.

Definition 1. A symmetric monoidal category is called rigid if $\underline{\mathrm{Hom}}(X, Y)$ exists for every pair of objects X, Y, ever object is reflexive, and the above morphism is always an isomorphism.

In this case, the functor $C^{\mathrm{op}} \to C$ given by $X \mapsto X^\vee$ and $(f : X \to Y) \mapsto (f^t : Y^\vee \to X^\vee)$ is an equivalence (of symmetric monoidal categories).

For every object X in a rigid symmetric monoidal category, we have maps

$$\underline{\mathrm{Hom}}(X, X) \xrightarrow{\sim} X^\vee \otimes X \xrightarrow{\mathrm{ev}_X} \mathbf{1},$$

where the existence of all objects and the fact that the first map is an isomorphism use the axioms of rigidity. Applying $\mathrm{Hom}(\mathbf{1}, -)$ gives a map

$$\mathrm{Tr}_X : \mathrm{End}(X) \to \mathrm{End}(\mathbf{1}),$$

called the trace morphism. Applied to the identity, we get the rank: $\mathrm{rank}(X) = \mathrm{Tr}_X(\mathrm{id}_X)$. One can check that

$$\mathrm{Tr}_{X\otimes Y}(u \otimes v) = \mathrm{Tr}_X(u) \cdot \mathrm{Tr}_Y(v),$$

since the trace map comes from the tensor product of evaluation maps, which on $\mathrm{End}(\mathbf{1})$ is just multiplication, so in particular $\mathrm{rank}(X \otimes Y) = \mathrm{rank}(X)\,\mathrm{rank}(Y)$ and $\mathrm{rank}(\mathbf{1}) = \mathrm{id}_{\mathbf{1}}$.

The explanation for the name "rigid" comes from the following fact.

Proposition 2. *Suppose that $F, G : C \to D$ are functors of symmetric monoidal categories, and that C and D are rigid. Then any natural transformation of symmetric monoidal functors $\lambda : F \to G$ is an isomorphism.*

Proof. Such a natural transformation λ is a collection (λ_X) of compatible morphisms $\lambda_X : F(X) \to G(X)$. Applying this to $X^\vee$, the compatibility of F and G with the tensor structure implies that $F(X^\vee) \simeq F(X)^\vee$ and similarly for G, so there exists a unique morphism μ_X^t making the diagram

$$\begin{array}{ccc} F(X^\vee) & \xrightarrow{\lambda_{X^\vee}} & G(X^\vee) \\ \downarrow \simeq & & \downarrow \simeq \\ F(X)^\vee & \xrightarrow{\mu_X^t} & G(X)^\vee \end{array},$$

and the corresponding family of morphisms (μ_X) assembles to a natural transformation $\mu : G \to X$ which is inverse to λ. □

For categories of representations, we want our underlying categories to also be abelian, and we can extend the theory in the natural way: since all our functors should respect the abelian structure, we just require the underlying category to be abelian and $\otimes$ to be bi-additive (or over a field k, k-bilinear). This automatically makes it compatible with direct and inverse limits, and therefore exact (this follows since it has adjoints in both C and its opposite).

Proposition 3. *In a rigid abelian tensor category, $\mathbf{1}$ is a simple object if its endomorphism ring is a field.*

Proof. Let U be a subobject of $\mathbf{1}$, and let V be the cokernel of $U \hookrightarrow \mathbf{1}$ so that we have an exact sequence

$$0 \to U \to \mathbf{1} \to V \to 0.$$

Since the tensor product is exact, tensoring with $U \hookrightarrow \mathbf{1}$ gives a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U \otimes U & \longrightarrow & U & \longrightarrow & V \otimes U & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & U & \longrightarrow & \mathbf{1} & \longrightarrow & V & \longrightarrow & 0 \end{array}$$

with both rows exact. The composition $U \to V \otimes U \to V$ is trivial; since $U \to V \otimes U$ is surjective, it follows that $V \otimes U \to V$ must be trivial. All of the vertical maps are injective since they arise from tensoring with an injection, so it follows that $V \otimes U = 0$, so $U \otimes U = U$ as a subobject of $\mathbf{1}$.

Let $U^\perp$ be the kernel of the dual morphism $\mathbf{1} \to U^\vee$. Tensoring our exact sequence with $U^\perp$ gives

$$0 \to U^\perp \otimes U \to U^\perp \to U^\perp \otimes V \to 0.$$

For any object T, the tensor product $T \otimes U$ is 0 if and only if the map $T \to U^\vee \otimes T$ given by tensoring with $\mathbf{1} \to U^\vee$ is zero; this is because $T \otimes U \hookrightarrow T$ and $\mathrm{Hom}(T \otimes U, T) = \mathrm{Hom}(T, U^\vee \otimes T)$, so this injection is mapped to this map $T \to U^\vee \otimes T$. The kernel of $T \to U^\vee \otimes T$ is $U^\perp \otimes T$, so in particular since $U \to U^\vee \otimes U$ is an isomorphism (adjoint to $U \otimes U \simeq U$) we have $U^\perp \otimes U = 0$ and therefore the above exact sequence gives $U^\perp \simeq U^\perp \otimes V$. Similarly since $V \otimes U = 0$ and $U^\vee \otimes V = \underline{\mathrm{Hom}}(U, V) = 0$, the kernel of $V \to U^\vee \otimes V$ is all of V and so $U^\perp \otimes V \simeq V$ and therefore $U^\perp \simeq V$. This gives V as both the cokernel of $U \hookrightarrow \mathbf{1}$ and the kernel of $\mathbf{1} \twoheadrightarrow U^\vee$, which gives a splitting of our short exact sequence, so $\mathbf{1} = U \oplus U^\perp$, so that the endomorphism ring cannot be a field unless $U = 0$ or $U = \mathbf{1}$. □

Some examples: $\mathbf{Vect}_k$ is a rigid abelian symmetric monoidal category, with $\mathrm{End}(\mathbf{1}) = k$.

For any commutative ring R, this generalizes to $\mathbf{Mod}_R$, with $\mathrm{End}(\mathbf{1}) = R$; this is an abelian symmetric monoidal category. However it is not always rigid: not every R-module is necessarily reflexive. If we restrict to projective modules, it is rigid, but only additive, not abelian.

The main example we are concerned with is $\mathrm{Rep}_k(G)$, where G is an affine group scheme over a field. In this case $\mathrm{Rep}_k(G)$ is a rigid abelian tensor category with $\mathrm{End}(\mathbf{1}) = k$.

3. Tannakian formalism

Let's explore this example further. In this situation we have a forgetful functor $\omega : \mathrm{Rep}_k(G) \to \mathbf{Vect}_k$, which preserves all the structure. We can define a functor $\mathrm{Aut}^\otimes(\omega)$ sending a k-algebra R to the set of R-linear automorphisms of ω compatible with the tensor product structure, i.e. families of R-linear G-automorphisms λ_X of $X \otimes R$ for every object $X \in \mathrm{Rep}_k(G)$, such that $\lambda_{X_1 \otimes X_2} = \lambda_{X_1} \otimes \lambda_{X_2}$ and $\lambda_{\mathbf{1}} = \mathrm{id}_R$. Every $g \in G(R)$ defines an element of $\mathrm{Aut}^\otimes(\omega)(R)$; we can view G as a functor of k-algebras sending R to the image of G in $\mathrm{Aut}^\otimes(\omega)(R)$.

Proposition 4. *The inclusion $G \to \mathrm{Aut}^\otimes(\omega)$ is an isomorphism of functors of k-algebras.*

In other words, it is possible to recover G from the data of $(\mathrm{Rep}_k(G), \omega)$.

Proof. For each $X \in \mathrm{Rep}_k(G)$, restricting to representations generated by X and its dual $X^\vee$ gives $\mathrm{Aut}^{\otimes}(\omega)(R)$ restricted to these representations as a subgroup of $\mathrm{GL}(X \otimes R)$, with G_X the image of G under this restriction given as a further subgroup. By the G-equivariance requirement, $\mathrm{Aut}^{\otimes}(\omega)$ restricted to X is just the group fixing the tensors fixed by G_X, which is a determining property of G_X, so they are equal; taking the limit over all X gives the proposition. □

This assignment to $(\mathrm{Rep}_k(G), \omega)$ of the functor $\mathrm{Aut}^{\otimes}(\omega)(R)$ can be done for any pair (C, ω) satisfying similar properties, which we can now pin down: we need C to be rigid abelian symmetric monoidal with $\mathrm{End}(\mathbf{1}) = k$, and $\omega : C \to \mathbf{Vect}_k$ to be an exact, faithful, k-linear functor of symmetric monoidal categories. The hope is that this functor $\mathrm{Aut}^{\otimes}(\omega)$ will be representable by some affine group scheme; this is in fact the case.

Theorem 5. *With the notation above,* $\mathrm{Aut}^{\otimes}(\omega)$ *is representable by an affine group scheme* G *over* k*, and there is an equivalence of categories* $C \to \mathrm{Rep}_k(G)$*.*

This is proven by constructing a coalgebra A from the k-linear category C, equip it with an algebra structure using the tensor structure on C, and show that $G = \mathrm{Spec}\, A$ is a group scheme (i.e. has inverses in addition to the product from the coalgebra structure) using the rigidity of C. Once the representability is established, it is clear that the resulting functor is an equivalence: the map from pairs (C, ω) to the corresponding functor $\mathrm{Aut}^{\otimes}(\omega)$ is injective up to equivalence.

When C is a rigid abelian symmetric monoidal category for which there exists such a functor ω, we say that C is a neutral Tannakian category, and ω is a neutralization; thus we can express Theorem 5 as stating that every neutral Tannakian category is of the form $\mathrm{Rep}_k(G)$ (possibly in multiple ways), and the category of neutralized Tannakian categories over k, i.e. pairs (C, ω), is equivalent to the category of affine group schemes G over k (with inverse $G \mapsto \mathrm{Rep}_k(G)$). The semisimplicity of C corresponds to G being proreductive.

Examples: consider the category of graded vector spaces (V_n) with finite-dimensional sum $V = \bigoplus V_n$. This is a rigid abelian symmetric monoidal category, and the functor ω sending (V_n) to V is a fiber functor neutralizing the category, so this is the representation category of some G. This is easy to describe: it is for $G = \mathbb{G}_{\mathrm{m}}$ over k, acting by $\lambda \mapsto \lambda^n$.

Similarly, the category of real Hodge structures is Tannakian and has the same neutralization over $\mathbb{R}$, with representing group the Deligne torus $\mathbb{S} = \mathrm{Res}_{\mathcal{C}/\mathbb{R}}\, \mathbb{G}_{\mathrm{m}}$.

Non-example: consider the category of $\mathbb{Z}/2$-vector spaces, with graded commutative tensor product given by a sign. This is a rigid abelian symmetric monoidal category. However, in this case the categorical trace (i.e. rank) of a vector space is not its dimension but the graded sum $\dim V^0 - \dim V^1$, which may be negative. But the rank of a representation will always be nonnegative, so this cannot be equivalent to $\mathrm{Rep}_k(G)$ for any G.

4. General theory

In general, we may have Tannakian categories which are not neutral. What does this mean?

We say that a category is Tannakian if it has a fiber functor $C \to \mathbf{Vect}_K$ for any extension K over k. The neutral case is when $K = k$.

In general, when C is Tannakian but not neutral, we replace the group G by a gerbe over the category of affine schemes over k. I don't understand this very well and won't spend time on it, but roughly this is a stack which behaves like a G-torsor. For a suitable fiber functor H with values in K, taking the fiber of this gerbe at H should give a proreductive group over K.

5. Motives

Let's finally come back to motives. We discussed last time how the notion of numerical equivalence is the best one for these sorts of categorical properties: it uniquely makes the category of motives $\mathbf{Mot}_k$ into a semisimple abelian category, which is naturally symmetric monoidal, and it is not too hard to check that it is rigid, with $\mathrm{End}(\mathbf{1}) = \mathbb{Q}$. We might therefore hope that it is Tannakian.

However, this fails, fortunately for stupid reasons similar to those we've seen before. The rank (categorical trace) of a motive is its Euler characteristic, which may be negative, e.g. for curves of genus $g > 1$; as for superspaces, this is impossible for anything of the form $\mathrm{Rep}_k(G)$. However, it is easily fixed: we need to change the tensor product structure by a sign, which makes the rank the *sum* of the Betti numbers. In order to make this make sense on the level of motives, we need to assume standard conjecture C. We call this modified category $\widetilde{\mathbf{Mot}}_k$.

The next problem is in finding suitable fiber functors. We have many functors to various vector space-like categories which respect all the structure of motives: these are Weil cohomology theories. However, by definition homological equivalence is the finest relation through which we expect these to nicely factor! Thus to have any hope of being Tannakian at least in a useful way we also need to assume standard conjecture D, equivalence of numerical and homological equivalence.

It turns out that this is enough, due to work of Janssen. In particular in the case where k is algebraic over a finite field, or for abelian varieties, conjecture C is known and so $\widetilde{\mathbf{Mot}}_k$ is unconditionally Tannakian.

Let's also assume conjecture D, so that we can take a Weil cohomology theory with values in K as our fiber functor. This gives us a proreductive group $\mathrm{GMot}_{H,k}$ as a fiber of the gerbe GMot_k associated to the Tannakian category $\widetilde{\mathbf{Mot}}_k$. We can also do this for various (thick rigid) subcategories, which will give various quotients of the motivic Galois group: for example, taking the category generated by some motive $h(X)$ of finite type gives the motivic Galois group of X, $\mathrm{GMot}_{H,k}(X)$.

For example, we can look at Artin motives, generated by the motives of finite extensions of k. In this case the motivic Galois group is just the regular absolute Galois group G_k.

We can also look at pure Tate motives $\mathrm{GMot}_k(\mathcal{L})$, generated by the Tate motive $\mathcal{L}$ (or the projective line); this is an algebraic group, in fact $\mathbb{G}_\mathrm{m}$. We could also combine these with Artin motives, which just gives the product of these groups.

The comparison theorems between different cohomology theories corresponds (assuming conjecture D) to isomorphisms between various motivic Galois groups for different fiber functors.

In characteristic p, the category of motives is still Tannakian, but provably not neutral so long as k contains $\mathbb{F}_{p^2}$: if K is contained in either $\mathbb{R}$ or $\mathbb{Q}_p$, any fiber functor would have to contain the endomorphisms of a supersingular elliptic curve, which are not contained in either.

In general, we have a short exact sequence

$$1 \to \mathrm{GMot}_{\bar{k}} \to \mathrm{GMot}_k \to G_k \to 1,$$

where $G_k = \mathrm{Gal}(\bar{k}/k)$. This last term corresponds to motives of zero-dimensional varieties, i.e. Artin motives, and is discrete, while $\mathrm{GMot}_{\bar{k}}$ is connected and so is the connected component of GMot_k.

Conjecture C gives a weight grading on $\mathbf{Mot}_k$, i.e. a functor to graded vector spaces; on the group side of the equivalence, this is a central homomorphism $w : \mathbb{G}_\mathrm{m} \to \mathrm{GMot}_k$. On the other hand, there is a canonical map $t : \mathrm{GMot}_k \to \mathbb{G}_\mathrm{m}$ corresponding to the Lefschetz motive, with $t \circ w = 2$.

Over finite fields, it turns out that $\mathbf{Mot}_k$ (up to numerical equivalence) is generated by Artin motives and motives of abelian varieties, so everything is fairly well-behaved. The essential image of the ℓ-adic realization is the set of ℓ-adic representations of G_k whose eigenvalues are Weil numbers. In particular the (connected component of the) motivic Galois group is essentially determined by the Galois action on Weil numbers, and in particular is abelian, so motives over finite fields should be reasonably manageable.

For the number theorists: the Langlands program relates representations of G_k in some Lie group to representations of a dual group. One could also expand the Galois side to the whole motivic Galois group (wildly conjecturally); the corresponding Tannakian group on the automorphic side is the Langlands group, which is even bigger (so on the motivic side something even larger may be needed).