Fermat's Last Theorem

Josua Kugler

July 12, 2022

Contents

1 Introduction

2 An Overview of Wiles' proof

3 Wiles' numerical criterion

Wiles has discovered a criterion for two rings in a specific category to be isomorphic that only depends on some numerical invariants of these rings. The aim of this section is to prove that criterion in its purely algebraic form.

3.1 Preliminaries

Let $\mathcal{O}$ be the ring of integers of a finite extension K of $\mathbb{Q}_\ell$. As K is a local field, its ring of integers is a discrete valutation ring (DVR), i.e. $\mathcal{O}$ is a local, noetherian Dedekind ring with maximal ideal λ. It is complete with resp server usedect to the λ-adic topology, a principal ideal domain (PID) and has residue field $k := \mathcal{O}/\lambda$ to name some properties that we will use in the course of the proof.

$\mathbb{Z}_\ell$ is the ring of integers of $\mathbb{Q}_\ell$ and $\mathbb{F}_\ell = \mathbb{Z}_\ell/\ell\mathbb{Z}_\ell$ its residue field. As $K/\mathbb{Q}_\ell$ is finite, the residue field of $\mathcal{O}$ is a finite extension of $\mathbb{F}_\ell$ and therefore finite.

The categories $\mathcal{C}_\mathcal{O}$ and $\mathcal{C}_\mathcal{O}^\bullet$ In this section, we will mostly deal with very specific rings. Therefore we define the category $\mathcal{C}_\mathcal{O}$ where objects of $\mathcal{C}_\mathcal{O}$ are local complete noetherian $\mathcal{O}$-algebras with residue field k and the morphisms are local $\mathcal{O}$-algebra morphisms. Often, we even need some extra structure. We obtain the category $\mathcal{C}_\mathcal{O}^\bullet$ from $\mathcal{C}_\mathcal{O}$ by equipping an object A with an additional surjective map

$$\pi_A \colon A \twoheadrightarrow \mathcal{O},$$

the so-called augmentation map. Objects in $\mathcal{C}_\mathcal{O}^\bullet$ are often called *augmented rings*. The morphisms in $\mathcal{C}_\mathcal{O}^\bullet$ are local $\mathcal{O}$-algebra morphisms that respect the augmentation map structure, i.e. for a morphism $f \colon A \to B$ we have the commutative diagram

$$\begin{tikzcd} A \arrow[rr, "f"] \arrow[dr, two heads, "\pi_A"'] & & B \arrow[dl, two heads, "\pi_B"] \\ & \mathcal{O} & \end{tikzcd} .$$

In order to state Wiles' criterion, we need some more definitions.

Definition 3.1. $A \in \mathcal{C}_\mathcal{O}$ is *finite flat*, if A is finitely generated and torsion-free as an $\mathcal{O}$-module. Note that $\mathcal{O}$ is a PID and therefore being torsion-free is equivalent to being flat as an $\mathcal{O}$-module.

Definition 3.2 (complete intersection)**.** A finite flat ring $A \in \mathcal{C}_\mathcal{O}$ is called a *complete intersection*, if A is isomorphic as an $\mathcal{O}$-algebra to a quotient

$$A \cong \mathcal{O}[[X_1, \dots, X_n]]/(f_1, \dots, f_n),$$

where there are as many relations as there are variables.

Let's take a look at an example.

Example 3.1. $A = \{(a,b) \in \mathcal{O}\times\mathcal{O},\ a \equiv b \pmod{\lambda^n}\} \cong \mathcal{O}[[T]]/(T(T-\lambda^n))$ is a finite flat complete intersection in $\mathcal{C}_{\mathcal{O}}^{\bullet}$. The projection π_A is given by $\pi_A(a,b) = a$.

Proof. Consider the map

$$\begin{aligned}\phi\colon \mathcal{O}[[T]]/(T(T-\lambda^n)) &\to A\\ f &\mapsto (f(0), f(\lambda^n)).\end{aligned}$$

ϕ is welldefined and respects the $\mathcal{O}$-algebra structure: Let f_0 be the constant term of a polynomial f and $f_1 \coloneqq T^{-1}(f - f_0)$, s.t. $f = f_0 + T \cdot f_1(T)$. Because of

$$f(0) - f(\lambda^n) = (f_0 + 0 \cdot f_1(0)) - (f_0 + \lambda^n \cdot f_1(\lambda^n)) = -\lambda^n \cdot f_1(\lambda^n),$$

$f(0) \equiv f(\lambda^n) \pmod{\lambda^n}$ as required. Furthermore,

$$\phi(T(T-\lambda^n)) = (0(-\lambda^n), \lambda^n(\lambda^n - \lambda^n)) = (0,0).$$

Finally, we need to think about series in $\mathcal{O}[[T]]$ with infinitely many terms. For the first component $f(0)$ this doesn't matter, as ϕ just takes the constant term. As $\mathcal{O}$ is complete with respect to the λ-adic topology, the map $\tilde{\phi}_2\colon \mathcal{O}[[T]] \to \mathcal{O},\ f \mapsto f(\lambda^n)$ is clearly welldefined and thus ϕ is welldefined. Let $o \in \mathcal{O}$. Then

$$\phi(of) = ((of)(0), (of)(\lambda^n)) = (of(0), of(\lambda^n)) = o(f(0), f(\lambda^n)) = o\phi(f)$$

Injectivity: Let $\phi(f) = 0$. Then $f(0) = 0 \implies T|f$ and $f(\lambda^n) = 0 \implies (T-\lambda)|f$. As a result, $f \in T(T-\lambda)$.

Surjectivity: Let $(a,b) \in A$. As $a \equiv b \bmod \lambda^n$, we can write $b = a + b' \cdot \lambda^n$. Because of

$$\phi(\overline{a + b'T}) = (a, a + b'\lambda^n) = (a,b),$$

ϕ is surjective.

$A \in \mathcal{C}_{\mathcal{O}}^{\bullet}$: $\mathcal{O}$ is noetherian, so $\mathcal{O}[T]/(T(T-\lambda^n))$ is noetherian as well. (λ, T) is a maximal ideal in $\mathcal{O}[T]/(T(T-\lambda^n))$, because

$$\left(\mathcal{O}[T]/(T(T-\lambda^n))\right)/(\lambda,T) = \mathcal{O}/(\lambda) = k.$$

Therefore, the completion $\mathcal{O}[T]/(T(T-\lambda^n))^{\wedge(\lambda,T)}$ of $\mathcal{O}[T]/(T(T-\lambda^n))$ with respect to (λ,T) is a local ring with maximal ideal $\widehat{(\lambda,T)}$. Consider the SES of finitely generated $\mathcal{O}$-modules

$$0 \to (T(T-\lambda^n))\mathcal{O}[T] \to \mathcal{O}[T] \to \mathcal{O}[T]/(T(T-\lambda^n)) \to 0.$$

As completion of finitely generated $\mathcal{O}$-modules is exact (because $\mathcal{O}$ is noetherian), we get the SES

$$0 \to (T(T-\lambda^n))\mathcal{O}[[T]] \to \mathcal{O}[[T]] \to \mathcal{O}[T]/(T(T-\lambda^n))^{\wedge(\lambda,T)} \to 0.$$

by completing with respect to (λ, T). As a result, we have

$$\mathcal{O}[T]/(T(T-\lambda^n))^{\wedge(\lambda,T)} = \mathcal{O}[[T]]/(T(T-\lambda^n)).$$

As a result, $\mathcal{O}[[T]]/(T(T-\lambda^n))$ is a local ring with maximal ideal (λ, T). Therefore, its residue field is

$$\mathcal{O}[[T]]/(T(T-\lambda^n))/(\lambda,T) = \mathcal{O}[T]/(T(T-\lambda^n))/(\lambda,T) = \mathcal{O}/(\lambda) = k.$$

As $\mathcal{O}[T]/(T(T-\lambda^n))$ is noetherian, its (λ, T)-completion $\mathcal{O}[[T]]/(T(T-\lambda^n))$ is again noetherian.

In total, we get that $A \cong \mathcal{O}[[T]]/(T(T-\lambda^n))$ is a local, complete, noetherian $\mathcal{O}$-algebra with residue field $k \implies A \in \mathcal{C}_{\mathcal{O}}$.

A **is a finite flat complete intersection:** A is generated by $(1,1)$ and $0, \lambda^n$ because

$$(a,b) = a(1,1) + (0, \underbrace{b-a}_{\in\lambda^n}) = a(1,1) + c(0,\lambda^n).$$

Also, A is torsion-free because $\mathcal{O}$ is an integral domain. As there is one variable and one relation in $A \cong \mathcal{O}[[T]]/(T(T-\lambda^n))$, A is a complete intersection. $\square$

Example 3.2. *$U = \mathcal{O}[[X_1,\dots,X_n]]$ with projection $\pi_U\colon U \to \mathcal{O},\ f \mapsto f(0)$ lies in $\mathcal{C}_{\mathcal{O}}^{\bullet}$.*

Proof. $\mathcal{O}$ is noetherian, so $\mathcal{O}[X_1,\dots,X_n]$ is noetherian as well. $(\lambda, X_1,\dots,X_n)$ is a maximal ideal in $\mathcal{O}[X_1,\dots,X_n]$, because

$$\left(\mathcal{O}[X_1,\dots,X_n]\right)/(\lambda, X_1,\dots,X_n) = \mathcal{O}/(\lambda) = k.$$

Therefore, the completion

$$\mathcal{O}[X_1,\dots,X_n]^{\wedge(\lambda,X_1,\dots,X_n)} = \mathcal{O}[[X_1,\dots,X_n]]$$

of $\mathcal{O}[X_1,\dots,X_n]$ with respect to $(\lambda, X_1,\dots,X_n)$ is a local ring with maximal ideal $(\lambda, \widehat{X_1,\dots,} X_n)$. Its residue field is $\mathcal{O}[X_1,\dots,X_n]/(\lambda, X_1,\dots,X_n) = k$, as required. As $\mathcal{O}[X_1,\dots,X_n]$ is noetherian, its $(\lambda, X_1,\dots,X_n)$-completion is again noetherian. $\square$

Remark 1. *In example 3.1 we could write A as a quotient of $\mathcal{O}[[X]]$. This is possible in a more general setting, in fact every $A \in \mathcal{C}_{\mathcal{O}}$ can be written as a quotient of $U = \mathcal{O}[[X_1,\dots,X_n]]$ for suitable n.*

Proof. As A is a noetherian ring and $\ker \pi_A$ is an ideal in A, it is finitely generated and therefore also finitely generated as an A-module. Consider the map

$$\begin{aligned} \Phi\colon U = \mathcal{O}[[X_1,\dots,X_n]] &\to A \\ X_i &\mapsto a_i, \end{aligned}$$

where $\ker \pi_A = (a_1, \ldots, a_n)$ and π_U is given by $f \mapsto f(0)$. As $(X_1, \ldots, X_n)$ generate the kernel of π_U, this is a map in $\mathcal{C}_{\mathcal{O}}^{\bullet}$. We have the short exact sequences

$$0 \to \ker \pi_A \to A \to \operatorname{im} \pi_A \cong \mathcal{O} \to 0$$

and

$$0 \to \ker \pi_U \to U \to \operatorname{im} \pi_U \cong \mathcal{O} \to 0$$

As both corresponding sequences split via the inclusion $\mathcal{O} \hookrightarrow A$ resp. $\mathcal{O} \hookrightarrow U$, we can write $A \cong \mathcal{O} \oplus \ker \pi_A$ and $A[[X_1, \ldots, X_n]] \cong A \oplus \ker \pi_A$. Φ by definition induces an equality on the first component, a surjection on the second and therefore is surjective on the direct sum. $\square$

Definition 3.3. Let $A \in \mathcal{C}_{\mathcal{O}}^{\bullet}$. Then

$$\phi_A \coloneqq (\ker \pi_A)/(\ker \pi_A)^2.$$

The reader with background in algebraic geometry might notice that this can be though of as a tangent space, in particular it is the cotangent space of the scheme $\operatorname{spec}(A)$ at the point $\ker \pi_A$. However this point of view is not necessary in the following, it might be more a hint of how Wiles came to investigate this specific invariant.

Example 3.3. Remember the definition of U in example 3.2. The tangent space $\phi_U = \ker \pi_U/(\ker \pi_U)^2$ is

$$\mathcal{O}X_1 \oplus \cdots \oplus \mathcal{O}X_n.$$

Indeed, elements of $f \in \ker \pi_U$ have no constant term as $f(0) = 0$ and therefore are multiples of X. Elements in $\ker \pi_U^2$ are multiples of X^2. As a result, we receive elements $\overline{f} \in \phi_U$ by cutting of all higher terms of a power series $f \in \ker \pi_U$.

Remark 2. *Write A as a quotient of U, $A = U/(f_1, \ldots, f_n)$. We then get $\phi_A = \phi_U/(\overline{f_1}, \ldots, \overline{f_n})$. As a quotient of ϕ_U its a finitely generated $\mathcal{O}$-module.*

Proof.

$$\phi_A = \frac{\ker \pi_A}{\ker \pi_A^2} = \frac{\ker \pi_U/(f_1, \ldots, f_n)}{(\ker \pi_U/(f_1, \ldots, f_n))^2} = \frac{\ker \pi_U/(\ker \pi_U)^2}{(\overline{f_1}, \ldots, \overline{f_n})}??$$

$\square$

Example 3.4. We now compute ϕ_A where A was defined in example 3.1. Remember that $f = T(T - \lambda^n) = -\lambda^n T + T^2$. Therefore,

$$\phi_A = \mathcal{O}T/(-\lambda^n T) = \mathcal{O}/\lambda^n.$$

Definition 3.4. Let $A \in \mathcal{C}_{\mathcal{O}}^{\bullet}$. Then

$$\eta_A \coloneqq \pi_A(\operatorname{Ann}_A(\ker \pi_A))$$

is an ideal in $\mathcal{O}$.

Example 3.5. We now compute η_U for U from example 3.2.

$$\begin{aligned}\eta_U &= \pi_U(\operatorname{Ann}\ker\pi_U)\\ &= \pi_U(\operatorname{Ann}\mathcal{O}X_1 \oplus \cdots \oplus \mathcal{O}X_n)\\ &= \pi_U(0) = 0.\end{aligned}$$

Lemma 3.1. *Let $\mathfrak{a} \subset \mathcal{O}$ be an ideal. Then*

$$\mathfrak{a} \neq 0 \implies \mathcal{O}/\mathfrak{a} \textit{ finite.}$$

Proof. As $\mathcal{O}$ is a DVR, $\mathfrak{a} = \lambda^n$ for some $n \in \mathbb{N}$ where λ is the maximal ideal in $\mathcal{O}$. Therefore, $\mathcal{O}/\mathfrak{a} = \mathcal{O}/\lambda^n$.

Using the fact that $\lambda = (t)$ for some uniformizer t, we get $\forall i \geq 1$ the isomorphism $\lambda^i/\lambda^{i+1} \cong \mathcal{O}/\lambda = k$ and thereby also the short exact sequence

$$0 \to \mathcal{O}/\lambda \cong \lambda^i/\lambda^{i+1} \to \mathcal{O}/\lambda^{i+1} \to \mathcal{O}/\lambda^i \to 0.$$

As $k = \mathcal{O}/\lambda$ is finite, we can use induction

$$\#\mathcal{O}/\lambda^{i+1} = \#\mathcal{O}/\lambda \cdot \#\mathcal{O}/\lambda^i = \#k \cdot (\#k)^i = (\#k)^{i+1}$$

and get $\#\mathcal{O}/\mathfrak{a} = \#\mathcal{O}/\lambda^n = (\#k)^n$. □

Example 3.6. We now compute η_A for A from example 3.1.

$$\begin{aligned}\eta_A &= \pi_A(\operatorname{Ann}\ker\pi_A)\\ &= \pi_A(\operatorname{Ann}\{(0,b) \subset \mathcal{O}\times\mathcal{O} | b \equiv 0 \mod \lambda^n\})\\ &= \pi_A(\{(a,0) \subset \mathcal{O}\times\mathcal{O} | a \equiv 0 \mod \lambda^n\})\\ &= \pi_A((\lambda^n)\times\mathcal{O})\\ &= (\lambda^n)\end{aligned}$$

With these results at hand, we can state

Theorem 3.1 (Wiles' numerical criterion)**.** *Let $R \twoheadrightarrow T$ a surjective morphism of augmented rings, T finite flat and $\eta_T \neq 0$ (i.e. $\mathcal{O}/\eta_T$ finite). Then the following are equivalent*

(a) $\#\phi_R \leq \#(\mathcal{O}/\eta_T)$,

(b) $\#\phi_R = \#(\mathcal{O}/\eta_T)$,

(c) R and T are complete intersections, and $R \to T$ is an isomorphism.

3.2 Basic properties of ϕ_A and η_A

In this subsection we prove the equivalence (a) ⇔ (b) in Theorem 3.1 by investigating the invariants ϕ_A and η_A that we defined last section.