ZARISKI CANCELLATION

PEDRO NÚÑEZ

Abstract. Following [Hoc72] we provide an example of rings (commutative and with 1 throughout the talk) R, R' such that $R[t] \cong R'[t]$ but $R \ncong R'$, where t is an indeterminate. As a preparation for this counterexample we also discuss the notion of projective module and the hairy ball theorem.

Contents

Many thanks to Oliver Braunling for his help while preparing this talk!
—parts in gray will be omitted during the talk—

1. Introduction

Let R be a ring. We can form the polynomial ring $R[t]$ in one variable t with coefficients in R. This construction is functorial, and hence

$$R \cong R' \Longrightarrow R[t] \cong R'[t].$$

The goal of this talk is to show with an explicit counterexample due to Hochster [Hoc72] that the converse is not true.

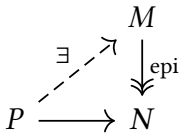
In the process of constructing this counterexample we will come across a projective module which, as a consequence of the hairy ball theorem, is not a free module. Therefore we will discuss projective modules and the hairy ball theorem before jumping into the counterexample.

Date: 4th November 2020.

The author gratefully acknowledges support by the DFG-Graduiertenkolleg GK1821 "Cohomological Methods in Geometry" at the University of Freiburg.

2. Projective modules

Definition 2.1. Let C be a category. An object $P \in C$ is called *projective* if the following lifting problem can always be solved:



Lemma 2.2. *Let $\mathcal{A}$ be an abelian category and let $P \in \mathcal{A}$ be an object. The following are equivalent:*

(1) P is projective.
(2) $\mathrm{Hom}(P, -)$ is exact.
(3) Every short exact sequence of the form

$$0 \to N \to M \to P \to 0$$

splits.

Proof. We start with (1) $\Rightarrow$ (2). Assume P is projective and consider a short exact sequence

$$0 \to A \to B \to C \to 0.$$

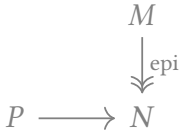
Since $\mathrm{Hom}(P, -)$ is always left exact, we only need to show that the induced map $\mathrm{Hom}(P, B) \to \mathrm{Hom}(P, C)$ is surjective. But $B \to C$ is an epimorphism, so this is precisely what P being projective means by definition.

Next we show (2) $\Rightarrow$ (3). Assume $\mathrm{Hom}(P, -)$ is exact and consider a short exact sequence

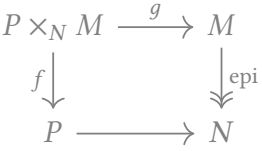
$$0 \to N \to M \to P \to 0.$$

Applying $\mathrm{Hom}(P, -)$ we get a surjection $\mathrm{Hom}(P, M) \to \mathrm{Hom}(P, P)$, and the identity on P comes then from the desired section $\sigma \colon P \to M$.

The implication (3) $\Rightarrow$ (1) is left as an exercise during the talk. The hint is that epimorphisms are stable under pullback in abelian categories, and the solution follows in gray. We are given the following situation:



All finite limits exist in $\mathcal{A}$, so we may consider the cartesian square



Epimorphisms are stable under pullback in abelian categories, so f is also an epimorphism. By assumption, we can find a section $\sigma\colon P \to P \times_N M$ splitting the corresponding short exact sequence. The composition $g \circ \sigma\colon P \to M$ is then the desired lift. □

Let us look now at the abelian category of modules over a ring R. What does it mean for an R-module to be projective?

Proposition 2.3 ([Fra18, Lemma 1.1.2])**.** *Let R be a ring. An R-module P is projective if and only if it is a direct summand of some free module F. Moreover, if P is finitely generated, then we can also choose F to be finitely generated.*

Proof. If P is (finitely generated) projective, consider a surjection $F \twoheadrightarrow P$ from a (finitely generated) free module F. The resulting short exact sequence

$$0 \to K \to F \to P \to 0$$

splits by Lemma 2.2, hence P is a direct summand of F.

Suppose conversely that $F = P \oplus K$ and consider a surjection $\varphi\colon M \twoheadrightarrow N$ and a morphism $f\colon P \to N$. Then $\varphi \oplus \mathrm{id}_K\colon M \oplus K \twoheadrightarrow N \oplus K$ is again a surjection. Since F is free, the morphism $f \oplus \mathrm{id}_K\colon F \to N \oplus K$ can be lifted to a morphism $\bar{f} \oplus \mathrm{id}_K\colon F \to M \oplus K$, so that $\bar{f}\colon P \to M$ is the desired lifting of the original surjection $\varphi\colon M \twoheadrightarrow N$. □

Corollary 2.4 ([Fra18, Cor. 1.1.28])**.** *Let R be a ring. A finitely presented R-module P is projective if and only if $\mathrm{Ext}^1_R(P, T) = 0$ for every finitely generated R-module T.*

Proof. Pick a presentation

$$0 \to K \to F \to P \to 0$$

with K finitely generated and F finitely generated and free. Then take the long exact sequence of $\mathrm{Ext}^\bullet_R(P, -)$ and use the assumption to find a splitting of the short exact sequence, exhibiting therefore P as a direct summand of a free module. □

To close this section, let us briefly discuss the relation between being projective and other well-known module properties. Let R be a ring. Recall that an R-module M is called *flat* if the functor $M \otimes_R (-)$ is exact, and it is called *torsion-free* if $0 \in M$ is the only element $m \in M$ such that there exist a non-zero divisor $r \in R$ with $rm = 0$. Then we have the following inclusions:

$$\{\text{Free}\} \quad \subseteq \quad \{\text{Projective}\} \quad \subseteq \quad \{\text{Flat}\} \quad \subseteq \quad \{\text{Torsion-free}\}$$

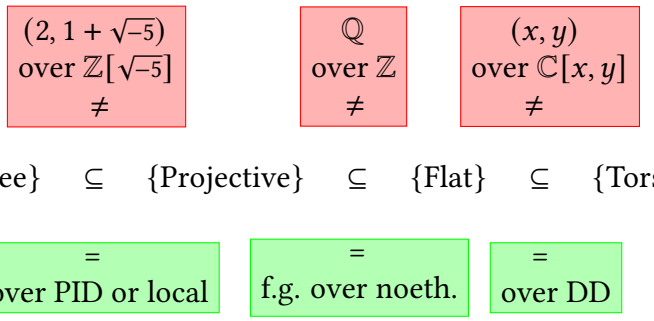
The axiom of choice is equivalent to the statement that every set is a projective object in the category of sets. This implies in turn that every free module is projective. Indeed, this follows then from the universal property of free modules. More generally, if a functor R preserves epimorphisms and $L \dashv R$, then L preserves projective objects [Fra18, Dual of Fact 1.1.1]. In our case R would be the forgetful functor and L would be the functor sending a set to the free module over this set.

The fact that projective modules are flat follows again from the axiom of choice, but in this case it is a strictly weaker statement. What we need now is the existence of enough projective objects in the category of R-modules, so that we can compute the left-derived functors of $M \otimes_R (-)$ with projective resolutions and argue as in [Fra18, Fact 1.2.1]. In order to ensure this, it would suffice to have enough projective objects in the category of sets.

Finally, that every flat module M is torsion-free follows from the computation in [Fra18, Example 1.2.1]. Namely, if $r \in R$ is not a zero-divisor, then

$$0 = \operatorname{Tor}_1^R(M, R/rR) \cong \ker(M \xrightarrow{r\cdot(-)} M).$$

A natural question at this point could be: to what extent are the previous inclusions of sets of R-modules strict? In the following diagram we give some examples in which these inclusions are strict, shown in red above the corresponding inclusion. We also point out under what assumptions we do get an equality, shown in green below the corresponding inclusion.



In the above diagram, DD stands for Dedekind domain [Neu99, Definition I.3.2], PID stands for principal ideal domain, f.g. stands for finitely generated and noeth. stands for noetherian.

Proof. The first example is discussed in [Gat14, Example 13.8]. Instead of reproducing Gathmann's discussion here, let us argue why the ring $\mathbb{Z}[\sqrt{-5}]$ was a natural place to look for such an examples, which also hints on how to produce similar examples. Dedekind domains are hereditary rings, so

all ideals are projective. One would then try to find a non principal ideal on a Dedekind domain, which then has necessarily a minimal number of two generators [Neu99, Exercise I.3.6]. A Dedekind domain is a principal ideal domain if and only if it is a unique factorisation domain [Gat14, Proposition 13.27], so good candidates to look for non principal ideals are rings of integers of number fields in which we know that there is no unique factorisation into prime elements, such as $\mathbb{Z}[\sqrt{-5}]$.

That every projective module over a principal ideal domain is free follows from Proposition 2.3 and the fact that submodules of free modules over principal ideal domains are again free [Rot02, Theorem 9.8]. That every projective module over a local ring is free is a result due to Kaplansky [AF92, Corollary 26.7], but easier proofs are available in the case of finitely generated modules over a noetherian local ring [Fra18, Proposition 1.3.1].

If R is a Dedekind domain, then every localization $R_\mathfrak{p}$ at a prime ideal $\mathfrak{p} \in \operatorname{Spec}(R)$ is a principal ideal domain. If M is a torsion-free module and $\mathfrak{p} \in \operatorname{Spec}(R)$, then $M_\mathfrak{p}$ is a torsion-free $R_\mathfrak{p}$-module, because localisation is an exact functor. Since flatness can be checked locally [Fra18, Fact 1.2.6], it suffices to argue that torsion-free modules over a principal ideal domain are flat. This follows from the analogue of Baer's criterion for flatness [Fra18, Proposition 1.2.3] and the isomorphisms

$$\operatorname{Tor}_1^R(M, R/rR) \cong \ker(M \xrightarrow{r\cdot(-)} M)$$

for all $r \in R \setminus \{0\}$ mentioned earlier.

That flat implies projective for finitely generated modules over a noetherian ring is a consequence of the local nature of both properties under these assumptions and the corresponding statement for finitely generated modules over a noetherian local ring, see [Fra18, Proposition 1.3.2].

The rationals $\mathbb{Q}$ are torsion-free over the Dedekind domain $\mathbb{Z}$, so they are also flat. Every two rational numbers are linearly dependent over $\mathbb{Z}$, so $\mathbb{Q}$ cannot be a free $\mathbb{Z}$-module. Therefore they cannot be a projective module either, because $\mathbb{Z}$ is a principal ideal domain.

Since $\mathbb{C}[x, y]$ is torsion-free over itself and (x, y) is a submodule, (x, y) is also torsion-free over $\mathbb{C}[x, y]$. Since $\mathbb{C}[x, y]$ is noetherian and (x, y) is finitely generated over $\mathbb{C}[x, y]$, if (x, y) was flat over $\mathbb{C}[x, y]$, then it would also be projective over $\mathbb{C}[x, y]$. By the Quillen–Suslin theorem mentioned below, this would in turn imply that (x, y) is a free $\mathbb{C}[x, y]$, so (x, y) cannot be a flat $\mathbb{C}[x, y]$-module. □

Remark 2.5. The example of $\mathbb{Q}$ as a $\mathbb{Z}$-module also shows that being projective depends on the base ring, since $\mathbb{Q}$ is projective over itself.

Remark 2.6. A ring R is called *hereditary* if all submodules of projective R-modules are again projective. Dedekind domains can be characterised as hereditary integral domains.

Remark 2.7. The example of (x, y) as a $\mathbb{C}[x, y]$-module also shows that $\mathbb{C}[x, y]$ is not an hereditary ring, since $\mathbb{C}[x, y]$ is projective over itself but its submodule (x, y) is not flat, thus not projective.

Remark 2.8. The left-most equality in the diagram above is also true for finitely generated modules over polynomial rings with coefficients in a field. This is an important result first asked by Serre and later proven independently by Quillen and Suslin, see the Wikipedia article on the Quillen–Suslin theorem. It corresponds geometrically to the statement that vector bundles on affine space are all trivial.

Remark 2.9 ([Fra18]). Baer's criterion [Fra18, Prop. 1.1.1] allows us to characterise the dual notion of injective modules only in terms of Ext^1_R and quotients of the ring R by its ideals. Indeed, an R-module N is injective if and only if $\mathrm{Ext}^1_R(R/I, N) = 0$ for every ideal $I \subseteq R$ [Fra18, Prop. 1.1.4]. In general, no similar criterion for projectivity is available. In fact, the *Whitehead problem* —which states that if an abelian group A has $\mathrm{Ext}^1_{\mathbb{Z}}(A, \mathbb{Z}) = 0$, then it is free— is undecidable in ZFC due to a result of Shelah.

On the other hand, the notion of flatness, which agrees with projectivity for finitely generated modules over noetherian rings, does have a general criterion similar to the previous one for injectivity. Namely, an R-module M is flat if and only if $\mathrm{Tor}^R_1(M, R/I) = 0$ for every ideal $I \subseteq R$ [Fra18, Prop. 1.2.3].

3. Hairy ball theorem

In this section we will prove the hairy ball theorem following [EG79].

Theorem 3.1 (Hairy ball theorem)**.** *Every continuous vector field on the 2-sphere $\mathbb{S}^2$ has at least one zero.*

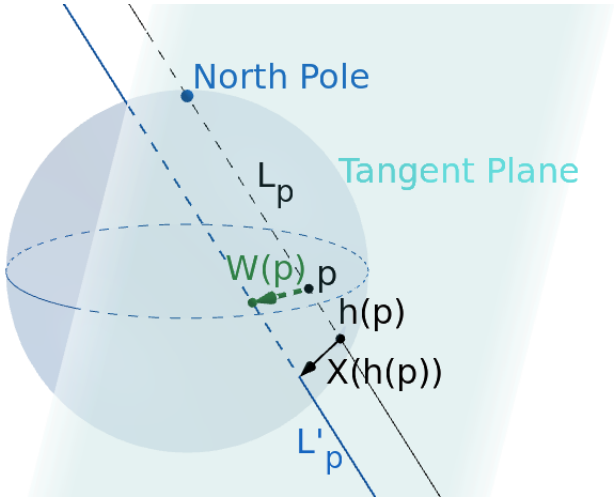
Proof. We will use the following definitions and identifications:

- $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$.
- $\mathbb{R}^2 = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$.
- $\mathbb{D}^2 = \{(x, y, 0) \in \mathbb{R}^2 \subseteq \mathbb{R}^3 \mid x^2 + y^2 \leqslant 1\}$.
- $\mathbb{S}^1 = \{(x, y, 0) \in \mathbb{R}^2 \subseteq \mathbb{R}^3 \mid x^2 + y^2 = 1\}$.
- $\mathbb{S}^2_+ = \{(x, y, z) \in \mathbb{S}^2 \subseteq \mathbb{R}^3 \mid z \geqslant 0\}$.
- $\mathbb{S}^2_- = \{(x, y, z) \in \mathbb{S}^2 \subseteq \mathbb{R}^3 \mid z \leqslant 0\}$.
- For $p = (x, y) \in \mathbb{S}^1 \subseteq \mathbb{R}^2$ we denote by $R_p \colon \mathbb{R}^2 \to \mathbb{R}^2$ the linear reflection with axis the direcction defined by the vector $(-y, x)$, tangent to $\mathbb{S}^1$ at p.

We prove the theorem by contradiction, so suppose that we are given a non-vanishing continuous vector field $X\colon \mathbb{S}^2 \to T\mathbb{S}^2$ on $\mathbb{S}^2$.

Let $h\colon \mathbb{D}^2 \to \mathbb{S}^2_-$ be the inverse of the homeomorphism $\mathbb{S}^2_- \cong \mathbb{D}^2$ induced by the stereographic projection from the north pole $(0,0,1) \in \mathbb{S}^2$. Let $p \in \mathbb{D}^2$ be a point. Let L_p be the line joining $(0,0,1)$ and p, which then intersects $\mathbb{S}^2_-$ precisely at $h(p)$. We consider the line L'_p which is parallel to L_p and passes through the point $h(p) + X(h(p)) \in \mathbb{R}^3$. With this notation, we define a new function

$$\begin{aligned} W\colon \mathbb{D}^2 &\longrightarrow \mathbb{R}^2 \\ p &\longmapsto (\mathbb{R}^2 \cap L'_p) - p. \end{aligned}$$

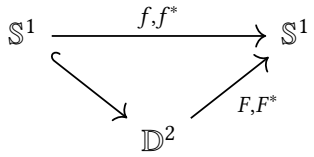


Since X is a non-vanishing continuous vector field, $W\colon \mathbb{D}^2 \to \mathbb{R}^2$ is a non-vanishing continuous function. Therefore it induces a continuous function

$$\begin{aligned} F\colon \mathbb{D}^2 &\longrightarrow \mathbb{S}^1 \\ p &\longmapsto \frac{F(p)}{\|F(p)\|}. \end{aligned}$$

We repeat the same process with the south pole and $\mathbb{S}^2_+$. This gives us a non-vanishing continuous function $W^*\colon \mathbb{D}^2 \to \mathbb{R}^2$ that we can again normalise into a continuous function $F^*\colon \mathbb{D}^2 \to \mathbb{S}^1$.

We denote by f and f^* the restrictions to $\mathbb{S}^1 \subseteq \mathbb{D}^2$ of F and F^* respectively. By definition f and f^* factor through the contractible space $\mathbb{D}^2$:



Therefore, they are both nullhomotopic.

To obtain the desired contradiction, we will next find a certain relation between the functions f and f^*, which is in turn induced by a relation between $W|_{\mathbb{S}^1}$ and $W^*|_{\mathbb{S}^1}$. Namely, given a point $p = h(p) = (x_0, y_0, 0) \in \mathbb{S}^1 \subseteq \mathbb{D}^2$, we claim that

$$W^*(p) = R_p(W(p)). \tag{1}$$

Recall that this means that the vector $W^*(p)$ is the linear reflection of $W(p)$ across the axis given by the direction defined by $(-y_0, x_0, 0) \in \mathbb{R}^2$. In order to prove this, we fix an arbitrary $p = (x_0, y_0, 0) \in \mathbb{S}^1$ and we change the coordinate system, making p the origin, $(x_0, y_0, 0)$ the first basis vector in $\mathbb{R}^2$ and $(-y_0, x_0, 0)$ the second basis vector in $\mathbb{R}^2$. We express $W(p)$ and $W^*(p)$ with respect to these coordinates:

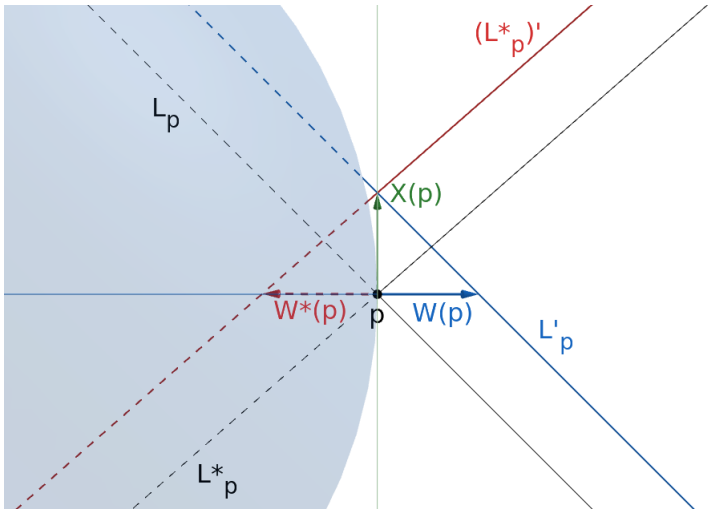
$$W(p) = (w_1, w_2, 0) \text{ and } W^*(p) = (w_1^*, w_2^*, 0).$$

The claim is then that $w_1 = -w_1^*$ and $w_2 = w_2^*$.

We check first that $w_2 = w_2^*$. Let L_p resp. L_p^* denote the line joining p to the north resp. south pole. These two lines intersect at p and define a plane Π which intersects $\mathbb{R}^2$ precisely along the x-axis of our current coordinate system. Let L_p' resp. $(L_p^*)'$ be the lines with directions equal to those of L_p resp. L_p^* containing the point $p + X(p)$. The plane Π' defined by these two lines is then parallel to Π, and therefore the intersection $\Pi' \cap \mathbb{R}^2$ is parallel to the x-axis of our current coordinate system. This implies that $w_2 = w_2^*$.

To check that $w_1 = w_1^*$ we argue projecting the whole picture into the plane spanned by the vectors $(x_0, y_0, 0)$ and $(0, 0, 1)$. From this point of view we see the following picture:

From Thales' theorem we know that L_p and L_p^* intersect at a right angle. From Proposition 29 in Book I of Euclid's *Elements* we deduce that all other angles that seem to be right angles in the picture are in fact right angles. From the fact that the inner angles of any triangle add up to two right angles we deduce in turn that all angles that look like half a right angle in the picture are in fact half a right angle. Indeed, we can start by ensuring that the inner angle between L_p and $W^*(p)$ is half a right angle, which follows from the fact that the triangle with vertices p, the usual origin of $\mathbb{R}^3$ and the north pole is isosceles with a right angle at the usual origin of $\mathbb{R}^3$. Then, knowing already that the inner angle between $W^*(p)$ and $X(p)$ is a right angle, we deduce from this that the inner angle between L_p and



$X(p)$ is also half a right angle, and so on. This implies that all four small triangles in the picture are congruent, and in particular $w_1 = w_1^*$.

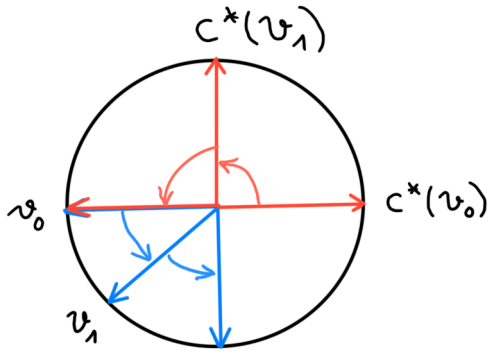
We are finally ready to exhibit the desired contradiction. Let $H\colon \mathbb{S}^1 \times [0,1] \to \mathbb{S}^1$ be a homotopy from f to the constant map $c\colon \mathbb{S}^1 \to \mathbb{S}^1$ with constant value the poitn $(-1,0,0) \in \mathbb{S}^1 \subseteq \mathbb{R}^3$. The formula

$$\begin{aligned} H^*\colon \mathbb{S}^1 \times [0,1] &\longrightarrow \mathbb{S}^1 \\ (p,t) &\longmapsto R_p(H(p,t)) \end{aligned}$$

defines a homotopy between the nullhomotopic map $f^*\colon \mathbb{S}^1 \to \mathbb{S}^1$ and the map

$$\begin{aligned} c^*\colon \mathbb{S}^1 &\longrightarrow \mathbb{S}^1 \\ p &\longmapsto R_p((-1,0,0)) \end{aligned}$$

But this is a contradiction, because $c^*\colon \mathbb{S}^1 \to \mathbb{S}^1$ is the morphism going twice around $\mathbb{S}^1$ at constant speed, which has degree 2.



□

If we take the Whitney embedding theorem [Bre93, Theorem II.10.7], the tubular neighbourhood theorem [Bre93, Theorem II.11.4] and the Lefschetz–Hopf fixed point theorem [Bre93, Theorem IV.23.4] for granted, we can prove a more general statement:

Theorem 3.2 ([Bre93, Corollary IV.23.6])**.** *If M is a compact smooth manifold with Euler characteristic $\chi(M) \neq 0$, then any continuous vector field on M has a zero.*

Proof. We show that if M admits a non-vanishing continuous vector field, then it has Euler characteristic $\chi(M) = 0$. So let $X\colon M \longrightarrow TM$ be a non-vanishing continuous vector field on M.

The Whitney embedding theorem [Bre93, Theorem II.10.7] allows us to assume that $M \subseteq \mathbb{R}^N$ is a compact smooth submanifold. The tubular neighbourhood theorem [Bre93, Theorem II.11.4] ensures the existence of a small enough real number $\varepsilon > 0$ such that the sum in $\mathbb{R}^N$ yields a diffeomorphism

$$\theta\colon \{(x,v) \in M \times \mathbb{R}^N \mid v \perp T_xM, \|v\| < \varepsilon\} \cong \{y \in \mathbb{R}^N \mid \operatorname{dist}(M,y) < \varepsilon\}$$

from the open subset $N_\varepsilon M$ of the normal bundle consisting of normal vectors with norm less than ε to an ε-neighbourhood $B_\varepsilon M$ of M in $\mathbb{R}^N$, which we refer to as a *tubular neighbourhood* of M in $\mathbb{R}^N$. The projection from the normal bundle $\pi\colon NM \longrightarrow M$ induces a smooth strong deformation retraction $r\colon B_\varepsilon M \longrightarrow M \subseteq B_\varepsilon M$ given by $y \mapsto \pi(\theta^{-1}(y))$. The smooth homotopy that shows that r is a smooth strong deformation retraction [Bre93, Definition I.14.8] is

$$\begin{aligned} F\colon B_\varepsilon M \times [0,1] &\longrightarrow B_\varepsilon M \\ (\theta(x,v),t) &\longmapsto \theta(x,tv) \end{aligned}$$

Since M is compact, we may assume—up to multiplying X by a small enough scalar—that $\|X(x)\| < \varepsilon$ for all $x \in M$. We define then

$$\begin{aligned} f\colon M &\longrightarrow M \\ x &\longmapsto r(x + X(x)). \end{aligned}$$

Geometrically, we are projecting the point $x + X(x) \in \mathbb{R}^N$ onto M along the normal direction, in a way that is made precise by the tubular neighborhood theorem.

If $f(x) = x$, then the tangent direction of $X(x)$ is zero. But $X(x)$ was by assumption a tangent vector at $x \in M$, so $X(x) = 0$, contradicting the assumption that X is non-vanishing. Thus f has no fixed points.

We have also shown as a consequence of the Whitney embedding theorem that M is an euclidean neighbourhood retract. Namely, M is a retract

of the open subset $B_\varepsilon M \subseteq \mathbb{R}^N$. Therefore we may apply the Lefschetz–Hopf fixed point theorem [Bre93, Corollary IV.23.5] to deduce that $L(f) = 0$. Recall that the Lefschetz number $L(f)$ is defined as the alternating sum of traces of the morphisms induced by f in homology with coefficients in $\mathbb{Q}$, so it is an homotopy invariant. By scaling down the vectors in our vector field and repeating the process above we obtain a homotopy $f \simeq \mathrm{id}_M$, so $0 = L(f) = L(\mathrm{id}_M)$. The trace of the identity on a vector space is its dimension, so $L(\mathrm{id}_M) = \chi(M) = 0$. □

Remark 3.3. The converse is also true if we further assume orientability and connectedness, see [Bre93, Corollary VII.14.5].

4. Hochster's example

In this section we discuss the example in [Hoc72], which shows that $R \cong R'$ does not necessarily follow from $R[t] \cong R'[t]$.

4.1. **Constructing the relevant rings.** Define $A := \mathbb{R}[x,y,z]/(x^2+y^2+z^2-1)$, which is an integral domain because $x^2+y^2+z^2-1$ is irreducible in $\mathbb{R}[x,y,z]$ and thus $\mathfrak{p} := (x^2+y^2+z^2-1) \subseteq \mathbb{R}[x,y,z]$ is a prime ideal. Let $\phi\colon A^{\oplus 3} \to A$ be the A-linear homomorphism given by

$$(f+\mathfrak{p}, g+\mathfrak{p}, h+\mathfrak{p}) \longmapsto xf+yg+zh+\mathfrak{p}$$

for all $f,g,h \in \mathbb{R}[x,y,z]$. This homomorphism is surjective, because

$$\phi(x+\mathfrak{p}, y+\mathfrak{p}, z+\mathfrak{p}) = x^2+y^2+z^2+\mathfrak{p} = 1+\mathfrak{p}.$$

Consider the short exact sequence of A-modules

$$0 \to E \to A^{\oplus 3} \xrightarrow{\phi} A \to 0,$$

with $E = \ker(\phi)$. Since A is projective over itself, the sequence splits and $A^{\oplus 3} \cong E \oplus A$. Therefore E is a projective module.

Suppose that E was a free module. Non-zero commutative rings have the invariant basis number property, so E would necessarily be a rank 2 free A-module. Let e_1 and e_2 be a basis of E over A, say

$$e_i = (f_{i,1}+\mathfrak{p}, f_{i,2}+\mathfrak{p}, f_{i,3}+\mathfrak{p})$$

for $i \in \{1,2\}$. A section of $\phi\colon A^{\oplus 3} \to A$ is given by

$$\sigma\colon f+\mathfrak{p} \longmapsto (xf+\mathfrak{p}, yf+\mathfrak{p}, zf+\mathfrak{p}).$$

Therefore we can consider a new basis of $A^{\oplus 3}$ given by e_1, e_2 and $(x+\mathfrak{p}, y+\mathfrak{p}, z+\mathfrak{p})$. The matrix transforming the canonical basis of $A^{\oplus 3}$ into this new

basis is given by

$$M+\mathfrak{p} := \begin{pmatrix} f_{1,1}+\mathfrak{p} & f_{2,1}+\mathfrak{p} & x+\mathfrak{p} \\ f_{1,2}+\mathfrak{p} & f_{2,2}+\mathfrak{p} & y+\mathfrak{p} \\ f_{1,3}+\mathfrak{p} & f_{2,3}+\mathfrak{p} & z+\mathfrak{p} \end{pmatrix} \in \mathrm{GL}_3(A).$$

We regard each polynomial in $\mathbb{R}[x,y,z]$ as a continuous function $\mathbb{R}^3 \to \mathbb{R}^3$. Since the value of $f \in \mathbb{R}[x,y,z]$ at a point $(x_0,y_0,z_0) \in \mathbb{S}^2 \subseteq \mathbb{R}^3$ does not depend on the choice of representative modulo $\mathfrak{p}$ and $(x,y,z) \in \mathbb{R}[x,y,z]^{\oplus 3}$ yields an outward pointing vector at each point of $\mathbb{S}^2$, the basis elements $e_1, e_2 \in \ker(\phi)$ give us continuous tangent vector fields on the sphere $\mathbb{S}^2$. Now $\det(M+\mathfrak{p}) \in A^\times$ is a unit, so there is some other $g+\mathfrak{p} \in A$ such that $(g+\mathfrak{p})\det(M+\mathfrak{p}) = 1+\mathfrak{p}$. So any representative $\det(M) \in \mathbb{R}[x,y,z]$ of $\det(M+\mathfrak{p})$ takes non-zero values on any point $(x_0,y_0,z_0) \in \mathbb{S}^2$, because $\det(M)(x_0,y_0,z_0)g(x_0,y_0,z_0) = 1$. In particular, the vector fields e_1 and e_2 are non-vanishing, which contradicts Theorem 3.1. This contradiction shows that E cannot be a free A-module.

If E was generated by only two elements $\alpha, \beta \in E$ over A, then $A^{\oplus 3}$ would be generated by α, β and $(x+\mathfrak{p}, y+\mathfrak{p}, z+\mathfrak{p})$. This would imply that they form a basis of $A^{\oplus 3}$, because the "change of basis" morphism $A^{\oplus 3} \to A^{\oplus 3}$ would be surjective and therefore an isomorphism by [AM69, Exercise 1 in §6]. In particular α and β would be linearly independent over A, making them a basis of E. This would imply that E is a free A-module, but we have seen in the previous paragraph that this is not the case. Therefore E cannot be generated by less than three elements over A.

The next step in the construction of Hochster's counterexample is to consider the symmetric algebras of our A-modules. Recall that there is a canonical A-algebra isomorphism $S(A) \cong A[t]$. Indeed, both A-algebras have the same universal property, because the functor

$$S\colon \mathrm{Mod}(A) \to \mathrm{Alg}(A)$$

from A-modules to (commutative) A-algebras is left adjoint to the forgetful functor. In particular $S(-)$ preserves colimits, so that

$$S(N \oplus N') \cong S(N) \otimes_A S(N')$$

is an isomorphism of A-algebras. Therefore, defining $B := A[\alpha,\beta]$ and $C := S(E)$ we have

$$B[t] = A[\alpha,\beta,t] \cong S(A^{\oplus 3}) \cong S(E \oplus A) \cong S(E) \otimes_A A[t] \cong S(E)[t] = C[t]$$

as A-algebras. In particular, $B[t] \cong C[t]$ as rings.

In order to achieve our goal it remains to show that $B \ncong C$ as rings, which we will show by contradiction. So suppose that $h\colon B \to C$ is a ring isomorphism.

4.2. h **is an $\mathbb{R}$-algebra isomorphism.** Both B and C are $\mathbb{R}$-algebras, because they are A-algebras. In this subsection we make sure that h has to be an $\mathbb{R}$-algebra isomorphism.

Lemma 4.1 ([Swa87, Lemma 9.1])**.** *The only invertible elements of A are (the equivalence classes of) the non-zero real numbers.*

Proof. We first consider the complexification $A_{\mathbb{C}} := A \otimes_{\mathbb{R}} \mathbb{C}$, which remains an integral domain because $x^2+y^2+z^2-1$ remains irreducible over $\mathbb{C}[x,y,z]$. We compute its group of units. Over the complex number, we may rewrite the equation $x^2+y^2+z^2-1=0$ as $uv+z^2-1=0$, where $u=x+iy$ and $v=x-iy$, so that $A_{\mathbb{C}}$ can be rewritten as

$$\mathbb{C}[z,u,v]/(uv+z^2-1).$$

We invert u and all its powers for a moment, and since

$$v=\frac{1-z^2}{u},$$

we obtain

$$(A_{\mathbb{C}})_u \cong \mathbb{C}[z,u,u^{-1}].$$

That is, $(A_{\mathbb{C}})_u$ are Laurent polynomials over the integral domain $\mathbb{C}[z]$. So the units of $(A_{\mathbb{C}})_u$ are precisely of the form λu^m with $\lambda \in \mathbb{C}[z]^\times = \mathbb{C}^\times$ and $m \in \mathbb{Z}$. Since $A_{\mathbb{C}}$ was an integral domain and $u \neq 0$, the localization $A_{\mathbb{C}} \to (A_{\mathbb{C}})_u$ is injective, so $A_{\mathbb{C}}^\times \subseteq (A_{\mathbb{C}})_u^\times$. And $u \notin A_{\mathbb{C}}^\times$, because $A_{\mathbb{C}}/(u) \cong \mathbb{C}[v,z]/(z^2-1) \neq 0$. Therefore $A_{\mathbb{C}}^\times \subseteq \mathbb{C}^\times$. But $\mathbb{C}$ is a field, so $\mathbb{C} \to A_{\mathbb{C}}$ is injective and $\mathbb{C}^\times \subseteq A_{\mathbb{C}}^\times$ as well, hence

$$A_{\mathbb{C}}^\times = \mathbb{C}^\times.$$

Now we consider our $\mathbb{R}$-algebra A again. It is a subalgebra of $A_{\mathbb{C}}$, so $A^\times \subseteq A_{\mathbb{C}}^\times = \mathbb{C}^\times$. Again because $\mathbb{R}$ is a field we also have that $\mathbb{R}^\times \subseteq A^\times$, so it remains to check that conversely every unit in $A^\times$ is the equivalence class of a non-zero real number. Let $a \in A^\times$ and look at its image in $A_{\mathbb{C}}$, which is then the equivalence class of some non-zero complex number $\lambda \in \mathbb{C}^\times$. This element comes from A, so it has to be invariant under the Galois action of $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ on $A_{\mathbb{C}}$, thus has to be the equivalence class of a non-zero real number and $A^\times = \mathbb{R}^\times$. □

We have $C \subseteq C[t] \cong B[t]$ over A, so we can see both B and C as A-subalgebras inside the same A-algebra $B[t] = A[\alpha,\beta,t]$. Since A is an integral domain, the units in $B[t]$ are the invertible elements in A [AM69, Exercise 2 of §1], hence the non-zero real numbers by Lemma 4.1. Therefore the units in B and in C are also the non-zero real numbers, and h and its inverse send real numbers to real numbers. The real numbers have no field automorphisms, so h must be an R-algebra isomorphism.

4.3. **We may assume that h is an A-algebra isomorphism.** Again, both B and C are A-algebras, but it is not clear a priori whether or not h is an A-algebra homomorphism. In fact, in this case it turns out that it need not be one, but we can reduce to the case in which it is.

Definition 4.2 ([Fre17, §10.1.5])**.** An integral domain D is called a *formally real domain* if for all $m \geqslant 1$ and all $a_1, \ldots, a_m \in D$ we have

$$a_1^2 + \ldots + a_m^2 = 0 \Rightarrow a_1 = \ldots = a_m = 0.$$

Let D be a formally real domain. Then any subring of its fraction field is again a formally real domain, because if a sum of squares of fractions is zero then the sum of the squares of the numerators is also zero. It follows by looking at the leading coefficients in a sum of squares that $D[t]$ is also a formally real domain, hence so are all polynomial rings $D[t_1, \ldots, t_m]$. These two observations already imply that A is a formally real domain. Indeed, since $\mathbb{R}$ is a formally real domain, so are the polynomial ring $\mathbb{R}[x, y]$ and its fraction field $\mathbb{R}(x, y)$. The real 2-sphere is birational to the real plane under the stereographic projection, so the fraction field of A is isomorphic to $\mathbb{R}(x, y)$ by [GW10, Lemma 9.33], hence also a formally real domain.

Again by looking at the leading coefficients in a sum of squares we see that any three polynomials $f, g, h \in D[t]$ satisfying $f^2 + g^2 + h^2 = 1$ must be in D, i.e. must be constant polynomials. This observation applied several consecutive times implies that any three elements satisfying such an equation in $A[\alpha, \beta, t] = B[t]$ must be in the subring A. Therefore any three elements in B or in C satisfying such an equation are also in the subring A. As a consequence, $h(A) \subseteq A$. Indeed, since h is an $\mathbb{R}$-algebra homomorphism, it suffices to show that the generators of A over $\mathbb{R}$ land in A, i.e. that $h(x + \mathfrak{p}), h(y + \mathfrak{p}), h(z + \mathfrak{p}) \in A$. This follows from the previous remark and the observation that

$$h(x + \mathfrak{p})^2 + h(y + \mathfrak{p})^2 + h(z + \mathfrak{p})^2 = 1.$$

Similarly, $h^{-1}(A) \subseteq A$. We may precompose h with the automorphism of $B = A[\alpha, \beta]$ given by applying h^{-1} to the coefficients of polynomials. This gives then an A-algebra isomorphism $A[\alpha, \beta] = B \cong C$.

4.4. **Exhibiting the desired contradiction.** Since $C \cong A[\alpha, \beta]$ as A-algebras, C can be generated as an A-algebra by two elements $c, c' \in C$, i.e. $C = A[c, c']$. Since $C = S(E)$ is a graded A-algebra, we may write $c = c_0 + c_1 + \ldots + c_k$ and $c' = c'_0 + c'_1 + \ldots + c'_{k'}$ in terms of their homogeneous components. Then $A[c, c'] = A[c - c_0, c' - c'_0] \cong C$.

Let now $e \in C$ be a homogeneous element in degree 1. We may then write it as

$$\begin{aligned} e &= a_{0,0} + a_{1,0}(c_1 + \ldots + c_k) + a_{0,1}(c'_1 + \ldots + c'_{k'}) + [\ldots]_1 \\ &= a_{0,0} + a_{1,0}c_1 + a_{0,1}c'_1 + [\ldots]_2 \end{aligned}$$

with $[\ldots]_2$ consisting of higher degree homogeneous terms. In the previous equalities, $a_{0,0}$ is an element of $A = S^0(E)$, so it must be zero by our assumption on the degree of e. And again for degree reasons we must have $[\ldots]_2 = 0$ as well. This shows that the degree 1 part of $C = S(E)$ is generated by c_1 and c'_1 over A. But the degree 1 part of C is $S^1(E) \cong E$, so this contradicts our previous conclusion that E requires at least three generators.

Therefore B and C cannot be isomorphic as rings, even though $B[t]$ and $C[t]$ are isomorphic as A-algebras.

5. Historical context

In this subsection we just mention a few key points, refering to the survey [Gup15] for a much more detailed discussion. If A and B are R-algebras, we use $A \cong_R B$ to denote an R-algebra isomorphism from now on.

Zariski originally raised his cancellation problem in 1949 in the following setting:

Question 5.1. Let L and L' be two finitely generated fields over a field k and let $L(x)$ and $L'(x)$ be simple transcendental extensions of L and L' respectively. Does $L(x) \cong L'(x)$ imply $L \cong L'$?

The answer is affirmative only under extra assumptions, but not in general. In the early 70's mathematicians started to think about the analogous problems over rings. To simplify the remaining discussion, let us introduce the following definition:

Definition 5.2. Let k be a field and let A be a k-algebra. A is said to be *cancellative* if for every k-algebra B we have

$$A[x] \cong_k B[x] \Rightarrow A \cong_k B,$$

where x is an indeterminate.

Abhyankar *et al.* showed in 1972 that one dimensional integral domains of transcendence degree one over a field k are cancellative. But in that context Hochster published also his example [Hoc72], showing that not all algebras are cancellative. So the k-algebra version of the question also had a negative answer in general.

Therefore we weaken a bit more the question and ask ourselves:

Question 5.3. Are polynomial rings $k[x_1, \dots, x_n]$ over a field k cancellative?

This is what then became known as the Zariski cancellation problem. The answer is affirmative for $n \in \{1, 2\}$ independently of the base field k. In positive characteristic the answer is negative for any $n \geqslant 3$. Over characteristic zero, the problem is still open for all $n \geqslant 3$.

References

[AF92] Frank W. Anderson and Kent R. Fuller. *Rings and categories of modules*, volume 13 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1992. ↑ 5

[AM69] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969. ↑ 12, 13

[Bre93] Glen E. Bredon. *Topology and geometry*, volume 139 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993. ↑ 10, 11

[EG79] Murray Eisenberg and Robert Guy. A proof of the hairy ball theorem. *Amer. Math. Monthly*, 86(7):572–574, 1979. ↑ 6

[Fra18] Jens Franke. Homological Methods in Commutative Algebra, 2018. Lecture notes by Ferdinand Wagner, available at github.com/Nicholas42/AlgebraFranke. ↑ 3, 4, 5, 6

[Fre17] Gene Freudenburg. *Algebraic theory of locally nilpotent derivations*, volume 136 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, second edition, 2017. Invariant Theory and Algebraic Transformation Groups, VII. ↑ 14

[Gat14] Andreas Gathmann. Commutative Algebra, 2014. Class Notes TU Kaiserslautern 2013/14 available at mathematik.uni-kl.de/~gathmann/class/commalg-2013/commalg-2013.pdf. ↑ 4, 5

[Gup15] Neena Gupta. A survey on Zariski cancellation problem. *Indian J. Pure Appl. Math.*, 46(6):865–877, 2015. ↑ 15

[GW10] Ulrich Görtz and Torsten Wedhorn. *Algebraic geometry I*. Advanced Lectures in Mathematics. Vieweg + Teubner, Wiesbaden, 2010. Schemes with examples and exercises. ↑ 14

[Hoc72] M. Hochster. Nonuniqueness of coefficient rings in a polynomial ring. *Proc. Amer. Math. Soc.*, 34:81–82, 1972. ↑ 1, 11, 15

[Neu99] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder. ↑ 4, 5

[Rot02] Joseph J. Rotman. *Advanced modern algebra*. Prentice Hall, Inc., Upper Saddle River, NJ, 2002. ↑ 5

[Swa87] Richard G. Swan. Vector bundles, projective modules and the K-theory of spheres. In *Algebraic topology and algebraic K-theory (Princeton, N.J., 1983)*, volume 113 of *Ann. of Math. Stud.*, pages 432–522. Princeton Univ. Press, Princeton, NJ, 1987. ↑ 13

Pedro Núñez
Albert-Ludwigs-Universität Freiburg, Mathematisches Institut
Ernst-Zermelo-Strasse 1, 79104 Freiburg im Breisgau (Germany)
Email address: pedro.nunez@math.uni-freiburg.de
Homepage: https://home.mathematik.uni-freiburg.de/nunez