Advanced Studies in Pure Mathematics 45, 2006
Moduli Spaces and Arithmetic Geometry (Kyoto, 2004)
pp. 327–386

Rigid geometry and applications

Kazuhiro Fujiwara and Fumiharu Kato

Dedicated to Professor Masaki Maruyama on the occasion of his 60th birthday

Abstract.

In this paper we present a survey of rigid geometry. Here, special emphasis is put on the so-called "birational approach" to rigid geometry, which adopts classical methods of birational geometry to the theory of rigid spaces. The paper is divided into three parts. Part I is a general introduction to rigid geometry a la J. Tate and M. Raynaud. In Part II we are to overview the birational approach to rigid geometry, which combines the idea of Raynaud and that of O. Zariski, as one of the conceptual starting points of rigid geometry. In Part III we discuss some applications, which reveal the effectiveness of the ideas in rigid geometry that arise from our viewpoint.

CONTENTS

Received November 4, 2005.
The first-named author was partially supported by the 21st century COE program of JSPS.

Part I. Classical rigid analytic geometry

§1. What is rigid geometry?

1.1. Introduction

It is well-known that the field $\mathbb{Q}$ of rational numbers admits for any prime number p a so-called *p-adic norm* $|\cdot|_p$, and they together with the usual absolute value norm $|\cdot|_\infty$ constitute the complete list of non-trivial norms on $\mathbb{Q}$ up to equivalence. The completion of $\mathbb{Q}$ by the usual absolute value $|\cdot|_\infty$ yields the field $\mathbb{R}$ of real numbers, and its algebraic closure $\mathbb{C}$, the field of complex numbers. These complete fields are at the bases of real and complex analytic geometries. As the absolute value norm is merely one of infinitely many possible norms on $\mathbb{Q}$, it is only natural to imagine a similar realm of analytic geometries arising from p-adic norms. The completion of $\mathbb{Q}$ by the p-adic norm $|\cdot|_p$ is the field $\mathbb{Q}_p$ of p-adic numbers, and the p-adic counterpart of the field $\mathbb{C}$ of complex numbers, denoted by $\mathbb{C}_p$, is the completion of the algebraic closure of $\mathbb{Q}_p$. Note that it is not simply the algebraic closure $\overline{\mathbb{Q}}_p$, since that turns out not to be complete with respect to the unique extension of the p-adic norm. Assuming the existence of analytic geometry based on the complete fields $\mathbb{Q}_p$ and $\mathbb{C}_p$ corresponding to real and complex analysis, one would, thus, finally arrive at the diagram starting from $\mathbb{Q}$ as in Figure 1. The vacant slot in the diagram is actually occupied by rigid

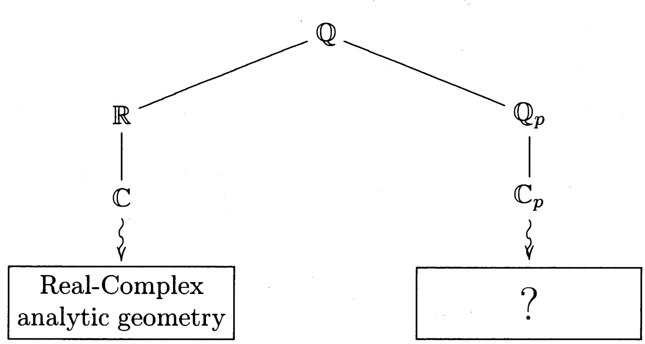

Fig. 1. Dichotomy between real-complex world and p-adic world

geometry,[1] which provides a systematic theory for analytic geometry over complete non-archimedean valued fields, not only $\mathbb{Q}_p$ and $\mathbb{C}_p$.

Table 1 shows points of similarity between the fields $\mathbb{C}$ and $\mathbb{C}_p$, which are considered to be important in the genesis of analytic geometry. As

Table 1. $\mathbb{C}$ vs $\mathbb{C}_p$

$\mathbb{C}$	$\mathbb{C}_p$
Algebraically closed	Algebraically closed
Complete with respect to absolute value $\lvert\cdot\rvert_\infty$	Complete with respect to p-adic norm $\lvert\cdot\rvert_p$

the table shows, $\mathbb{C}_p$ is algebraically closed[2] and complete. By completeness one can speak of convergent power series and functions expressed by them, which are, as in complex analysis, the fundamental things to consider also in rigid analytic geometry.

1.2. Why analytic geometry?

But, already having nice analytic geometry on the real-complex side, why do we need to consider analytic geometry also on the p-adic side? It turns out that the reason mainly comes from number-theoretic considerations. This is best explained in the context of *uniformization*, which is one of the useful techniques that reveal already in complex analytic geometry the true value of analytic methods.

Let us first briefly recall complex analytic uniformization of elliptic curves:

- regarded as a compact Riemann surface, an elliptic curve over $\mathbb{C}$ is realized as a quotient $\mathbb{C}/\Lambda$, where Λ is a lattice in $\mathbb{C}$ of the form $\Lambda = 2\pi\sqrt{-1}(\mathbb{Z} + \mathbb{Z}\cdot\tau)$ for $\tau \in \mathbb{H} = \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$;
- another way of analytic representation is provided by the quotient $\mathbb{C}^\times \to \mathbb{C}^\times/q^{\mathbb{Z}} = \mathbb{C}/\Lambda$, where $q = \exp(2\pi\sqrt{-1}\tau)$, which factorizes the previously mentioned quotient map $\mathbb{C} \to \mathbb{C}/\Lambda$ through the exponential mapping $\exp(\,\cdot\,)\colon \mathbb{C} \to \mathbb{C}^\times$.

Whereas rigid analytic geometry over $\mathbb{C}_p$ fails to have an analogue of the first uniformization, it actually affords that of the second, the so-called *Tate's uniformization*, given by a quotient of the form $\mathbb{C}_p^\times \to$

[1]The reason for the adjective "rigid" will be explained later (cf. §2.4).

[2]The non-trivial fact that $\mathbb{C}_p$ is algebraically closed is due to Krasner.

$\mathbb{C}_p^\times/q^{\mathbb{Z}}$ with $q \in \mathbb{C}_q^\times$, $|q|_p < 1$, of which Tate was able to give an analytic description [41]. In fact, we have the Weierstrass $\wp$-function on $\mathbb{C}_p^\times$ defined by the usual (but transcribed by the coordinate change $w = e^{2\pi\sqrt{-1}z}$) formula, which induces the following commutative diagram

$$\begin{array}{ccc} \mathbb{C}_p^\times & \xrightarrow{(\wp:\wp':1)} & \mathbb{P}^2(\mathbb{C}_p) \\ & \searrow \quad \nearrow & \\ & \mathbb{C}_p^\times/q^{\mathbb{Z}}, & \end{array}$$

where the dashed arrow embeds $\mathbb{C}_p^\times/q^{\mathbb{Z}}$ in $\mathbb{P}^2(\mathbb{C}_p)$ onto a cubic curve. The analytic curve $\mathbb{C}_p^\times/q^{\mathbb{Z}}$ thus obtained is called a *Tate curve.*

Remark 1.1. Contrary to the complex case, not all elliptic curves can be realized as Tate curves. It is known that an elliptic curve E over $\mathbb{C}_p$ is realized as a Tate curve if and only if $|j(E)|_p > 1$, where $j(E)$ denotes the j-invariant of E; note that the last condition is equivalent to E having multiplicative reduction.

Now we return to our first question: *why do we need analytic geometry on the p-adic side?* Consider an elliptic curve E over $\mathbb{Q}$. The complex analytic method tells us that the Riemann surface $E(\mathbb{C})$ is a complex torus, and gives us several useful analytical and topological properties. On the p-adic side, on the other hand, assuming that there exists a prime p at which E has multiplicative reduction, we know that $E(\mathbb{Q}_p)$ is written in the form $\mathbb{Q}_p^\times/q^{\mathbb{Z}}$ (by the $\mathbb{Q}_p$-rational version of Tate's uniformization). This representation of E allows one to have a good grasp on rational points on E; for example, one is able to show at a glance that the torsion part of $E(\mathbb{Q})$ is a finite group (Nagell-Lutz Theorem; this is, however not the way they proved it).

One can therefore expect in general that for an algebraic variety X over a number field, rigid analytic geometry reveals number-theoretic information hidden behind X, and thus compensates for properties that complex analytic geometry fails to capture. This is the reason why rigid analytic geometry is useful.

The *Tate curve* $\mathbb{C}_p^\times/q^{\mathbb{Z}}$ is our first example of a rigid analytic space, which will appear again and again in the sequel. Also for Tate, this curve was actually the starting point that led him to discover rigid analytic geometry. In the next section, we will overview Tate's theory of rigid analytic geometry [40], and will see at the end (in Example 2.15) how the above picture is justified.

§2. Tate's rigid analytic geometry

2.1. Non-archimedean valued fields

The above-mentioned normed fields $(\mathbb{Q}_p, |\cdot|_p)$ and $(\mathbb{C}_p, |\cdot|_p)$ are examples of so-called *complete non-archimedean valued fields with non-trivial valuation*, which are one of the basic cornerstones of Tate's rigid analytic geometry.

By a *non-archimedean valued field* we mean a pair $(K, |\cdot|)$ consisting of a field K and a non-archimedean norm $|\cdot|$, that is, a mapping $|\cdot|: K \to \mathbb{R}_{\geq 0}$ such that

(1) $|x| = 0 \Longleftrightarrow x = 0$;
(2) $|xy| = |x||y|$;
(3) $|x + y| \leq \max\{|x|, |y|\}$,

for any $x, y \in K$. The norm $|\cdot|$ is said to be *non-trivial* if $|K^\times| \neq \{1\}$. Finally, we need to assume that $(K, |\cdot|)$ is *complete*, that is, K is complete with respect to the norm $|\cdot|$.

Example 2.1. Let V be a complete discrete valuation ring, and K its field of fractions. As usual, the field K comes with a discrete valuation $v: K \to \mathbb{Z} \cup \{\infty\}$, which induces the corresponding norm $|\cdot|_v: K \to \mathbb{R}_{\geq 0}$ by the formula $|x|_v = e^{-v(x)}$ for any $x \in K$, where e is a real number with $e > 1$. Then the pair $(K, |\cdot|_v)$ is a complete non-archimedean valued field with non-trivial valuation.

The p-adic number field $(\mathbb{Q}_p, |\cdot|_p)$ together with the p-adic norm is an example of this kind. Another such example is provided by $(k((x)), |\cdot|_x)$, where $k((x))$ is the fractional field of $V = k[[x]]$, the ring of formal power series over a field k endowed with the x-adic valuation. These are examples of complete discrete valuation fields, which the reader is invited to always bear in mind.

Similarly to the construction of $\mathbb{C}_p$, the completion $\mathbb{C}_K$ of the algebraic closure of the valuation field K from Example 2.1 is algebraically closed, and the resulting pair $(\mathbb{C}_K, |\cdot|_v)$ provides another example of a complete non-archimedean valued field with non-trivial valuation.

Notice that one can perform a similar construction as in Example 2.1 starting from a valuation ring V of *height* 1 (but not necessarily discrete), that is, the fractional field K has a valuation of the form $v: K \to \mathbb{R} \cup \{\infty\}$.[3]

In the sequel of this section, K denotes a complete non-archimedean valued fields with non-trivial valuation.

[3]For generalities of valuations, we refer to [7, Chap. VI] and [46, Chap. VI].

2.2. Basic idea

Tate modeled his rigid analytic geometry on the geometry of schemes in the sense that his rigid analytic spaces are constructed by gluing certain "affine" objects. As such objects are defined, like affine schemes, as certain spectra of rings of some kind, one can say that Tate's rigid analytic geometry belongs to the general trend of understanding spaces as spectra of rings, the historical origin of which can be traced back to Gelfand. A consequence of this is the seemingly strange-looking fact that Tate's rigid analytic geometry is better understood in analogy with classical *algebraic* geometry over a field k than with complex analytic geometry.

Table 2. Comparison between algebraic geometry and rigid geometry (the italic-written items are explained in the text.)

	Algebraic geometry $/k$	Rigid geometry $/K$
Function algebra	Finitely generated algebra A/k	Topologically finitely generated algebra A/K (called: *affinoid algebra*)
Points (Naive)	Maximal ideals of A (with Zariski topology)	Maximal ideals of A (with *admissible topology*)
Building block	Affine variety (Specm $A, \mathscr{O}_X$)	*Affinoid* (Spm $A, \mathscr{O}_X$)

The rings that rigid analytic geometry deals with, which in algebraic geometry correspond to finitely generated algebras over k, are the so-called *affinoid algebras*,[4] which are by definition topologically finitely generated algebras over K (cf. Definition 2.4). Similarly to algebraic geometry, Tate's rigid analytic geometry takes the maximal ideals of A as the spectrum. As a counterpart of Zariski topology, we have the so-called *admissible topology*, which is, however, not a topology in the naive sense, but is actually a Grothendieck topology.[5] Finally, the maximal

[4]In some literature, affinoid algebras are called *Tate algebras*. Here we follow the terminology of [5], where Tate algebra means affinoid algebra of a special kind; cf. Definition 2.3.

[5]That one has to use a Grothendieck topology is a fatal drawback of Tate's theory, which makes the theory look extremely difficult. It is one of our aims

spectrum $\operatorname{Spm} A$ together with a suitably defined structure sheaf with respect to the admissible topology provides the basic building block of general analytic spaces in a similar way that varieties in algebraic geometry are constructed by gluing affine varieties. The building block thus obtained is called an *affinoid.*

Remark 2.2. Notwithstanding the perfect looking comparison with algebraic geometry, there is in fact no a priori reason in rigid geometry why one should take maximal ideals as points, and one could even say that here lies a serious problem of Tate's approach. In fact, Tate's rigid analytic spaces in general are severely deficient in points, and it is for this reason that one has to use Grothendieck topology as the natural topology to think about. This mismatching of points and topology leads to several problems: for instance, points of Tate's rigid analytic spaces are not enough to detect abelian sheaves with respect to the admissible topology.

As a matter of fact, there are many more approaches to rigid geometry, including ours (which will be explained later), and one of the most important differences between these approaches lies in what to choose as points. Namely, the notion of points in rigid analytic geometry depends entirely on the way one approaches it. Thus one can say that it is only due to Tate's way of approaching rigid geometry that one takes maximal ideals as points. This means, in other words, that another choice of points would avoid Grothendieck topology. We will see that this is in fact the case.[6]

2.3. Affinoid algebras

The most important example of affinoid algebras, which plays the role of polynomial rings in algebraic geometry, is the so-called *Tate algebra.*

Definition 2.3 (Tate algebra).

$$K\langle\langle T_1,\ldots,T_n\rangle\rangle = \left\{ \begin{array}{r|l} \sum_{\nu_1,\ldots,\nu_n\geq 0} a_{\nu_1,\ldots,\nu_n} T_1^{\nu_1}\cdots T_n^{\nu_n} & |a_{\nu_1,\ldots,\nu_n}| \to 0 \text{ as} \\ \in K[[T_1,\ldots,T_n]] & \nu_1+\cdots+\nu_n\to\infty \end{array} \right\}.$$

of this paper to show that it is by no means essential to use Grothendieck topologies in developing rigid geometry. See Remark 2.2.

[6]Here we would like to stress that, nevertheless, it is not our intension to defy Tate's approach; each approach has its own advantage and drawback. Rather, we believe that a good attitude is to have various approaches at one's disposal and to feel free in choosing one of them depending on the situation.

The similarity with the polynomial ring comes from the fact that the Tate algebra $K\langle\langle T_1, \ldots, T_n\rangle\rangle$ is the K-algebra consisting of power series converging absolutely and uniformly on the closed unit polydisk $\{(z_1, \ldots, z_n) \in K^n \mid |z_i| \leq 1 \text{ for } 1 \leq i \leq n\}$ in K^n.[7] Assume for simplicity that K is algebraically closed. Then the set of all maximal ideals of $K\langle\langle T_1, \ldots, T_n\rangle\rangle$ coincides with the closed unit polydisk (this follows from the weak Nullstellensatz for affinoid algebras stated below). The corresponding affinoid is, therefore, underlain by this set. Table 3 shows the dictionary for comparison between the polynomial ring and the Tate algebra.

Table 3. Polynomial ring vs Tate algebra

Algebraic geometry $/k = \overline{k}$	Rigid geometry $/K = \overline{K}$
$k[X_1, \ldots, X_n]$	$K\langle\langle X_1, \ldots, X_n\rangle\rangle$
k^n	$(z_1, \ldots, z_n) \in K^n$ with $\lvert z_i\rvert \leq 1$
$\mathbb{A}^n_k$ affine space	$\mathbb{D}^n_K$ closed unit polydisk

Basic properties. Here we list some basic properties of the Tate algebra; one finds more in [5, Chap. 5]:

- it is a K-Banach algebra endowed with the so-called *Gauss norm:*

$$\Big\| \sum_{\nu_1,\ldots,\nu_n \geq 0} a_{\nu_1,\ldots,\nu_n} T_1^{\nu_1} \cdots T_n^{\nu_n} \Big\| = \sup_{\nu_1,\ldots,\nu_n \geq 0} |a_{\nu_1,\ldots,\nu_n}|;$$

- it is Noetherian, and every ideal is closed with respect to the topology induced by the Gauss norm.

Definition 2.4 (Affinoid algebra). An *affinoid algebra* is a K-algebra of the form

$$A = K\langle\langle T_1, \ldots, T_n\rangle\rangle / I$$

for some n, where I is an ideal. This is a K-Banach algebra by the norm induced from the Gauss norm.

[7]Note that this set is an *open* subset of K^n with respect to the metric topology.

Among several basic properties of affinoid algebras, we mention the analogue of Noether's normalization theorem ([5, 6.1.2]):

- (Noether's normalization theorem for affinoid algebras) for any affinoid algebra A over K there exists a finite injective K-algebra homomorphism

$$K\langle\langle T_1, \ldots, T_d\rangle\rangle \longrightarrow A$$

for some $d \geq 0$.

By this we have the following property, which implies the functoriality of taking the maximal spectrum:

- (Weak Nullstellensatz for affinoid algebras) for any maximal ideal $\mathfrak{m}$ of A, the residue field $A/\mathfrak{m}$ is a finite extension of K.

2.4. Wobbly topology

For an affinoid algebra A we set $\operatorname{Spm} A$ to be the set of all maximal ideals of A. For any K-algebra homomorphism $A \to B$ between affinoid algebras[8] we have an induced mapping $\operatorname{Spm} B \to \operatorname{Spm} A$. As usual, any element f of A is regarded as a function on the set $\operatorname{Spm} A$; since for any $x \in \operatorname{Spm} A$ the residue field at x is a finite extension of K and thus admits a unique extension of the norm $|\cdot|$, one can put $|f(x)| = |f \bmod x|$. For any $f, g \in A$ we set

$$R(f, g) = \{x \in \operatorname{Spm} A \mid |f(x)| \leq |g(x)|\}.$$

As a subset of $\operatorname{Spm} A$, we have

$$R(f, g) = \operatorname{Spm} A\langle\langle X\rangle\rangle/(gX - f),$$

where $A\langle\langle X\rangle\rangle$ denotes the ring $A\widehat{\otimes}_K K\langle\langle X\rangle\rangle$. The ring $A\langle\langle X\rangle\rangle/(gX - f)$, which is again an affinoid algebra, is often abbreviated as $A\langle\langle \frac{f}{g}\rangle\rangle$.

Definition 2.5 (Wobbly topology)**.** The *wobbly topology* on the set $\operatorname{Spm} A$ is the topology having $\{R(f, g)\}_{f,g\in A}$ as open basis.

Example 2.6. Suppose for simplicity that K is algebraically closed, and consider $\mathbb{D}^n_K = \operatorname{Spm} K\langle\langle X_1, \ldots, X_n\rangle\rangle$, which is identified as a set with the closed unit polydisk in K^n. Then one sees easily that the wobbly topology on $\mathbb{D}^n_K$ coincides with the topology induced from the metric topology on K^n (which is, as is well-known, totally disconnected).

[8] Any K-algebra homomorphism between affinoid algebras is automatically continuous.

Difficulties. As Example 2.6 indicates, the wobbly topology is not such a good topology; for example:

- $\mathrm{Spm}\,A$ with the wobbly topology is, in most cases, not quasi-compact, which would be troublesome when one considers gluing;
- the presheaf $R(f,g) \mapsto A\langle\langle \frac{f}{g} \rangle\rangle$, which comes as the most natural candidate for the structure sheaf on $\mathrm{Spm}\,A$, is in general not a sheaf.

These difficulties come from the fact that the wobbly topology is somewhat too fine. Indeed, considering the sheafification of the above presheaf, we would get a ring of functions on $\mathrm{Spm}\,A$ that is much larger than A itself, which contradicts our basic requirement that A should be the ring of all "holomorphic" functions on $\mathrm{Spm}\,A$. In other words, the wobbly topology leads to a very feeble notion of analytic functions. Hence, to obtain a reasonable theory of analysis, one has to "rigidify" the notion of analytic functions,[9] and, to this end, one wants to replace the wobbly topology with a more legitimate one.

2.5. Admissible topology

In 1961 Tate [40] overcame the above-mentioned difficulties by introducing the so-called *admissible topology.* The admissible topology is, in short, a Grothendieck topology that is

- weaker than the wobbly topology,
- the strongest one that makes each $R(f,g)$ quasi-compact.

The actual definition is given as follows.

Definition 2.7 (Admissible site). Let $\mathfrak{A}_K$ be the category of affinoid algebras over K and K-algebra homomorphisms. For any object A of $\mathfrak{A}_K$, we denote by $\underline{\mathrm{Spm}}\,A$ the same object considered as an object of the opposite category $\mathfrak{A}_K^{\mathrm{opp}}$. We define a Grothendieck topology on $\mathfrak{A}_K^{\mathrm{opp}}$ as follows: a finite collection $\{\underline{\mathrm{Spm}}\,A_i \to \underline{\mathrm{Spm}}\,A\}_{i\in I}$ of morphisms in $\mathfrak{A}_K^{\mathrm{opp}}$ is a *covering* of $\underline{\mathrm{Spm}}\,A$ if and only if

(1) each A_i is étale over A (see, for example, [17, §8.1] for the definition of étaleness);
(2) $\mathrm{Spm}\,A_i \to \mathrm{Spm}\,A$ is injective for each i and induces an isomorphism between the residue fields at each point of $\mathrm{Spm}\,A_i$;
(3) $\mathrm{Spm}\,A = \bigcup_{i\in I} \mathrm{Spm}\,A_i$.

We denote the resulting site by $\mathfrak{A}_{K,\mathrm{ad}}^{\mathrm{opp}}$.

---

[9] This is the reason for the name "rigid" geometry.

Here is a typical example of coverings in the admissible site. Let A be an affinoid algebra over K, and $f_0, \ldots, f_n \in A$ elements of A such that $(f_0, \ldots, f_n) = A$. Set

$$A_i = A\langle\langle X_0, \ldots, \widehat{X}_i, \ldots, X_n\rangle\rangle/(f_iX_j - f_j \mid j \neq i).$$

Then the collection $\{\underline{\mathrm{Spm}}\, A_i \to \underline{\mathrm{Spm}}\, A\}_{0\leq i\leq n}$ is a covering in the site $\mathfrak{A}^{\mathrm{opp}}_{K,\mathrm{ad}}$; for each i, the image of $\mathrm{Spm}\, A_i \to \mathrm{Spm}\, A$ is given as $\{x \in \mathrm{Spm}\, A \mid |f_i(x)| \geq |f_j(x)| \text{ for } j \neq i\}$. Let us denote this covering by $\mathscr{R}(f_0, \ldots, f_n)$.

Theorem 2.8 (Gerritzen-Grauert). *Any covering family* $\{\underline{\mathrm{Spm}}\, A_i \to \underline{\mathrm{Spm}}\, A\}_{i\in I}$ *in the site* $\mathfrak{A}^{\mathrm{opp}}_{K,\mathrm{ad}}$ *has a refinement to a covering of the form* $\mathscr{R}(f_0, \ldots, f_n)$ *for some* $f_0, \ldots, f_n \in A$.

Another version of the Gerritzen-Grauert theorem will be stated in Corollary 6.21 below.

Affinoids and general rigid spaces. For an affinoid algebra A over K, consider the presheaf $\mathscr{O}_{\mathrm{Spm}\, A}$ on the comma site $(\mathfrak{A}^{\mathrm{opp}}_{K,\mathrm{ad}})_{\underline{\mathrm{Spm}}\, A}$ defined by

$$\mathscr{O}_{\mathrm{Spm}\, A}\colon (\underline{\mathrm{Spm}}\, B \to \underline{\mathrm{Spm}}\, A) \mapsto B.$$

The following theorem says that the admissible topology defined above is the good one in the sense that it gives rise to the correct notion of "holomorphic" functions.

Theorem 2.9 (Tate's acyclicity theorem). *The presheaf* $\mathscr{O}_{\mathrm{Spm}\, A}$ *is a sheaf on* $(\mathfrak{A}^{\mathrm{opp}}_{K,\mathrm{ad}})_{\underline{\mathrm{Spm}}\, A}$ *with respect to the admissible topology.*

Definition 2.10 (Tate's rigid analytic space). (1) A representable sheaf on the site $\mathfrak{A}^{\mathrm{opp}}_{K,\mathrm{ad}}$ is called an *affinoid.*

(2) A map $\mathscr{Y} \hookrightarrow \mathscr{X}$ between affinoids is said to be an *open immersion* if, identified with a morphism in the category $\mathfrak{A}^{\mathrm{opp}}_{K}$, it satisfies the conditions (1) and (2) in Definition 2.7.

(3) A sheaf $\mathscr{X}$ of sets on the site $\mathfrak{A}^{\mathrm{opp}}_{K,\mathrm{ad}}$ is called a (*Tate's*) *rigid analytic space* if there exists a surjective map of sheaves

$$\coprod_{i\in I} \mathscr{Y}_i \longrightarrow \mathscr{X},$$

where $\mathscr{Y}_i$ for each $i \in I$ is an affinoid, such that, for each $i, j \in I$, the projection $\mathscr{Y}_i \times_{\mathscr{X}} \mathscr{Y}_j \to \mathscr{Y}_i$ is isomorphic to the limit of a filtered direct system $\{\mathscr{U}_\lambda \to \mathscr{Y}_i\}_{\lambda\in\Lambda}$ of maps between affinoids such that all maps in the commutative diagram for $\mu \leq \lambda$

$$\begin{array}{ccc} \mathscr{U}_\lambda & \longrightarrow & \mathscr{Y}_i \\ \uparrow & \nearrow & \\ \mathscr{U}_\mu & & \end{array}$$

are open immersions.

In other words, Tate's rigid analytic spaces are constructed by gluing affinoids. As the definition indicates, it allows non-separated or non-quasi-separated gluing.

Remark 2.11. In Tate's original approach, rigid analytic spaces are regarded as local ringed spaces with Grothendieck topology.[10] For example, an affinoid is such a space isomorphic to the one given by the data $(\mathrm{Spm}\,A, \mathscr{T}_A, \mathscr{O}_{\mathrm{Spm}\,A})$ consisting of the set $\mathrm{Spm}\,A$, the Grothendieck topology $\mathscr{T}_A$ (equivalent to the admissible topology in our sense), and the sheaf of rings (essentially the same as the one that we have given above). General rigid analytic spaces are obtained by gluing these spaces with respect to what is called the *strong topology.* This viewpoint of rigid analytic geometry is surely useful. But one has to be careful, since, as we have already seen in Remark 2.2, the point set $\mathrm{Spm}\,A$ is not the correct "underlying set" for the affinoid $\underline{\mathrm{Spm}}\,A$.

In the sequel, for brevity and conformity with the usual notation, we denote the affinoid $\underline{\mathrm{Spm}}\,A$ *simply by* $\mathrm{Spm}\,A$.

2.6. Examples

Example 2.12 (Annulus). An annulus is an affinoid that is, if K is algebraically closed, supported on the set

$$\{z \in K \mid |a| \leq |z| \leq |b|\}$$

with $a, b \in K$. The corresponding affinoid algebra is given by

$$K\langle\langle \tfrac{a}{z}, \tfrac{z}{b} \rangle\rangle = K\langle\langle X, Y \rangle\rangle / (XY - \tfrac{a}{b}).$$

Note that, since it is an affinoid, it is quasi-compact.

In general, a rigid analytic space is said to be *quasi-compact* if it has an admissible covering consisting of finitely many affinoids.

[10] See [5, 9.1] for what "Grothendieck topology" means here.

Example 2.13 (Affine line). An affine line $\mathbb{A}_K^{1,\mathrm{an}}$ in rigid analytic geometry is realized as, for example, the limit of concentric closed disks, each of which is an affinoid:

$$\mathbb{A}_K^{1,\mathrm{an}} = \varinjlim_{n\geq 1} \operatorname{Spm} K\langle\langle a^n z\rangle\rangle,$$

where a is an element of K with $|a| < 1$. This of course reflects the equality

$$K = \bigcup_{n\geq 1} \mathbb{D}(0, |a|^{-n}),$$

where $\mathbb{D}(0,r) = \{z \in K \mid |z| \leq r\}$. The affine line $\mathbb{A}_K^{1,\mathrm{an}}$ is not quasi-compact.

Example 2.14 (Multiplicative group). Let $a \in K$ be as above. The multiplicative group $K^\times$ is regarded as the union of countably many annuli:

$$\begin{aligned} K^\times &= \bigcup_{n\geq 1} \{z \in K \mid |a|^n \leq |z| \leq |a|^{-n}\} \\ &= \bigcup_{n\in\mathbb{Z}} \{z \in K \mid |a|^{n+1} \leq |z| \leq |a|^n\}. \end{aligned}$$

Taking up, for example, the latter description, one defines

$$\mathbb{G}_{m,K}^{\mathrm{an}} = \textstyle\bigcup_{n\in\mathbb{Z}} \operatorname{Spm} K\langle\langle \frac{a^{n+1}}{z}, \frac{z}{a^n}\rangle\rangle.$$

This is again a rigid analytic space that is not quasi-compact.

Example 2.15 (Tate curve). The last description of the rigid analytic multiplicative group $\mathbb{G}_{m,K}^{\mathrm{an}}$ allows one to display the analytic structure of the Tate curve discussed in §1.2. For $q \in K$ with $|q| < 1$, the Tate curve is given by $\mathbb{G}_{m,K}^{\mathrm{an}}/q^{\mathbb{Z}}$. In order to describe an analytic covering, take $a \in K$ such that $|a|^k = |q|$ for some $k \geq 2$ and the analytic covering $\mathbb{G}_{m,K}^{\mathrm{an}} = \bigcup_{n\in\mathbb{Z}} A_n$ considered in Example 2.14, where $A_n = \operatorname{Spm} K\langle\langle \frac{a^{n+1}}{z}, \frac{z}{a^n}\rangle\rangle$. Multiplication by q maps each A_n isomorphically onto A_{n+k}. Thus, $\mathbb{G}_{m,K}^{\mathrm{an}}/q^{\mathbb{Z}}$ is written as the union of k annuli, glued together by identifying the "exterior" boundary component with the "interior" boundary component of another one. In particular, it is a quasi-compact rigid analytic space.

As mentioned at the end of §1, one of Tate's goals in formulating rigid analytic geometry was to give a legitimate way of regarding $\mathbb{C}_p^\times/q^{\mathbb{Z}}$ as an "analytification" of an elliptic curve. This was done in the last example.

§3. Raynaud's approach to rigid geometry

3.1. Formal models of affinoids

The moral basis of the p-adic counterpart of real-complex analytic geometry, leading to the saga of Tate's theory of rigid analytic geometry, was, as we have seen in §1.1, the similarity between the complex number field $\mathbb{C}$ and its p-adic counterpart $\mathbb{C}_p$, as listed in Table 1. Now we change our view to the completely opposite direction, and rather pay attention to differences between $\mathbb{C}$ and $\mathbb{C}_p$. The most important difference is that $\mathbb{C}_p$ has, while $\mathbb{C}$ does not, the sub*ring* consisting of integral elements, that is, elements of norm ≤ 1 (Table 4). Similarly, any affinoid algebras,

Table 4. $\mathbb{C}$ vs $\mathbb{C}_p$ (continued)

$\mathbb{C}$	$\mathbb{C}_p$
$\nexists$ integer ring	$\exists$ integer ring

unlike function algebras in real-complex analysis, have a "model" over the integer ring. This observation, however simple it might look, is the starting point of Raynaud's approach to rigid analytic geometry, which, as we will see, leads to a bold shift of viewpoint.

Situation. In the sequel of this section we work in the following situation:

- V is a valuation ring of height 1 that is complete with respect to the a-adic topology for an element a belonging to the maximal ideal $\mathfrak{m}_V$;
- we set $K = \mathrm{Frac}(V)$ (the field of fractions), which has the a-adic norm $|\cdot|$ and is complete with respect to the metric topology induced from this norm.

Note that a valuation ring V of arbitrary height is a-adically separated if and only if $V[\frac{1}{a}]$ is a field (and hence coincides with $\mathrm{Frac}(V)$).

Example 3.1. The typical example is provided by a complete discrete valuation ring V with π-adic topology, where π is a generator of the maximal ideal (uniformizer). The corresponding norm $|\cdot|$ on the fractional field K coincides with the one as in Example 2.1 up to equivalence.

In this situation, for any topologically finitely generated V-algebra A, we obtain an affinoid algebra

$$A_K = A \otimes_V K$$

over K. Here, a V-algebra A is said to be *topologically finitely generated* if it is a quotient by an ideal of an algebra of the form $V\langle\langle X_1, \dots, X_n\rangle\rangle$, the a-adic completion of the polynomial ring $V[X_1, \dots, X_n]$.

In general, let $\mathscr{A}$ be an affinoid algebra over K. A *formal model* of $\mathscr{A}$ is a topologically finitely generated V-algebra A such that $A_K = A \otimes_V K$ is isomorphic to $\mathscr{A}$ as a K-algebra. If, in addition, A is flat over V, or what amounts to the same, A is a-torsion free, then we say that A is a *distinguished* (or *flat*) formal model of $\mathscr{A}$. For example, the algebra $V\langle\langle X_1, \dots, X_n\rangle\rangle$ is a distinguished formal model of the Tate algebra $K\langle\langle X_1, \dots, X_n\rangle\rangle$. Any affinoid algebra has a distinguished formal model; indeed, if it is given as $K\langle\langle X_1, \dots, X_n\rangle\rangle/I$ with $I = (f_1, \dots, f_r)$ finitely generated (recall that the Tate algebra is Noetherian), by multiplying each f_i with a power of a, one can assume $f_i \in V\langle\langle X_1, \dots, X_n\rangle\rangle$ and then $A/A_{a\text{-tor}}$, where $A = V\langle\langle X_1, \dots, X_n\rangle\rangle/(f_1, \dots, f_r)$, gives a desired formal model.

Remark 3.2. It is known that, if a topologically finitely generated V-algebra A is flat, then it is actually topologically finitely *presented* ([6]). As it is also known that any finitely generated ideal of $V\langle\langle X_1, \dots, X_n\rangle\rangle$ is closed with respect to the a-adic topology (i.e. the Artin-Rees lemma is valid for finitely generated ideals; cf. [18]), any topologically finitely generated flat V-algebra is complete with respect to the a-adic topology.

3.2. Raynaud's functor

Let $X = \operatorname{Spf} A$ be an affine flat formal scheme of finite type over $\operatorname{Spf} V$, where V is considered with some a-adic topology. Then, as we have seen, $A_K = A \otimes_V K$ is an affinoid algebra, and thus we can consider the corresponding affinoid $X_K = \operatorname{Spm} A_K$ over K. This correspondence $X \mapsto X_K$ is globalized in the following way.

Consider the localization $A_{\{f\}}$ by an element $f \in A$, that is, the a-adic completion of A_f. Note that we have[11]

$$A_{\{f\}} = A\langle\langle X\rangle\rangle/(fX - 1).$$

Therefore, the corresponding affinoid $\operatorname{Spm}(A_{\{f\}})_K$ is nothing but $R(1, f)$ (cf. §2.4), an admissible open subset of $\operatorname{Spm} A_K$ with respect to the admissible topology. Hence, by patching, one obtains a functor

$$X \mapsto X_K$$

[11]Indeed, while the a-adic completion of $A\langle\langle X\rangle\rangle/(fX - 1)$ obviously coincides with $A_{\{f\}}$, as we have mentioned in Remark 3.2, $A\langle\langle X\rangle\rangle/(fX - 1)$ is already complete, whence the equality.

from the category of coherent (= quasi-compact and quasi-separated) flat formal schemes of finite type over V to the category of Tate's rigid spaces over K ([37]). This functor is called the *Raynaud functor*, and the rigid space X_K associated to X is the *Raynaud generic fiber* of X. Put in the other way, when a rigid space $\mathscr{X}$ is isomorphic to X_K for a flat formal scheme X as above, we say that X is a (*distinguished*) *formal model* of $\mathscr{X}$.

Note that, by the definition of the functor, if X has an affine covering $X = \bigcup_{i \in I} U_i$, then we get an admissible covering $X_K = \bigcup_{i \in I} U_{i,K}$. Let us call this covering of X_K the admissible covering induced from the affine covering $\{U_i\}_{i \in I}$ of X.

Example 3.3. Let V be as in Example 3.1, and consider a semi-stable curve $E \to \operatorname{Spec} V$ such that the generic fiber E_η is an elliptic curve over K, and that the closed fiber E_0 is the union of non-singular rational curves arranged as type I_k in Kodaira's classification. Consider the formal completion $\widehat{E}$ along the closed fiber E_0. It admits the affine covering $\widehat{E} = \bigcup_{n=1}^{k} U_n$, where each $U_n = \operatorname{Spf} V\langle\langle \frac{\pi^{n+1}}{z}, \frac{z}{\pi^n} \rangle\rangle$ is isomorphic to $\operatorname{Spf} V\langle\langle X, Y \rangle\rangle/(XY - \pi)$. The corresponding rigid space $\mathscr{E} = \widehat{E}_K$ is the Tate curve, which has the induced admissible covering $U_{n,K} = \operatorname{Spm} K\langle\langle \frac{\pi^{n+1}}{z}, \frac{z}{\pi^n} \rangle\rangle$. Note that this rigid analytic space $\mathscr{E}$, as well as the admissible covering, is nothing but the one we have already described in Example 2.15.

3.3. Zariski topology vs admissible topology

Needless to say, there may be many choices of formal models for a given rigid space, and this diversity of choice is, in fact, reflected in diversity of choice of admissible coverings of the rigid space. To see this, let us first establish a typical change of formal models.

Admissible blow-up. Let X be a formal scheme of finite type over V. An *admissible ideal* is a quasi-coherent open ideal $\mathscr{J}$ of $\mathscr{O}_X$ of finite type. For an admissible ideal $\mathscr{J}$, the *admissible blow-up* along $\mathscr{J}$ is the morphism of formal schemes

$$X' = \varinjlim_{k \geq 0} \operatorname{Proj}\left(\bigoplus_{n \geq 0} \mathscr{J}^n \otimes \mathscr{O}_{X_k}\right) \longrightarrow X,$$

where $X_k = (X, \mathscr{O}_X/a^{k+1}\mathscr{O}_X)$. If, for instance, $X = \operatorname{Spf} A$ is affine, then $\mathscr{J}$ is of the form[12] J^Δ for a uniquely determined finitely generated

[12]Here we followed the commonly used notation as in [EGA, $\mathrm{I}_{\mathrm{new}}$, (10.10)].

ideal J of A that contains a power of a, and the admissible blow-up $X' \to X$ is nothing but the a-adic completion of the usual blow-up $Y' = \mathrm{Proj} \bigoplus_{n \geq 0} J^n \to Y = \mathrm{Spec}\, A$.

Example 3.4. Consider $X = \mathrm{Spf}\, V\langle\langle z \rangle\rangle$. The corresponding rigid space X_K is the "closed unit disk" $\mathbb{D}^1_K = \mathrm{Spm}\, K\langle\langle z \rangle\rangle$. Consider the admissible ideal $J = (X, a)$ of $V\langle\langle z \rangle\rangle$. The admissible blow-up X' along J is the union of two affine subsets $U = \mathrm{Spf}\, V\langle\langle \frac{z}{a} \rangle\rangle$ and $W = \mathrm{Spf}\, V\langle\langle z, \frac{a}{z} \rangle\rangle = \mathrm{Spf}\, V\langle\langle z, w \rangle\rangle/(zw - a)$. The resulting rigid space X'_K is therefore covered by two admissible open subsets $U_K = \mathrm{Spm}\, K\langle\langle \frac{z}{a} \rangle\rangle$ and $W_K = \mathrm{Spm}\, K\langle\langle z, \frac{a}{z} \rangle\rangle$; U_K is again a closed disk but having a different radius equal to $|a|$, and W_K is a closed annulus "$\{z \in K \mid |a| \leq |z| \leq 1\}$". Thus the rigid space X'_K is isomorphic to X_K. The difference is that, while X_K was considered as rigid space by the trivial covering (the covering by itself), X'_K has the non-trivial induced covering $\{U_K, W_K\}$.

As indicated in Example 3.4, whereas an admissible blow-up does not change the Raynaud generic fiber, viz. for a coherent (= quasi-compact and quasi-separated) flat formal V-scheme of finite type X and an admissible blow-up $X' \to X$ we have $X'_K = X_K$, it replaces the admissible covering by a refinement. Raynaud's very important insight is that this fact is the key point for comparing admissible topology and Zariski topology.

Consider, for example, the affine case $X = \mathrm{Spf}\, A$, and let U be a quasi-compact open subset of an admissible blow-up X' of X:

$$\begin{array}{ccc} U & \hookrightarrow & X' \\ & & \downarrow \\ & & X. \end{array}$$

Then we have the open immersion

$$U_K \hookrightarrow \mathrm{Spm}\, A_K$$

identifying U_K with a quasi-compact open subset of $\mathrm{Spm}\, A_K$ with respect to the admissible topology. Due to the Gerritzen-Grauert theorem (Theorem 2.8), the open subsets of the form U_K constructed as above constitute an open basis for the admissible topology. Thus one can recover the admissible topology on X_K from the Zariski topology of formal models. The important fact is that, in order to recover the admissible topology, one has to vary the formal model.

3.4. Raynaud's viewpoint

The important point of the above observation is that it explains the admissible topology entirely in terms of formal models. Based on this, we can now illustrate Raynaud's viewpoint of rigid analytic geometry; this is itemized as follows:

- from this viewpoint, rigid analytic geometry in totality is induced from a geometry of "models" (Figure 2);
- as the geometry of models, Raynaud suggests geometry of formal schemes over valuation rings.

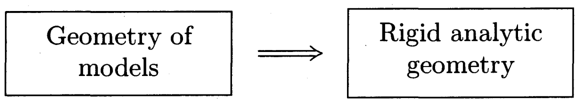

Fig. 2. Raynaud's viewpoint

For instance, if K is the fractional field of a complete discrete valuation ring V, then theorems in rigid analytic geometry over K should follow from theorems in formal geometry over V, already stated in [EGA, **III**, §4, §5].

In practice, this program goes along the following thread. Starting from a coherent formal scheme X of finite type over V, we obtain the rigid analytic space $\mathscr{X} = X_K$ over K, whose topology, points, and structure sheaf are characterized as follows.

- Topology: a quasi-compact admissible open subset of $\mathscr{X}$ is of the form $\mathscr{U} = U_K$ where U is a quasi-compact open subset of an admissible blow-up X' of X;
- Points:

$$\begin{aligned} \mathscr{X}(K) &= \{\text{sections } \operatorname{Spf} V \to X\}, \\ \mathscr{U}(K) &= \{\text{sections that factors through } U\}; \end{aligned}$$

- Structure sheaf: when $U = \operatorname{Spf} A$, then $\Gamma(\mathscr{U}, \mathscr{O}_{\mathscr{X}}) = A_K$.

This viewpoint culminates in the following theorem.

Theorem 3.5 (Raynaud 1972 [37]). *The Raynaud functor $X \mapsto X_K$ gives rise to the categorical equivalence*

$$\left\{\begin{matrix}\text{Coherent formal}\\ \text{schemes of finite}\\ \text{type over } V\end{matrix}\right\} \Big/ \left\{\begin{matrix}\text{Admissible}\\ \text{blow-ups}\end{matrix}\right\} \xrightarrow{\sim} \left\{\begin{matrix}\text{Coherent rigid}\\ \text{analytic spaces of}\\ \text{finite type over } K\end{matrix}\right\}.$$

Here the left-hand category is the quotient category, that is, the category consisting of the same objects as the category of coherent formal schemes of finite type over V but of arrows with all admissible blow-ups inverted.

Remark 3.6. (1) The objects in the right-hand category are defined a priori by "patching affinoids" (cf. Definition 2.10). The equivalence shows that this patching turns out to be equivalent to, so to speak, "birational patching" (birational up to admissible blow-ups). This invokes a birational viewpoint in rigid geometry, which will play a very important role in our approach (to be explained) to rigid geometry.

(2) Let us briefly mention something about the proof of Theorem 3.5. There are two important ingredients:

- existence of formal birational patching,
- comparison of topologies.

The last point was already mentioned in connection with the Gerritzen-Grauert Theorem.

3.5. Significance of Raynaud's viewpoint

Perhaps the most significant aspect of Raynaud's viewpoint (and Raynaud's theorem) lies in the *shift* from "analysis" to "geometry". To be more precise, whereas Tate's rigid analytic geometry is motivated by "analysis" over non-archimedean fields, Raynaud's approach starts totally differently, namely from formal "geometry," and is developed entirely as a geometric theory with seemingly no flavor of analysis. Consequently, contrary to Tate's rigid analytic geometry, which aims at something similar to complex analytic geometry, Raynaud's approach forces one to think that rigid geometry is entirely *not* similar to complex analytic geometry.

§4. Our approach: brief announcement

In the next part, we are to exhibit our approach to rigid geometry, which is different both from Tate's and Raynaud's approaches. Our General Policy is the following.

> **General Policy:** rigid geometry is a hybrid of *formal geometry* and *birational geometry.*

There is little doubt that our approach has been largely influenced by Raynaud's approach. But, nevertheless, it differs much from Raynaud's in how to deal with birational geometry, on which our approach puts much more stress. For the general treatment of birational geometry, we will take up Zariski's classical idea that deals with the so-called *Zariski-Riemann spaces* as its foremost objects. Schematically shown, our approach is an "amalgam" of Raynaud's approach and Zariski's classical approach to birational geometry (Figure 3).

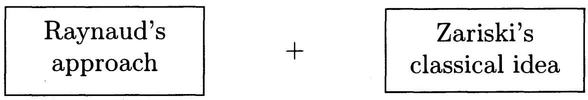

Fig. 3. Our approach

For this reason, we will start the next part of this paper with a brief recap of birational geometry from Zariski's classical viewpoint.

Part II. Birational approach to rigid geometry

Part II consists of two sections. In §5 we describe some birational geometry in the spirit of Zariski's classical viewpoint. What we do in this section is a preparation for the next section, §6, where we will outline our approach to rigid geometry.

§5. Birational geometry from Zariski's viewpoint

5.1. Basic Question: Extension problem

Throughout this section we work in the following situation:

- S: a coherent scheme,
- $\mathscr{I} = \mathscr{I}_D$: a quasi-coherent ideal sheaf of finite type such that $U = S \setminus D$ is a dense open subset of S, where $D = V(\mathscr{I})$.

Here a scheme is said to be *coherent* if it is quasi-compact and quasi-separated.[13] Note that, as the ideal sheaf $\mathscr{I}_D$ is of finite type, the open subset U is quasi-compact.

The basic problem we are concerned with is of the following type.

[13]Coherent schemes are the analogue of compact Hausdorff topological spaces in the category of schemes.

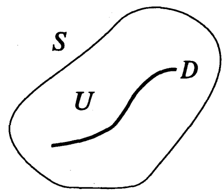

Fig. 4. Situation for the extension problem

Problem 5.1 (Extension problem). *Let P be a property of morphisms (e.g. $P =$ "flat"). Let $f_U \colon X_U \to U$ be a morphism of schemes of finite presentation with the property P. Suppose there exists at least one morphism $f \colon X \to S$ such that $f \times_S U = f_U$. Then, can one find such an f that satisfies the property P?*

This problem may have a trivial solution; for instance, if $P =$"flat", then $f = j \circ f_U$, where $j \colon U \to S$ is the open immersion, gives a solution. Such a solution is, needless to say, not the one we want to have. We like to find a "good" solution. However, if we like to clarify what "good" means, we find that the problem itself is not well-posed (or, say, not reasonable). For instance, if, trying to make the problem well-posed, we put $P =$"proper and flat", then a moment thought immediately gives a negative answer in practically important cases (e.g. family of curves over a surface S with D a normal crossing divisor), and hence we find that the problem in this case is not reasonable.

5.2. Admissible modifications and modified extension problem

In order to make the extension problem more reasonable, one needs to allow birational changes of S that preserve the dense open part U. Thus we are naturally led to the following notion.

Definition 5.2 (U-admissible modification). (1) A *U-admissible modification* of S is a diagram

$$\begin{array}{ccc} U & \hookrightarrow & S' \\ & \searrow & \downarrow \\ & & S \end{array}$$

such that the vertical arrow is proper and the other arrows are open immersions onto dense open subsets (hence the vertical arrow is birational).

(2) A *morphism* between two U-admissible modifications $S' \to S$ and $S'' \to S$ is an S-morphism $S' \to S''$.

U-admissible modifications constitute the category $\mathbf{MD}_{(S,U)}$, which is cofiltered; indeed, for two U-admissible modifications $S' \to S$ and $S'' \to S$ one constructs the diagram in $\mathbf{MD}_{(S,U)}$

$$\begin{array}{ccc} & S''' & \\ \swarrow & & \searrow \\ S' & & S'', \end{array}$$

where S''' is the closure of the image of the diagonal mapping $U \hookrightarrow S' \times_S S''$.[14]

The following special class of U-admissible modifications will be of particular importance.

Definition 5.3 (U-admissible blow-up). A *U-admissible blow-up* of S is a blow-up $S' \to S$ whose center is given by a quasi-coherent ideal $\mathscr{J}$ of $\mathscr{O}_S$ of finite type such that the corresponding closed subscheme $V(\mathscr{J})$ is set-theoretically contained in D, or what amounts to the same, there exists a positive integer n such that $\mathscr{I}_D^n \subseteq \mathscr{J}$.

Here is an example: when $S = \operatorname{Spec} A$ is affine, and $D = V(I)$, then a U-admissible blow-up is given by

$$S' = \operatorname{Proj} \bigoplus_{n \geq 0} J^n \to S,$$

where J is a finitely generated ideal of A that contains I^k for some $k > 0$.

We denote by $\mathbf{BL}_{(S,U)}$ the full subcategory of $\mathbf{MD}_{(S,U)}$ consisting of U-admissible blow-ups. To state the modified extension problem, we need yet one more concept.

Definition 5.4 (Strict transform). Let $S' \to S$ be a U-admissible modification, and $f\colon X \to S$ an S-scheme. The *strict transform* $f'\colon X' \to S'$ of f is the S'-scheme defined by the commutative diagram

$$\begin{array}{ccccc} X & \longleftarrow & X_{S'} = X \times_S S' & \hookleftarrow & X' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f_{S'}} & \swarrow{\scriptstyle f'} & \\ S & \longleftarrow & S', & & \end{array}$$

where the map $X' \hookrightarrow X_{S'}$ is the closed immersion given by dividing out $\mathscr{I}_D$-torsion.

[14]The scheme S''' might be called the *join* of S' and S''.

Having these notions on birational changes of schemes, we can now state the desired "modified" version of our basic question that we are going to consider.

Problem 5.5 (Modified extension problem). *Let $f_U \colon X_U \to U$ be a morphism of finite presentation that satisfies the property P. Suppose an extension $f \colon X \to S$ of f_U on S, that is, a morphism such that $f \times_S U = f_U$, is given. Then, can one find a U-admissible modification (resp. blow-up) $S' \to S$ such that the strict transform $f' \colon X' \to S'$ of f satisfies P?*

5.3. Flattening theorem

Problem 5.5 in the case $P =$"flat" is the so-called *flattening problem*, and was affirmatively solved by Raynaud and Gruson [38].

Theorem 5.6 (Raynaud-Gruson 1970 [38]). *Let $f \colon X \to S$ be a morphism of finite presentation such that $f \times_S U \colon X \times_S U \to U$ is flat. Then there exists a U-admissible blow-up $S' \to S$ such that the strict transform $f' \colon X' \to S'$ is flat of finite presentation.*

Among many valuable corollaries of this theorem, we refer to the following one.

Corollary 5.7 ([38, (5.7.12)]). *The full subcategory $\mathbf{BL}_{(S,U)}$ is cofinal in the category $\mathbf{MD}_{(S,U)}$.*

Remark 5.8. Here a few remarks on the flattening theorem are in order.

(1) The theorem is entirely clear in case $S = \operatorname{Spec} V$ where V is a discrete valuation ring. Indeed, in this case, one can take as $S' \to S$ the identity map $S' = S$, and thus the strict transform X' is the closed subscheme of X given by dividing out V-torsions.

(2) More generally, if $S = \operatorname{Spec} V$ where V is a (not necessarily discrete) valuation ring, then flatness of a similarly defined X' is clear by the same reasoning, whereas the finite presentation of f' is rather difficult to show.

5.4. Revival of Zariski's idea

In the rest of this section we are going to outline the proof of Theorem 5.6. The proof that we are going to present here is not the one in [38], but is done by Zariski's classical idea, which Zariski invented in order to apply it to the resolution of singularities of algebraic surfaces [44]. The keystone of Zariski's argument is the so-called *Zariski-Riemann space*, and the most crucial point of the proof is its quasi-compactness.

Definition 5.9 (Zariski-Riemann space; cf. [44][45]).

$$\langle U\rangle_{\mathrm{cpt}} = \varprojlim_{S'\in\mathbf{BL}_{(S,U)}} S',$$

where the projective limit is taken in the category of local ringed spaces.

Let us say that an ideal $\mathscr{J}$ of $\mathscr{O}_S$ is *admissible* if it is quasi-coherent of finite type and the corresponding closed subscheme $V(\mathscr{J})$ is set-theoretically contained in D. Then U-admissible blow-ups are exactly the morphisms of the form $\operatorname{Proj}\bigoplus_{n\geq 0}\mathscr{J}^n \to S$ by an admissible ideal $\mathscr{J}$. Hence the projective limit in Definition 5.9 is regarded as the filtered projective limit taken along the directed set of all admissible ideals with the ordering $\leq$ defined as follows: $\mathscr{J}\geq\mathscr{J}'$ if and only if there exists an admissible ideal $\mathscr{J}''$ such that $\mathscr{J}=\mathscr{J}'\mathscr{J}''$. This justifies Definition 5.9, for the category of local ringed spaces is closed under filtered projective limits. Note that the Zariski-Riemann space thus defined generalizes the so-called abstract Riemann surface, the introduction of which traces back to Dedekind-Weber in the 19th century, for if S is a regular curve then we have $\langle U\rangle_{\mathrm{cpt}} = S$.

Points. Let $x\in\langle U\rangle_{\mathrm{cpt}}$. The point x is, by definition, a compatible system of points $\{x_{S'}\}_{S'\in\mathbf{BL}_{(S,U)}}$ with $x_{S'}\in S'$ for any $S'\in\mathbf{BL}_{(S,U)}$.

- The topological space $\langle U\rangle_{\mathrm{cpt}}$ contains U. If $x\in U$, then the corresponding points $x_{S'}$ lie in the common U, and all of them are equal.
- If on the other hand $x\notin U$, then the system $\{x_{S'}\}$ is described in terms of a valuation ring[15] as follows: there exists a valuation ring V_x (of height ≥ 1) and a map $\alpha\colon \operatorname{Spec} V_x\to S$ of schemes mapping the closed point to x_S and the generic point to a point in U. For any U-admissible blow-up $S'\to S$, by the valuative criterion of properness, one has a unique arrow $\alpha'\colon \operatorname{Spec} V_x\to S'$ such that the resulting triangle

$$\begin{array}{ccc} & & S' \\ & \overset{\alpha'}{\nearrow} & \downarrow \\ \operatorname{Spec} V_x & \xrightarrow[\alpha]{} & S \end{array}$$

commutes. The point $x_{S'}$ is the image of the closed point by α'.

[15] See, for example, [46, Chap. VI] and [7, Chap. VI] for basics of valuation rings.

Local rings. The local rings of the structure sheaf $\mathscr{O}_{\langle U\rangle_{\mathrm{cpt}}}$ are best described by the following notion.

Definition 5.10. Let A be a ring, and I a finitely generated ideal. The ring A is said to be I*-valuative* if any finitely generated ideal J of A that contains I^k for some $k > 0$ (called an I*-admissible* ideal) is invertible.

In case A is a local ring, then A is I-valuative if and only if I is a principal ideal $I = (a)$ generated by a non-zero-divisor $a \in A$ and every I-admissible ideal is principal.

Proposition 5.11. (1) *Let A be a local ring, and $I = (a)$ a principal ideal generated by a non-zero-divisor $a \in A$. Set $J = \bigcap_{n\geq 1} I^n$. Suppose A is I-valuative. Then:*

(a) $B = A[\frac{1}{a}]$ *is a local ring, and $V = A/J$ is a valuation ring, which is $\overline{a}$-adically separated, where $\overline{a} = (a \bmod J)$;*
(b) $A = \{f \in B \,|\, (f \bmod \mathfrak{m}_B) \in V\}$, *where $\mathfrak{m}_B$ is the maximal ideal of B;*
(c) $J = \mathfrak{m}_B$.

(2) *Conversely, if B is a local ring and V is an $\overline{a}$-adically separated valuation ring for some non-zero $\overline{a} \in V$ such that the fractional field of V coincides with the residue field of B, then the subring $A = \{f \in B \,|\, (f \bmod \mathfrak{m}_B) \in V\}$ is an I-valuative local ring for any finitely generated ideal I such that $IV = (\overline{a})$, and $B = A[\frac{1}{a}]$.*

Proposition 5.11 shows that an I-valuative local ring is a "composite" of a local ring and a valuation ring. The following proposition follows from basic properties of U-admissible blow-ups, and is easy to verify.

Proposition 5.12. *For any point $x \in \langle U\rangle_{\mathrm{cpt}}$ the local ring $\mathscr{O}_{\langle U\rangle_{\mathrm{cpt}},x}$ is an $(\mathscr{I}_D\mathscr{O}_{\langle U\rangle_{\mathrm{cpt}},x})$-valuative ring.*

For $A_x = \mathscr{O}_{\langle U\rangle_{\mathrm{cpt}},x}$, we set $B_x = A_x[\frac{1}{a}]$ (where $I_x = \mathscr{I}_D\mathscr{O}_{\langle U\rangle_{\mathrm{cpt}},x} = (a)$) and $V_x = A_x/J_x$, where $J_x = \bigcap_{n\geq 1} I_x^n$. The local ring B_x is a local ring on U, and the valuation ring V_x is the one that describes the point $x = \{x_{S'}\}_{S'\in \mathbf{BL}_{(S,U)}}$ as above. In other words, each local ring of $\langle U\rangle_{\mathrm{cpt}}$ is a "composite" of a valution ring and a local ring of U.

Note that the above description of points and the local rings is essentially due to Zariski's original description of Zariski-Riemann space, which Zariski originally introduced not by projective limit of varieties, but as a certain space of places.

Remark 5.13. Notice that the valuation rings V_x that appear in the above context are not necessarily of height 1 even if the scheme S is Noetherian (e.g. an algebraic variety over a field). This is exactly the reason why valuation rings of higher height need to be considered in Zariski's argument. See Table 5 for the classification given by Zariski [44] of possible valuations that appear on algebraic surfaces.

Table 5. Valuation rings on algebraic surfaces

Height	Rational rank	
0	0	trivial valuation
1	1	divisorial
		non-divisorial
	2	non-divisorial
2	2	composite of two divisorial valuations

Intuitive description. Recall that the set of ideals of a valuation ring V is totally ordered by the inclusion order. In particular, the spectrum $\operatorname{Spec} V$ consists of points that are linearly configured as depicted in Figure 5. It can therefore be understood as a "long curve"[16] with

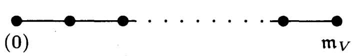

Fig. 5. Spectrum of valuation ring

the extremities (0), the generic point, and $\mathfrak{m}_V$, the closed points. Each point is the specialization of points sitting on its left (in the figure), and the generalization of points sitting on its right. In the finite height case, the height is the number of points minus one.

It is, therefore, appropriate to say that (the image of) a map $\operatorname{Spec} V \to S$ of schemes from a valuation ring is a "long path" in S. Intuitively, from what we have seen above in the description of points, one can say

[16]The adjective "long" indicates that it might be of large height.

that the space $\langle U\rangle_{\mathrm{cpt}}$ is like a "path space." More precisely, we have a set-theoretical decomposition

$$\langle U\rangle_{\mathrm{cpt}} = U \amalg T^{*}_{D/S},$$

where $T^{*}_{D/S}$ is the set of all "long paths" that pass through D, or is, so to speak, an analogue of a *tubular neighborhood*[17] of D in S; see Figure 6.

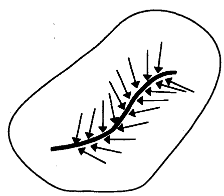

Fig. 6. Set-theoretical description of $\langle U\rangle_{\mathrm{cpt}}$

5.5. Quasi-compactness

The space $\langle U\rangle_{\mathrm{cpt}}$, being defined as the projective limit of all U-admissible blow-ups, would seem fairly gigantic. The following theorem, which turns out to be ineffably important, says that it is actually not.

Theorem 5.14 (Zariski 1944). *The space $\langle U\rangle_{\mathrm{cpt}}$ is quasi-compact.*

This theorem played one of the most essential roles in Zariski's proofs of resolution of singularities on algebraic surfaces (cf. §5.7) and Abhyankar's proof for three-folds. Also in our proof of the flattening theorem, quite similarly, this plays a very important role. The proof of Theorem 5.14 is by no means technical, but rather, one can say, the quintessence lies in a general principle applicable to a much wider situation.

One way of proof relies on the fact that the (2-categorical) filtered projective limit of coherent topoi with coherent transition maps is again coherent [SGA4-2, Exposé VI], which confers with the well-known fact that the filtered projective limit of compact Hausdorff spaces is again compact Hausdorff. Applying Deligne's theorem on the existence of points for locally coherent topoi, one shows the theorem.

[17]It might be more precise to say *deleted* tubular neighborhood.

A more handy way is provided by Stone's representation theorem, which asserts that the category of coherent topological spaces[18] and quasi-compact maps is categorically equivalent to the opposite category of unital distributive lattices (cf. [24]). As the latter category is closed under filtered direct limit, the theorem follows immediately (a minor point that should be confirmed here is that the direct limit taken in the category of topological spaces is equal to the one taken in the category of coherent topological spaces and quasi-compact maps).

In both proofs, the most important point is the following fact (existence of points): the filtered projective limit of non-empty coherent spaces with coherent transition maps is non-empty. An extensive use of this fact verifies the finite intersection property for open coverings, whence the quasi-compactness as desired. Notice that the above two ways of the proof are not entirely different from each other, and both arguments actually prove coherence, not only quasi-compactness.

5.6. Outline of the proof of Theorem 5.6

Now we can outline the proof of Theorem 5.6. The idea of the proof is the following.

> **Idea:** reduction to the case of "long curves" $\operatorname{Spec} V$ by means of quasi-compactness of Zariski-Riemann space.

This can be regarded as a "curve-cut" technique, which is quite often employed in algebraic geometry. In this sense, one can say that our approach is a quite geometric one.

First step. Observe first that the theorem is true for long curves $S = \operatorname{Spec} V$, where V is a valuation ring. As we have mentioned in Remark 5.8, the flattening theorem in this case is not easy, whereas the "flattening part" (without finiteness property) is trivial. The proof of the finiteness part has a quite different flavor from the other part; first, using composition of valuation rings, we reduce to the case of height 1, and then employ Gröbner basis arguments to show the finiteness. We omit the details here, and proceed to the general case, assuming the validity of the theorem in this case.

Second step. Observe next that the theorem is true for $S = \operatorname{Spec} A$, where A is the local ring at a point of $\langle U \rangle_{\mathrm{cpt}}$. Here we use the fact that the ring A is I-valuative, and the assertion follows from the previous

[18]A sober topological space is said to be *coherent* if it is quasi-compact, quasi-separated (i.e., the intersection of finitely many quasi-compact open subsets is quasi-compact), and has an open basis consisting of quasi-compact open subsets. Notice that this condition is equivalent to that the associated topos is coherent in the sense of [SGA4-2, Exposé VI].

step, the assumption that f is flat on U, and patching of flatness, where "patching" means composition in the sense of Proposition 5.11.

Third step. By the previous steps and the fact that the property $P =$"flat" is locally finitely presented, one deduces that the assertion is true locally on $\langle U\rangle_{\mathrm{cpt}}$, that is, for any point $x \in \langle U\rangle_{\mathrm{cpt}}$, there exists an admissible blow-up $S' \to S$ and a quasi-compact open subset U_x of S' such that

- U_x contains the image of x by the projection $\langle U\rangle_{\mathrm{cpt}} \to S'$;
- $f'|_{U_x}$ is flat and finitely presented, where $f'\colon X' \to S'$ is the strict transform of f.

Here we have tacitly used the following *extension of admissible ideals.*

Proposition 5.15. *Let $T \subset S$ be a quasi-compact open subset of S, and $\mathscr{J}$ an admissible ideal on T (with respect to $V = U \cap T$). Then there exists an admissible ideal $\widetilde{\mathscr{J}}$ on S such that $\widetilde{\mathscr{J}}|_T = \mathscr{J}$.*

Fourth step. Finally, by quasi-compactness (Theorem 5.14), the assertion follows by birational patching. More precisely, there exist finitely many points x_i $(i = 1, \ldots, n)$ such that $\langle U\rangle_{\mathrm{cpt}} = \bigcup_{i=1}^n p_i^{-1}(U_{x_i})$, where, for each i, $U_{x_i} \subset S_i$, and $p_i\colon \langle U\rangle_{\mathrm{cpt}} \to S_i$ is the projection map. Take $S' \in \mathbf{BL}_{(S,U)}$ that dominates the S_i's. Replacing S' by the blow-up along $\mathscr{I}_D\mathscr{O}_{S'}$, we may assume that $\mathscr{I}_D\mathscr{O}_{S'}$ is an invertible ideal. Let U'_i be the pull-back of U_{x_i} by the map $S' \to S_i$. Then $S' = \bigcup_{i=1}^n U'_i$, and thus the strict transform $f'\colon X' \to S'$ is flat and finitely presented.

5.7. Other applications

The argument of the above type, which uses the quasi-compactness of Zariski-Riemann spaces, was largely applicable to several other situations. Let us list some of them (which are, however, not new).

Resolution of singularities of quasi-excellent surfaces. This is the one to which Zariski originally applied this argument (in the case of algebraic sufaces). Similarly to the above-mentioned procedure, one first reduces the claim to the case of "long curves" to show that resolution can be done locally (local uniformization), and then patches the resulting local resolutions into a regular model by using quasi-compactness of Zariski-Riemann space. See, for example, [31, Chap. I] for more details.

Embedding theorem for algebraic spaces (cf. Nagata 1963). This asserts that a separated algebraic space of finite type over a coherent scheme can be embedded into a proper space. This was first proved by Nagata in 1963 for Noetherian separated schemes. Considering "long curves", one first observes that, locally, an appropriate embedding can be constructed (local extension lemma), and then birationally patches these locally extended data into a globally extended space, which is possible because of the quasi-compactness of Zariski-Riemann space. Finally by the valuative criterion of properness, one shows that the resulting space is proper.

Remark 5.16. There are two remarks in order on the Nagata's embedding theorem.

(1) The theorem is true not only for schemes but also for algebraic spaces. There are several motivations for this generalization. One of them will be seen below (Theorem 6.11). Another motivation is that the embedding theorem potentially has large applications to the compactification of moduli spaces that are usually not representable by schemes; e.g. M. Rapoport's Habilitationschrift. We also remark that this generalized form of Nagata's embedding theorem has an application to trace formula; cf. Remark 8.3.

(2) One can actually simplify the proof of the embedding theorem by using ideas from rigid geometry. The details will be shown in [21].

§6. Birational approach to rigid geometry

6.1. Introduction

Now we come to the stage of expounding our approach to rigid geometry. As we have briefly announced in §4 our general policy is that rigid geometry is a hybrid of formal geometry and birational geometry; here, in our approach to rigid geometry, we will see that Zariski's classical idea of birational geometry explained in §5 revives, and plays one of the most important roles. One can thus refine the picture from Figure 3 into the one from Figure 7.

As Raynaud's viewpoint of rigid geometry takes up geometry of formal schemes as the starting point, from which rigid geometry is supposed to arise in the way that birational changes by admissible blow-ups are inverted, the birational geometry on the right-hand side means, so to speak, *birational geometry of formal schemes*, which should be a theory of a formal analogue of Zariski-Riemann spaces. One arrives in this way at the following "central dogma," which realizes more concretely our general policy for approaching rigid geometry:[19]

$$\boxed{\text{Birational geometry of formal schemes}} \quad = \quad \boxed{\text{Rigid geometry}}.$$

The analogue of Zariski-Riemann spaces in this context gives rise to the so-called *Zariski-Riemann triples* (Definition 6.15), which provide

[19]There is no reason why we should deal only with formal schemes, and a perhaps more reasonable formulation would be given by allowing formal spaces (= formal algebraic spaces) to enter in. See Theorem 6.12.

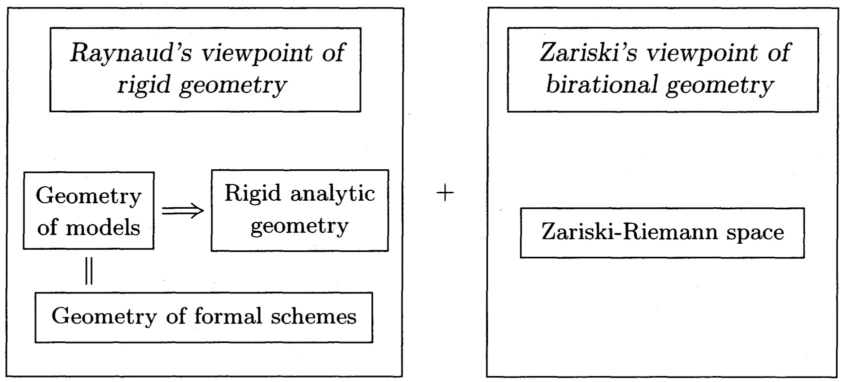

Fig. 7. Birational approach to rigid geometry

for each rigid space a topological space together with two sheaves of local rings, the *integral structure sheaf* and the *rigid structure sheaf*. We think of these objects as the most basic figure in rigid geometry in which rigid analytic and formal geometric aspects are amalgamated and crystalized in a certain canonical way. Moreover, the admissible topology of a rigid space is honestly represented by the topology of the underlying topological space of the corresponding Zariski-Riemann triple. In this sense, one can say that the Zariski-Riemann triple *visualizes* the rigid space (cf. Proposition 6.16).

Our basic dictionary of comparing the situation of birational geometry as in §5.1 with that of, say, p-adic rigid geometry is as follows:

- $S \longleftrightarrow$ formal scheme of finite type over $\mathrm{Spf}\,\mathbb{Z}_p$;
- $D \longleftrightarrow$ the closed fiber, that is, the closed subscheme defined by "$p = 0$."

Note that, by means of this comparison, the notion of U-admissible blow-ups as in §5.2 precisely correspond to the admissible blow-ups introduced in §3.3. The object corresponding to the classical Zariski-Riemann space $\langle U\rangle_{\mathrm{cpt}}$ is the underlying topological space of the Zariski-Riemann triple arising from formal schemes.

6.2. Adequate formal schemes

We have seen in Remark 5.13 that, in Zariski's approach to birational geometry, one needs to consider valuation rings of large height in general, even when dealing with Noetherian schemes. It turns out, for the same reason, that valuation rings of higher height have to be considered also in our situation. Indeed, even when we deal with Noetherian formal schemes to define rigid spaces, points are described by means of valuation rings. However, the valuation rings that enter in this situation may be of height greater than 1. Note that, without such valuation rings, or, as a result, without enough points, one cannot detect topology and sheaves. Hence, as such valuation rings are rarely Noetherian, one almost always has to deal with non-Noetherian formal schemes, of which we lack sufficiently practical knowledge; even in [EGA], apart from the generalities at the first set-up, most of the theorems, such as finitudes, GFGA, etc., are proven under the Noetherian hypothesis. Thus, one first has to establish a class of formal schemes that is wide enough to contain Noetherian and some other hitherto considered classes of formal schemes (such as formal spectra of a-adically complete valuation rings), and to generalize the necessary theorems.

The new class of adic formal schemes that we would like to offer here is that of so-called *adequate* formal schemes. We postpone the precise definition of them to another opportunity [21], and confine ourselves to the following rough explanation.

Basic properties.

- the definition is given ring theoretically;
- the rings are Noetherian outside the ideal of definition.

Objects. Let $\mathbf{Fs}^{\mathrm{adq}}$ denote the category of adequate formal schemes. It contains as objects

- $\operatorname{Spf} V$, where V is an a-adically complete valuation ring for some non-zero $a \in \mathfrak{m}_V$,
- Noetherian formal schemes.

Functoriality. The category $\mathbf{Fs}^{\mathrm{adq}}$ has the following pleasant functoriality: it is

- closed under finite type extensions;
- closed under base change by finite type morphisms.

Figure 8 depicts the category $\mathbf{Fs}^{\mathrm{adq}}$ together with some subcategories, where $\mathbf{Fs}^{\mathrm{Noe}}$, $\mathbf{Fs}^{\mathrm{fin}}_{/V}$, $\mathbf{Fs}^{\mathrm{fin}}_{/\mathrm{DVR}}$ denote respectively the categories of Noetherian formal schemes, of formal schemes of finite type over an a-adically complete valuation ring, and of formal schemes of finite type over a complete discrete valuation ring. Notice that:

- the height of valuation rings appearing in the category $\mathbf{Fs}^{\mathrm{adq}}$ is arbitrary, finite or infinite;

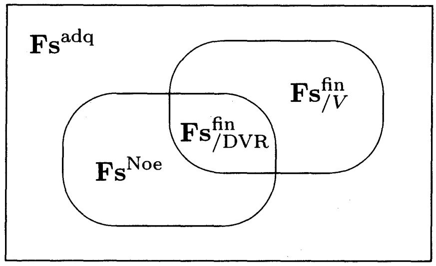

Fig. 8. Category of adequate formal schemes

- more importantly, the category $\mathbf{Fs}^{\mathrm{adq}}$ contains all objects of the form Spf A, where A is a formal model of an affinoid algebra in Tate's theory of rigid analytic geometry.

Among several nice points of adequate formal schemes, we would like to announce that most of the important theorems, such as finitudes, GFGA comparison, GFGA existence theorems, can be proved in this category, which therefore gives generalizations of the theorems in [EGA, **III**]. The details will be shown in [21]. There, these theorems are stated and proved entirely by using systematically the derived categorical framework.[20]

6.3. Coherent rigid spaces

Let us denote by $\mathbf{CFs}^{\mathrm{adq}}$ the category of coherent adequate formal schemes.

Proposition 6.1. (1) *Any coherent (= quasi-compact and quasi-separated) adequate formal scheme has an ideal of definition of finite type.*

[20]Let us list two reasons why it is necessary to work in the derived categorical language: (1) it is user-friendly for applications; (2) recently, the importance of derived categories has been more and more recognized in algebraic geometry and in mathematical physics. The last point is related to the cohomological mirror symmetries speculated on by Kontsevich-Soibelman and Fukaya et al., in which our theorems in terms of the derived categorical language, as well as our approach involving higher-height valuation rings, could be important.

(2) *Let X be a coherent adequate formal scheme, and $\mathscr{I}$ an ideal of definition. If $\mathscr{O}_X$ is $\mathscr{I}$-torsion free, then $\mathscr{O}_X$ is coherent*[21] *as a module over itself.*

Definition 6.2 (Admissible ideal). Let X be a coherent adequate formal scheme, and $\mathscr{J}$ an ideal of $\mathscr{O}_X$. Then $\mathscr{J}$ is said to be *admissible* if it is an adically quasi-coherent open ideal of finite type.

Here an $\mathscr{O}_X$-module $\mathscr{F}$ is said to be *adically quasi-coherent* if the following conditions are satisfied:

(a) $\mathscr{F}$ is complete with respect to $\mathscr{I}$-adic topology, where $\mathscr{I}$ is an ideal of definition of X;
(b) for any $k \geq 0$, the sheaf $\mathscr{F}_k = \mathscr{F}/\mathscr{I}^{k+1}\mathscr{F}$ is a quasi-coherent sheaf on the scheme $X_k = (X, \mathscr{O}_X/\mathscr{I}^{k+1})$.

Definition 6.3 (Admissible blow-up). Let X be a coherent adequate formal scheme, and $\mathscr{J}$ an admissible ideal. The *admissible blow-up* along $\mathscr{J}$ is the morphism of formal schemes

$$X' = \varinjlim_{k\geq 0} \operatorname{Proj}\left(\bigoplus_{n\geq 0} \mathscr{J}^n \otimes \mathscr{O}_{X_k}\right) \longrightarrow X,$$

where $X_k = (X, \mathscr{O}_X/\mathscr{I}^{k+1})$ is the scheme defined as above.

As X' is clearly of finite type over X, X' is again a coherent adequate formal scheme. Notice that the above definition of admissible blow-ups does not depend on the choice of an ideal of definition $\mathscr{I}$.

Having obtained a nice category of formal schemes and a nice notion of admissible blow-ups, we can now define rigid spaces in our approach by applying Raynaud's idea.

Definition 6.4 (Coherent rigid spaces). The category **CRf** of coherent rigid spaces is defined to be the quotient category of $\mathbf{CFs}^{\mathrm{adq}}$ where all admissible blow-ups are inverted:

$$\mathbf{CRf} = \mathbf{CFs}^{\mathrm{adq}}/\{\text{admissible blow-ups}\}.$$

We denote the quotient functor $\mathbf{CFs}^{\mathrm{adq}} \to \mathbf{CRf}$ by

$$X \longmapsto X^{\mathrm{rig}}.$$

[21] Perhaps the reader might complain that there is too much use of "coherent." Do not mix up the coherence of sheaves and the coherence of spaces.

For a coherent rigid space $\mathscr{X}$, a *formal model* of $\mathscr{X}$ is defined to be a coherent adequate formal scheme X such that $X^{\mathrm{rig}} \cong \mathscr{X}$. A formal model X of $\mathscr{X}$ is said to be *distinguished* if $\mathscr{O}_X$ is $\mathscr{I}$-torsion free, where $\mathscr{I}$ is an ideal of definition of X.

6.4. Admissible topology

Definition 6.5. (1) A morphism $\mathscr{U} \to \mathscr{X}$ of coherent rigid spaces is said to be a (*coherent*) *open immersion* if it has as a formal model an open immersion $U \hookrightarrow X$.

(2) Let $\{\mathscr{U}_\alpha \hookrightarrow \mathscr{X}\}$ be a family of open immersions between coherent rigid spaces. We say that the family is a covering with respect to the *admissible topology* if it has a finite refinement $\{\mathscr{V}_i \hookrightarrow \mathscr{X}\}$ satisfying the following condition: there exist a formal model X of $\mathscr{X}$ and formal models $V_i \hookrightarrow X$ of $\mathscr{V}_i \hookrightarrow \mathscr{X}$ such that $X = \bigcup V_i$.

The last notion gives rise to a topology on $\mathbf{CRf}$, called the admissible topology. The resulting site is denoted by $\mathbf{CRf}_{\mathrm{ad}}$.

6.5. General rigid spaces

The category $\mathbf{CFs}^{\mathrm{adq}}$ is a good category that allows "formal birational patching"; the following statement is a consequence of the existence of formal birational patching of morphisms.

Proposition 6.6. *Any representable presheaf on* $\mathbf{CRf}_{\mathrm{ad}}$ *is a sheaf.*

The proposition allows a consistent definition of more general rigid spaces.

Definition 6.7 (General rigid spaces). A *general rigid space* is a sheaf $\mathscr{F}$ of sets on the site $\mathbf{CRf}_{\mathrm{ad}}$ such that the following conditions are satisfied:

(1) there exists a surjective map of sheaves

$$\coprod_{i \in I} \mathscr{Y}_i \longrightarrow \mathscr{F},$$

where $\{\mathscr{Y}_i\}_{i \in I}$ is a collection of sheaves represented by coherent rigid spaces;

(2) for $i, j \in I$, the map $\mathscr{Y}_i \times_{\mathscr{F}} \mathscr{Y}_j \longrightarrow \mathscr{Y}_i$ is isomorphic to the direct limit of a direct system $\{\mathscr{U}_\lambda \to \mathscr{Y}_i\}_{\lambda \in \Lambda}$ of maps between coherent rigid spaces such that all maps in the commutative diagram for $\mu \leq \lambda$

$$\begin{array}{ccc} \mathscr{U}_\lambda & \longrightarrow & \mathscr{Y}_i \\ \uparrow & \nearrow & \\ \mathscr{U}_\mu & & \end{array}$$

are coherent open immersions.

We denote by **Rf** the category of general rigid spaces. It has **CRf** as a full subcategory.

Example 6.8. Here is an example of (coherent) rigid spaces that cannot be dealt with in classical rigid geometry (by Tate). Consider the ring $\mathbb{Z}[[q]]$ of formal power series with integral coefficients. This ring is not a valuation ring, but is a complete ring with respect to the q-adic topology. Hence we can consider the formal scheme $S = \operatorname{Spf} \mathbb{Z}[[q]]$, which is clearly adequate, since it is Noetherian. Any adic formal scheme X of finite type over S therefore gives rise to a rigid space $\mathscr{X} = X^{\mathrm{rig}}$ over $\mathscr{S} = S^{\mathrm{rig}}$. A particularly important example of this form, which we will discuss later in §7.1, is a *Tate curve* over $\mathscr{S}$.

Rigid spaces of the above form over $(\operatorname{Spf} \mathbb{Z}[[q]])^{\mathrm{rig}}$ (or higher dimensional adic rings) enter quite naturally in discussions on compactification of moduli spaces. Although such kinds of rigid spaces are ruled out in the classical rigid geometry, they come rather naturally in our approach to rigid geometry, and this proves to be one of the advantages of our approach.

6.6. Fiber products

A morphism $\varphi\colon \mathscr{X} \to \mathscr{Y}$ of coherent rigid spaces is said to be *of finite type* if it has a formal model $f\colon X \to Y$ that is of finite type. The notion of "locally of finite type" is defined for morphisms between general rigid spaces in an obvious way. The following proposition follows from the fact that the adequateness of formal schemes is closed under base change locally of finite type (as we have mentioned in §6.2).

Proposition 6.9. *Consider the diagram*

$$\mathscr{X} \xrightarrow{\varphi} \mathscr{S} \xleftarrow{\psi} \mathscr{Y}$$

in **Rf**. *If either one of the morphisms is locally of finite type, then the fiber product* $\mathscr{X} \times_{\mathscr{S}} \mathscr{Y}$ *is representable in* **Rf**.

Remark 6.10. As we will see later (Remark 6.17), for a rigid space $\mathscr{X}$, points (in a certain topos-theoretic sense) correspond to valuation rings of a certain kind; that is, points are represented by morphisms of the form $(\operatorname{Spf} V)^{\mathrm{rig}} \to \mathscr{X}$, where V is an a-adically complete valuation ring. Hence, in our rigid geometry, "fibers over points" are those fiber products taken with morphisms of this kind. The importance of studying rigid spaces over rigid spaces of the form $(\operatorname{Spf} V)^{\mathrm{rig}}$ thus arises. Notice that, even if we work in the categories of rigid spaces coming from Noetherian formal schemes, valuation rings V of higher height are inevitable.

Technically, the importance of the last remark lies in the fact that, by considering fibers over points, one can usually reduce quite a few geometric properties of rigid spaces of finite type to those of rigid spaces of finite type over valuation rings. In case the valuation ring is of finite height, one can further reduce to the case of height 1 (by the gluing method), where one can use some extra tools, such as Noether's normalization theorem, etc.

6.7. Relation with algebraic spaces

Let $\mathrm{Spf}\, A$ be an affine adequate formal scheme, and I a finitely generated ideal of definition of A. We set $U = \mathrm{Spec}\, A \setminus V(I)$, which is a Noetherian scheme (cf. §6.2). The precise meaning of the following somewhat vague statement will be clarified in [21].

Theorem 6.11 (GAGA functor). *The GAGA functor*

$$\left\{\begin{matrix}\text{Separated algebraic spaces}\\ \text{of finite type } /U\end{matrix}\right\} \longrightarrow \mathbf{Rf}_{\mathscr{S}}, \qquad X \mapsto X^{\mathrm{an}},$$

where $\mathscr{S} = (\mathrm{Spf}\, A)^{\mathrm{rig}}$, *exists.*

Notice that the GAGA functor of this general form has not been defined even in the classical rigid geometry (at least in literature). There are two main ingredients for the proof. One is the embedding theorem (of Nagata) for algebraic spaces, and the other one is the following.

Theorem 6.12 (Equivalence theorem). *Let S be a coherent adequate formal scheme. Then the natural functor*

$$\left\{\begin{matrix}\text{Formal schemes}\\ \text{of finite type } /S\end{matrix}\right\} \Big/ \left\{\begin{matrix}\text{Admissible}\\ \text{blow-ups}\end{matrix}\right\} \longrightarrow \left\{\begin{matrix}\text{Formal alge-}\\ \text{braic spaces of}\\ \text{finite type } /S\end{matrix}\right\} \Big/ \left\{\begin{matrix}\text{Admissible}\\ \text{blow-ups}\end{matrix}\right\}$$

is a categorical equivalence.

This follows from the following theorem.

Theorem 6.13. *Let S be as above, and $X \to S$ a formal algebraic space of finite type. Then there exists an admissible blow-up $X' \to X$ such that X' is a formal scheme.*

The proof of this theorem uses (again!) the technique of Zariski-Riemann spaces.

6.8. Tate's rigid analytic spaces

Tate's rigid analytic spaces are naturally objects of the category **Rf** via Raynaud's theorem (Theorem 3.5) and obvious patching arguments, that is, we have the natural functor

$$\left\{\begin{matrix}\text{Tate's rigid}\\ \text{spaces}\end{matrix}\right\} \longrightarrow \mathbf{Rf},$$

which maps affinoids to affinoids. Here by an affinoid in **Rf** we mean a coherent rigid space of the form $(\operatorname{Spf} A)^{\mathrm{rig}}$.

The essential image of the above functor considered on the category of Tate's rigid analytic spaces over K is the category of rigid spaces locally of finite type over $(\operatorname{Spf} V)^{\mathrm{rig}}$, where V is a complete valuation ring of height 1, and K is its fractional field. Note that this is essentially the assertion of Raynaud's theorem (Theorem 3.5).

6.9. Visualization

The moral basis of our (and hence Raynaud's) defining rigid spaces as "generic fibers" of formal schemes stems from the policy that rigid geometry is so to speak the birational geometry of formal schemes (cf. §6.1). It being so, one can say that the visualization of rigid spaces, which we are going to pursue below, is the way to enhance the birational geometric aspect of rigid geometry. It does this job by adopting Zariski's old idea of birational geometry, and the visualization itself is given by the so-called Zariski-Riemann triple. The pleasant thing is that the admissible topology attached to a rigid space is equivalent to the topology (in the usual sense) of the associated Zariski-Riemann space, the underlying topological space of the Zariski-Riemann triple. This is the origin of the name "visualization." As we can easily imagine, having the genuine ringed space that really represents the rigid space helps and streamlines discussions, and enables many applications.

Definition 6.14. Let $\mathscr{X} = X^{\mathrm{rig}}$ be a coherent rigid space.

(1) Define the projective limit

$$\langle \mathscr{X} \rangle = \varprojlim_{X' \to X} X'$$

along all admissible blow-ups of X taken in the category of local ringed spaces. Note that, by the similar reasoning as in §5.4, the projective limit can be replaced by the filtered projective limit taken along the directed set of all admissible ideals, and hence is well-defined as a local ringed space. The canonical projection map $\langle \mathscr{X} \rangle \to X'$ for any admissible blow-up X' of X is called the *specialization map*, and is denoted by

$$\mathrm{sp}_{X'}\colon \langle \mathscr{X} \rangle \longrightarrow X'.$$

This is a continuous map.

(2) The structure sheaf of $\langle \mathscr{X} \rangle$, which is the direct limit of the sheaf $\mathrm{sp}_{X'}^{-1}\mathscr{O}_{X'}$, is called the *integral structure sheaf*, and is denoted by $\mathscr{O}_{\mathscr{X}}^{\mathrm{int}}$.

(3) The *rigid structure sheaf* $\mathscr{O}_{\mathscr{X}}$ is the sheaf on $\langle \mathscr{X} \rangle$ defined by

$$\mathscr{O}_{\mathscr{X}} = \varinjlim_{n \geq 0} \mathscr{H}om_{\mathscr{O}_{\mathscr{X}}^{\mathrm{int}}}(\mathscr{I}^n, \mathscr{O}_{\mathscr{X}}^{\mathrm{int}});$$

here we take an ideal of definition $\mathscr{I}_X$ of X and set $\mathscr{I} = (\mathrm{sp}_X^{-1}\mathscr{I}_X)\mathscr{O}_{\mathscr{X}}^{\mathrm{int}}$.

Here the definition of $\mathscr{O}_{\mathscr{X}}$ calls for an explanation. It turns out that the sheaf $\mathscr{O}_{\mathscr{X}}^{\mathrm{int}}$ of local rings is $\mathscr{I}$-valuative, and due to Proposition 5.11, one sees that the sheaf $\mathscr{O}_{\mathscr{X}}$ is also a sheaf of local rings. For example, in the p-adic situation, we have $\mathscr{O}_{\mathscr{X}} = \mathscr{O}_{\mathscr{X}}^{\mathrm{int}}[\frac{1}{p}]$. As this particular example indicates, it is $\mathscr{O}_{\mathscr{X}}$ that plays the role of the structure sheaves of Tate's rigid analytic geometry. In fact, when $\mathscr{X}$ comes from a rigid analytic space in the sense of Tate via the functor as in §6.8, $\mathscr{O}_{\mathscr{X}}$ "is" the structure sheaf of the original Tate rigid space. Thus the realization of rigid spaces as a topological space $\langle \mathscr{X} \rangle$ naturally weaves its structure sheaf $\mathscr{O}_{\mathscr{X}}$ with, one can say, its "canonical" formal model $\mathscr{O}_{\mathscr{X}}^{\mathrm{int}}$.

Definition 6.15 (Zariski-Riemann triple)**.** We write

$$\mathbf{ZR}(\mathscr{X}) = (\langle \mathscr{X} \rangle, \mathscr{O}_{\mathscr{X}}^{\mathrm{int}}, \mathscr{O}_{\mathscr{X}}),$$

and call it the *Zariski-Riemann triple* associated to the rigid space $\mathscr{X}$.

One can, in fact, extend the above definition to general rigid spaces by gluing. It is worth remarking here that the idea of considering the triple as above, rather than merely a local ringed space, comes from the analogy between hermitian vector bundles $(\mathscr{E}, |\cdot|)$ and pairs $(\mathscr{E}, \mathscr{E}^{\mathrm{int}})$ of vector bundle with its integral model (which is at the center of the idea of, for example, Arakelov geometry).

Be that as it may, the main motivation for introducing Zariski-Riemann triples is that they really visualize the rigid spaces, as the following proposition indicates.

Proposition 6.16. *The topos associated to the topological space* $\langle \mathscr{X} \rangle$ *is isomorphic to the admissible topos* $\mathscr{X}_{\mathrm{ad}}^{\sim}$.[22]

[22]An essentially equivalent statement was proved by Huber [23]; a similar but different approach was taken by van der Put-Schneider [42].

Remark 6.17. The last proposition forces us to review the issue of points of rigid analytic spaces, which was already considered during the discussion of Tate's basic idea of approaching rigid geometry (see Remark 2.2). Even in case where $\mathscr{X}$ comes from a Tate rigid analytic space, the topological space $\langle\mathscr{X}\rangle$ was not considered by Tate, since Tate's notion of points only grasps points coming from maximal ideals of affinoid algebras, which occupies only a very small part of the space $\langle\mathscr{X}\rangle$. This is why Tate had to introduce the Grothendieck topology machinery to obtain the admissible topology.

Now, the space $\langle\mathscr{X}\rangle$ gives the correct notion of points for rigid analytic spaces; in fact, quite similarly to §5.4, points of $\langle\mathscr{X}\rangle$ are described in terms of a-adically complete valuation rings. It is based on this fact that we say that the Zariski-Riemann triple *visualizes* rigid spaces.

Remark 6.18. We would like to mention that, by using visualization, one can simplify the definition of the so-called "dagger-ring" that appears in the theory of rigid cohomology (cf. [4]). Let us give a simple example. Let $A = V\langle\langle X\rangle\rangle$, where V is a complete discrete valuation ring of mixed characteristic $(0, p)$ such that the residue field k is perfect, and consider $\mathbb{D} = (\operatorname{Spf} A)^{\mathrm{rig}}$ (closed unit disk). It is a coherent open rigid subspace of the projective line $\mathbb{P}^1_{\mathscr{S}} = ((\mathbb{P}^1_V)^{\wedge})^{\mathrm{rig}}$, where $\mathscr{S} = (\operatorname{Spf} V)^{\mathrm{rig}}$. Consider the closure $\overline{\langle\mathbb{D}\rangle}$ of $\langle\mathbb{D}\rangle$ in $\langle\mathbb{P}^1_{\mathscr{S}}\rangle$. Consider the sheaf $\mathscr{O}^{\dagger}_{\mathscr{D}}$ defined by the pull-back $i^*\mathscr{O}_{\mathbb{P}^1_{\mathscr{S}}}$ of $\mathscr{O}_{\mathbb{P}^1_{\mathscr{S}}}$ by the inclusion $i\colon \overline{\langle\mathbb{D}\rangle} \hookrightarrow \langle\mathbb{P}^1_{\mathscr{S}}\rangle$. The dagger-ring in this case, usually denoted by $A^{\dagger}_K$, is the ring $\Gamma(\overline{\langle\mathbb{D}\rangle}, i^*\mathscr{O}_{\mathbb{P}^1_{\mathscr{S}}})$.

6.10. Relation with other theories

Now let us mention something about the relation between our approach to rigid geometry and other hitherto known approaches.

(1) As we have already mentioned in §6.8, there exists a natural functor that maps Tate's rigid analytic spaces to rigid spaces in **Rf**.

(2) Zariski-Riemann triples are regarded as Huber's adic spaces, whence we have a natural functor

$$\mathbf{ZR}\colon \mathbf{Rf} \longrightarrow \left\{\begin{matrix}\text{Huber's adic}\\ \text{spaces}\end{matrix}\right\},$$

which is, however, *not fully faithful* in general; but it is fully faithful in practically important situations.

(3) Each Zariski-Riemann space $\langle\mathscr{X}\rangle$ admits, by means of maximal generalization[23] of all points, a so-called *separation map*

$$\mathrm{sep}_{\mathscr{X}}\colon \langle\mathscr{X}\rangle \longrightarrow [\mathscr{X}],$$

where $[\mathscr{X}]$ is the set of all points of $\langle\mathscr{X}\rangle$ of height 1. The map $\mathrm{sep}_{\mathscr{X}}$ is a continuous map. At least in case where $\mathscr{X}$ comes from a Tate rigid analytic space via the functor as in §6.8, the target space $[\mathscr{X}]$ (with more structure coming from $\mathscr{X}$) can naturally be regarded as a Berkovich space ([2][3]).

Figure 9 depicts the above mentioned relations.

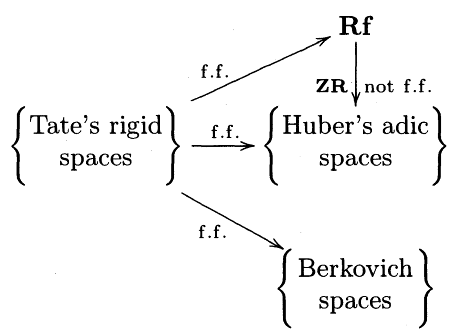

Fig. 9. Relation with other theories (f.f.= fully faithful)

6.11. Formal flattening theorem

Applying Zariski's idea explained in §5, but now using the Zariski-Riemann triple introduced as above, one can show the following theorem.

Theorem 6.19 (Bosch-Raynaud, Fujiwara). *Let $f\colon X \to S$ be a morphism of finite type between coherent adequate formal schemes. Then the following conditions are equivalent:*

(1) *$f^{\mathrm{rig}}\colon X^{\mathrm{rig}} \to S^{\mathrm{rig}}$ is flat, that is, $\langle f^{\mathrm{rig}}\rangle\colon \langle X^{\mathrm{rig}}\rangle \to \langle S^{\mathrm{rig}}\rangle$ is flat as a mapping of local ringed spaces (with the rigid structure sheaf);*

[23] It can be shown that any point of $\langle\mathscr{X}\rangle$ has a unique maximal generalization.

(2) *there exists an admissible blow-up $S' \to S$ such that the strict transform $f'\colon X' \to S'$ is flat.*

Corollary 6.20. *Admissible blow-ups are cofinal in the category of formal modifications.*

The following corollary is, as we have already seen in §3.3, important in Tate's rigid analytic geometry (see Theorem 2.8).

Corollary 6.21 (Gerritzen-Grauert). *Let $\varphi\colon \mathscr{X} \to \mathscr{Y}$ be a morphism of Tate's rigid analytic spaces over a complete non-archimedean valued field K with non-trivial valuation. Then the following conditions are equivalent:*

(1) *φ is an open immersion;*
(2) *φ is separated, étale, and injective, and induces an isomorphism between the residue fields at any point.*

6.12. Properness in rigid geometry

Definition 6.22. (1) A morphism $\varphi\colon \mathscr{X} \to \mathscr{Y}$ of rigid spaces is said to be *closed* if the induced map $\langle\varphi\rangle\colon \langle\mathscr{X}\rangle \to \langle\mathscr{Y}\rangle$ of topological spaces is closed.

(2) Let $\varphi\colon \mathscr{X} \to \mathscr{Y}$ be a morphism locally of finite type. The morphism φ is said to be *universally closed* if, for any morphism $\mathscr{Z} \to \mathscr{Y}$ of rigid spaces, the base change $\varphi_{\mathscr{Z}}\colon \mathscr{X} \times_{\mathscr{Y}} \mathscr{Z} \to \mathscr{Z}$ is closed.

Definition 6.23. A morphism $\varphi\colon \mathscr{X} \to \mathscr{Y}$ of rigid spaces is said to be *proper* if it is universally closed, separated, and of finite type.

In case $\mathscr{X}$ and $\mathscr{Y}$ are coherent, according to our general policy of regarding rigid geometry as birational geometry of formal schemes, the properness thus defined should be equivalent to that in formal geometry as follows.

Proposition 6.24. *Let $\varphi\colon \mathscr{X} \to \mathscr{Y}$ be a morphism of coherent rigid spaces. Then the following conditions are equivalent:*

(1) *φ is proper;*
(2) (Raynaud properness) *there exists a proper formal model $f\colon X \to Y$ of φ.*
(3) (Kiehl properness) *there exist affinoid enlargements* (*cf.* [26]) *of coverings for each relatively compact affinoid open subset; that is, there exists a finite admissible covering $\mathscr{X} = \bigcup \mathscr{U}_i$ consisting of affinoids together with a refinement $\mathscr{X} = \bigcup \mathscr{V}_i$ again consisting of affinoids with $\mathscr{U}_i \hookrightarrow \mathscr{V}_i$ such that, for each i, $\overline{\langle\mathscr{U}_i\rangle} \subset \langle\mathscr{V}_i\rangle$, where the closure is taken in $\langle\mathscr{X}\rangle$.*

Historically, properness in Tate's rigid geometry has been first defined by Kiehl in his work [26] on finiteness theorem; there, properness was defined by existence of enlargements according to the general idea by Cartan-Serre and Grauert for proving finiteness of cohomologies of coherent sheaves.

Whereas the implication (3) $\Rightarrow$ (2) is in general not difficult to show, the converse is a very difficult theorem; even in case where all rigid spaces are of finite type over $(\mathrm{Spf}\, V)^{\mathrm{rig}}$ with V being a complete discrete valuation ring, Lütkebohmert's 1990 paper [32] was the first for the proof. We claim (in [21]) that this is also valid in general. In the case over $(\mathrm{Spf}\, V)^{\mathrm{rig}}$ where V is an a-adically complete valuation ring, this amounts to showing the following statement.

Theorem 6.25. *Let $f\colon X \to \mathrm{Spf}\, V$ be a morphism of adequate formal schemes of finite type, and $U \subset X$ an affine open subset such that $\overline{U}$ is proper. Then there exists an admissible blow-up $\pi\colon X' \to X$ and an open subset $W \subset X'$ such that the following conditions are satisfied:*

(a) $\overline{\pi^{-1}(U)} \subseteq W$;
(b) *there exists a map $W \to \mathrm{Spf}\, A$ to an affine adequate formal scheme that is a contraction (that is, $W^{\mathrm{rig}} = (\mathrm{Spf}\, A)^{\mathrm{rig}}$).*

6.13. Cohomology theory

Let $\mathscr{X}$ be a rigid space, and $\mathscr{F}$ an abelian sheaf on the topological space $\langle \mathscr{X} \rangle$. We write

$$\mathrm{H}^q(\mathscr{X}, \mathscr{F}) = \mathrm{H}^q(\langle \mathscr{X} \rangle, \mathscr{F}).$$

Similarly, for a morphism $\varphi\colon \mathscr{X} \to \mathscr{Y}$ of rigid spaces, we write

$$\mathrm{R}^q\varphi_*\mathscr{F} = \mathrm{R}^q\langle\varphi\rangle_*\mathscr{F}.$$

As we have mentioned before, an affinoid is a coherent rigid space of the form X^{rig}, where X is an affine adequate formal scheme. For a rigid space $\mathscr{X}$, by a *coherent sheaf* on $\mathscr{X}$, we mean a coherent $\mathscr{O}_{\mathscr{X}}$-module on $\langle \mathscr{X} \rangle$.

Theorem 6.26. *For a rigid space $\mathscr{X}$, the rigid structure sheaf $\mathscr{O}_{\mathscr{X}}$ is coherent.*

Thus an $\mathscr{O}_{\mathscr{X}}$-module is coherent if and only if it is finitely presented.

Definition 6.27. An affinoid $\mathscr{X}$ is said to be a *Stein affinoid* if one of the following equivalent conditions is satisfied:

(1) $\mathrm{H}^1(\mathscr{X}, \mathscr{F}) = 0$ for any coherent sheaf $\mathscr{F}$;
(2) $\mathrm{H}^q(\mathscr{X}, \mathscr{F}) = 0$ for $q \geq 1$ and for any coherent sheaf $\mathscr{F}$;

(3) There exists a formal model X of $\mathscr{X}$ such that X is affine $X = \mathrm{Spf}\, A$ and that $\mathrm{Spec}\, A \setminus V(I)$ is an affine scheme, where I is an ideal of definition of A.

(4) There exists a distinguished formal model X of $\mathscr{X}$ such that X is affine $X = \mathrm{Spf}\, A$ and that $\mathrm{Spec}\, A \setminus V(I)$ is an affine scheme, where I is an ideal of definition of A.

The equivalence of the above conditions follows from the comparison theorem for affinoids and GFGA existence theorem. It can be shown that, for any rigid space $\mathscr{X}$, any admissible covering of $\mathscr{X}$ by affinoids can be refined by an admissible covering consisting of Stein affinoids. Combined with this fact, the next theorem shows that one can compute cohomology of coherent sheaves by means of Čech calculation using admissible covering by Stein affinoids.

Theorem 6.28 (Theorem A and Theorem B). *Let $\mathscr{X}$ be a Stein affinoid, and $\mathscr{F}$ a coherent sheaf on $\mathscr{X}$.*

(1) *If $X = \mathrm{Spf}\, A$ is a distinguished formal model of $\mathscr{X}$ such that $\mathrm{Spec}\, A \setminus V(I)$ (where I is an ideal of definition of A) is affine, then there exists a finitely presented A-module M such that*

$$\mathrm{H}^0(\mathscr{X}, \mathscr{F}) = \varinjlim_{n \geq 0} \mathrm{Hom}_A(I^n, M).$$

(2) *For $q \geq 1$, we have $\mathrm{H}^q(\mathscr{X}, \mathscr{F}) = 0$.*

Finally, we mention the finiteness theorem for proper morphisms.

Theorem 6.29 (Finiteness theorem for proper morphisms of rigid spaces). *Let $\varphi\colon \mathscr{X} \to \mathscr{Y}$ be a proper morphism between quasi-compact*[24] *rigid spaces. Then the functor $\mathrm{R}\varphi_*$ maps $\mathbf{D}^*_{\mathrm{coh}}(\mathscr{X})$ to $\mathbf{D}^*_{\mathrm{coh}}(\mathscr{Y})$ for $* = \emptyset, +, -, \mathrm{b}$.*

Here, for a rigid space $\mathscr{X}$, $\mathbf{D}^*_{\mathrm{coh}}(\mathscr{X})$ denotes the full subcategory of the derived category of the category of $\mathscr{O}_{\mathscr{X}}$-modules consisting of objects that have only coherent cohomologies.

Most of material presented in this part will be written in detail in the book [21] in preparation by the authors.

Part III. Applications

[24] A rigid space $\mathscr{X}$ is said to be quasi-compact if the topological space $\langle \mathscr{X} \rangle$ is quasi-compact; this is equivalent to the small admissible topos $\mathscr{X}^{\sim}_{\mathrm{ad}}$ being quasi-compact.

We expect that rigid geometry, especially that of our approach explained in the previous part, allows diverse applications. The applications within our scope at this moment include, at least, the following things:

- arithmetic geometry of Shimura varieties: p-adic period map and local models, p-adic automorphic representations, etc.,
- cohomology theory of algebraic varieties: ℓ-adic Lefschetz trace formulas, p-adic cohomology theory, etc.

In this final part, we discuss the applications of these things. In §7 we discuss arithmetic compactification of moduli of elliptic curves, and in §8, Lefschetz trace formulas.

Although we are not going to treat in this paper, one might moreover expect, in addition, the following applications:

- mirror symmetry (construction of mirror partner); cf. [28],
- p-adic Hodge theory (via theory of almost étale extensions); cf. [22],
- derived category equivalence,
- non-archimedean uniformization.

As for the last, we remark that, by means of the visualization, one can understand the known uniformization (e.g. [34], [35]) entirely as *topological uniformization*, that is, the uniformization by taking the universal covering. This point also streamlines the theory of orbifold uniformizations of rigid analytic curves developed in [25] (see also [10]).

§7. Application to compactification of moduli

In this section we discuss compactification of moduli spaces. We want to show that rigid geometry is useful in the analysis of moduli object near the boundary, and thus can be applied to the construction of the compactification. As the method that we are going to take is abstract enough, it affords the construction of the compactifications not only over fields, but over $\mathbb{Z}$, that is, *arithmetic compactifications*.

Here, at first, we would like to remind the reader of the fact that, in the classical theory of toroidal compactifications, complex analytic methods play an important role. The important point here is that the notion of rigid spaces is much broader than that of schemes, and rigid spaces are much flexible than schemes. In fact, there are several merits of using non-scheme theoretical geometric objects, such as rigid spaces, in application to the theory of moduli; among them are:

- topological feature: admissible topology of rigid spaces is finer than Zariski topology;

- it allows, in general, "construction by infinite repetition;" in this context, non-coherent objects play an essential role.

A typical example of the methods in the second point is the theory of p-adic uniformization, which provides, as we have already seen in §1.2, numerous nice techniques and viewpoints already in the classical rigid geometry by Tate.

First we fix some notations that we are going to use frequently in the sequel. We often consider pairs of the form $(\overline{X}, D)$, where $\overline{X} = \operatorname{Spec} A$ is an affine scheme, and D is a closed subscheme defined by a finitely generated ideal I of A. Let $\widehat{A}$ be the I-adic completion of A. We write $\overline{X}_{/D} = \operatorname{Spec} \widehat{A}$. It forms another pair $(\overline{X}_{/D}, D)$ of the form as above together with the closed subscheme defined by the ideal $I\widehat{A}$, which is, by a slight abuse of notation, again denoted by D.

In practice, the affine scheme $\overline{X}$ in the sequel appears as a "partial compactification" of a scheme X such that $\partial X = \overline{X} \setminus X = D$. In this situation, the complement $(\overline{X}_{/D}) \setminus D$ $(\cong \overline{X}_{/D} \times_{\overline{X}} X)$ is denoted by $X_{/D}$.

For a pair $(\overline{X}, D)$ as above we denote by $\widehat{\overline{X}}|_D$ the formal completion of $\overline{X}$ along D, i.e., $\widehat{\overline{X}}|_D = \operatorname{Spf} \widehat{A}$. The canonical morphism $\gamma_{\overline{X}} \colon \widehat{\overline{X}}|_D \to \overline{X}$ is factorized into the composite

$$\widehat{\overline{X}}|_D \xrightarrow{\beta_{\overline{X}}} \overline{X}_{/D} \xrightarrow{\alpha_{\overline{X}}} \overline{X}.$$

Notice that the formal completion of $(\overline{X}_{/D}, D)$ is the same as that of $(\overline{X}, D)$.

7.1. Analysis near cusps

In this section, we discuss the arithmetic compactification of the moduli space of elliptic curves over $\mathbb{Z}$, which is considered as one of the simplest but non-trivial examples, and then, later, indicate more general situation of Shimura varieties of PEL-type.

Let $\mathscr{M}$ be the moduli stack of elliptic curves over $\mathbb{Z}$. It is the algebraic stack characterized by the following condition: for a scheme S the category of Cartesian sections $\mathscr{M}(S)$ over S forms the groupoid consisting of elliptic curves over S, where morphisms are isomorphisms of the elliptic curves. We view $\mathscr{M}$ as a Deligne-Mumford stack, and denote by $f^{\mathrm{univ}} \colon \mathscr{E}^{\mathrm{univ}} \to \mathscr{M}$ the universal elliptic curve. We want to compactify the stack $\mathscr{M}$. To this end, first we are to analyze points near the cusps.

In order to do this, the construction of the Tate curves as in Example 3.3 provides a good picture. To apply it to our situation, we need to recast the construction in the universal form. Let W be the moduli space of maps of the form

$$u\colon \mathbb{Z} \longrightarrow \mathbb{G}_m$$

over $\mathbb{Z}$; that is, an affine scheme identified with $\operatorname{Spec}\mathbb{Z}[q,q^{-1}]$. Notice that the map as above is determined by its value at $u = 1$, and thus, the identification of W with $\operatorname{Spec}\mathbb{Z}[q,q^{-1}]$ is given by the universal map $u^{\mathrm{univ}}\colon \mathbb{Z} \to \mathbb{G}_m$ that maps 1 to q. We choose a torus embedding $W \hookrightarrow \overline{W} = \operatorname{Spec}\mathbb{Z}[q]$, which is seen as a partial compactification, and denote the infinity $\overline{W} \setminus W$ by D. The closed subscheme D is defined by the equation $q = 0$. We consider the pair $(\overline{W}, D)$.

By Tate construction (or the generalization due to Mumford), we have a semi-abelian scheme $\mathscr{A}$ over $\overline{W}_{/D}$, which has the following properties:

(1) the restriction $\mathscr{A}_{W_{/D}}$ of $\mathscr{A}$ to $W_{/D}$ is an elliptic curve, and the restriction to D is isomorphic to $\mathbb{G}_m$;
(2) $(\mathscr{A}_{W_{/D}})^{\mathrm{an}}$ is canonically isomorphic to the quotient $(\mathbb{G}_m)^{\mathrm{an}}$ by the subgroup generated by q as a rigid space over $\mathscr{W} = (\widehat{\overline{W}|_D})^{\mathrm{rig}}$;
(3) the construction is functorial in the sense as follows: for any complete valuation ring V of height 1 and an adic homomorphism $\Gamma(\overline{W}_{/D}, \mathscr{O}_{\overline{W}_{/D}}) = \mathbb{Z}[[q]] \to V$ that maps q to an element in $\mathfrak{m}_V$ (which we again denote by q), the base change of $\mathscr{A}$ to V corresponds to the Tate curve $(\mathbb{G}_{m,K})^{\mathrm{an}}/q^{\mathbb{Z}}$ as in Example 2.15, where K is the field of fractions of V.

Moreover the following property is known for the Tate construction:

Proposition 7.1 (Uniformization theorem, converse to Tate construction; cf. [16, Chap. II, §4]). *Assume that* $(S, D) = (\operatorname{Spec} V, V(I))$ *is a pair of affine schemes, where V is a Noetherian normal ring that is complete with respect to the I-adic topology. Let A be a semi-abelian scheme over S that satisfies following conditions:*

(1) *the relative dimension is* 1;
(2) *the restriction of A to D is a split torus;*
(3) *the restriction of A to $S \setminus D$ is an elliptic curve.*

Then there exists a morphism $g\colon S \to \overline{W}_{/D}$ such that A is isomorphic to the pullback $g^*\mathscr{A}$. Moreover, the morphism g is unique up to isomorphisms.

The proposition says that $\mathscr{A}$ is seen as the *universal Tate curve*, and $\overline{W}_{/D}$ is the classifying space for elliptic curves with split multiplicative reductions over complete base schemes. Moreover one can drop the assumption "normal" of S when S is of dimension 1. The uniformizaiton theorem, and the extension to the general 1-dimensional base schemes due to Raynaud, which we often abbreviate to "Raynaud-Tate theory", becomes very important later.

Now let us return to the moduli stack $\mathscr{M}$. We have a map

$$\epsilon\colon W_{/D} \longrightarrow \mathscr{M}$$

defined by the elliptic curve $\mathscr{A}_{W_{/D}}$ over $W_{/D}$. This map sits in the following 2-commutative diagram in a suitable 2-category of spaces:

$$\begin{array}{ccccc}
\overline{W}_{/D} & \hookleftarrow & W_{/D} & \longleftarrow & \mathscr{W} \\
\big\downarrow & & \big\downarrow{\scriptstyle \epsilon} & \swarrow & \\
\operatorname{Spec}\mathbb{Z} & \longleftarrow & \mathscr{M}, & &
\end{array}$$

The rigid space $\mathscr{W} = (\widehat{\overline{W}}|_D)^{\mathrm{rig}}$ is considered to be the family of "punctured unit disks" over $\mathbb{Z}$, or "(deleted) tubular neighborhood" of D inside $\overline{W}$. The desired compactification $\overline{\mathscr{M}}$ is obtained by patching the stack $\mathscr{M}$ and the scheme $\overline{W}_{/D}$ along the rigid space $\mathscr{W}$. This will be made more precise in the next two sections.

7.2. Arithmetic compactification

The following assertion provides the model case of the arithmetic compactifications in general.

Proposition 7.2. *There exists a proper smooth Deligne-Mumford stack $\overline{\mathscr{M}}$ over $\mathbb{Z}$ that contains $\mathscr{M}$ as an open substack enjoying the following properties*:

(1) *there exists a semi-abelian scheme $\overline{f}\colon \overline{\mathscr{E}}^{\mathrm{univ}} \to \overline{\mathscr{M}}$ that extends $\mathscr{E}^{\mathrm{univ}}$*;

(2) *the morphism $\epsilon\colon W_{/D} \to \mathscr{M}$ extends to $\overline{\epsilon}\colon \overline{W}_{/D} \to \overline{\mathscr{M}}$ in such a way that $\overline{\epsilon}^*\overline{\mathscr{E}}^{\mathrm{univ}} = \mathscr{A}$ holds*;

(3) *moreover, the morphism $\overline{\epsilon}$ induces a formally étale surjective morphism on passage to the formal completions.*

In fact, the rigid space $\mathscr{W} = (\widehat{\overline{W}_{/D}}|_D)^{\mathrm{rig}} = (\widehat{\overline{W}}|_D)^{\mathrm{rig}}$ is *almost* isomorphic to $(\widehat{\overline{\mathscr{M}}}|_{\partial\mathscr{M}})^{\mathrm{rig}}$; it gives an isomorphism when we introduce level structures to make the moduli problem fine. Thus the compatification $\overline{\mathscr{M}}$ in question should be constructed as the patching of $\mathscr{M}$ and $\overline{W}_{/D}$ along $\mathscr{W} = (\widehat{\overline{W}_{/D}}|_D)^{\mathrm{rig}}$. This gives the strategy for the construction that is quite similar to the complex analytic case. Notice that, to carry out this strategy, the framework of general rigid spaces (introduced in §6.5) is necessary.

7.3. Construction

The construction of $\overline{\mathscr{M}}$ takes three steps. The method exhibited here follows [20]. It is influenced by M. Rapoport work on Hilbert-Blumenthal varieties [36] and G. Faltings work on Siegel modular varieties [15].

First step (Algebraization). First we are to algebraize the family $\mathscr{A}$ over $\overline{W}_{/D}$ to a semi-abelian scheme over an affine scheme of finite type over $\mathbb{Z}$. For any $n \geq 1$, by Artin's approximation theorem, we have an affine smooth scheme $\overline{V}_n$ over $\mathbb{Z}$ and a closed subscheme $D_n \subset \overline{V}_n$ such that the formal completion of $\overline{V}_n$ along D_n is identified with $\widehat{\overline{W}|_D}$ (we fix this identification). Moreover, there is a semi-abelian scheme A_n of relative dimension 1 over $\overline{V}_n$ such that A_n is an elliptic curve over $V_n = \overline{V}_n \setminus D_n$ and is a split torus over D_n.

The family A_n is "very near" to $\mathscr{A}$ in the following sense: when we regard A_n as a quotient of $\mathbb{G}_m$ by ${q_n}^{\mathbb{Z}}$ over $\overline{V}_{n/D_n}$ for some $q_n \in \Gamma(\widehat{\overline{W}|_D}, \mathscr{O}_{\widehat{\overline{W}|_D}})$ by the uniformization theorem (Proposition 7.1), $q_n \equiv q \mod q^n$ holds. To achieve the last condition, one must approximate the semi-abelian scheme with a line bundle and sections, i.e. with theta functions.

When A_n is very near to $\mathscr{A}$ in the sense as above and $n \geq 2$, the morphism

$$\delta_n \colon \overline{V}_{n/D_n} \longrightarrow \overline{W}_{/D}, \qquad q \mapsto q_n$$

by the universality of $\overline{W}_{/D}$ (again by Proposition 7.1) is an isomorphism[25], and that the pull back $\delta_n^* \mathscr{A}$ is isomorphic to A_n.

One sets $\overline{V} = \overline{V}_n$, $D = D_n$, and $A_{\overline{V}} = A_n$ for some $n \geq 2$, and $V = \overline{V} \setminus D$.

Second step (Openness of versality). We show that the classifying morphism $V = \overline{V} \setminus D \to \mathscr{M}$ defined by $A_{\overline{V}}$ is étale by shrinking V around D if necessary. For this, it suffices to show that $V_{/D} \to \mathscr{M}$ is formally smooth at any closed point of $V_{/D}$ (note that the residue field at any closed point is a complete discrete valuation field). To show this, one uses the infinitesimal criterion of formal smoothness, and reduces to show the following assertion.

Proposition 7.3. *Let R_0 be a complete discrete valuation ring, π a uniformizer, and R a finite local algebra that is a thickening of R_0.*

---

[25]The identification by δ_n can be different from the one which is already chosen

Assume that we are given an elliptic curve E over $R[\frac{1}{\pi}]$ such that the restriction E_{R_0} to R_0 is a Tate curve over R_0. Then, by replacing R by a finite modification (= finite map that induces isomorphism outside the ideal (π)) if necessary, E is also a Tate curve.

This is a direct consequence of the uniformization theorem (Raynaud-Tate theory over 1-dimensional complete rings). Roughly speaking, the point is to show the deformations of an elliptic curve with split multiplicative reduction are the same as the deformations of corresponding 1-motives obtained by the Raynaud-Tate theory. (A related work for Mumford curves is in [11, §9].)

Third step (Patching). We construct $\overline{\mathscr{M}}$ by patching $\overline{V}$ (obtained in Step 2) and $\mathscr{M}$ along $(\overline{V}|_D)^{\mathrm{rig}}$. This is easy by using the openness of versality (Proposition 7.3). Since $\mathscr{M}$ is a Deligne-Mumford stack, there are an étale surjective morphism $P \to \mathscr{M}$ from a smooth affine scheme P, and a relation $R \to P \times_{\mathbb{Z}} P$ that defines $\mathscr{M}$ as a stack. Note that we have the pull back A_P to P of the universal elliptic curve. Together with $A_{\overline{V}}$ over $\overline{V}$, we have a semi-abelian scheme $A_{P \coprod \overline{V}}$ on $P \coprod \overline{V}$.

We take the normalization $\widetilde{R}$ of $(P \coprod \overline{V}) \times_{\mathbb{Z}} (P \coprod \overline{V})$ in R, and show that $\widetilde{R}$ defines an étale relation on $P \coprod \overline{V}$ and defines a Deligne-Mumford stack $\overline{\mathscr{M}}$. Semi-abelian scheme $A_{P \coprod \overline{V}}$ also descends to a semi-abelian scheme $\overline{\mathscr{E}}^{\mathrm{univ}} \to \overline{\mathscr{M}}$. The point here is that one can control the situation using the semi-abelian scheme on $P \coprod \overline{V}$ and the uniformization theorem. By the construction, the properties (1)–(3) of Proposition 7.2 follow.

The properness of $\overline{\mathscr{M}}$ follows from the valuative criterion, using Grothendieck's semi-stable reduction theorem for abelian varieties. Then we finish the construction.

7.4. General case: Shimura varieties of PEL-type

The method of the arithmetic compactification of the moduli of elliptic curves generalizes to more general Shimura varieties.

First, we need a good model of Shimura varieties over $\mathbb{Z}$. For this purpose, we must restrict ourselves to the so-called PEL-case, which can be seen as a moduli of abelian varieties with some rigidification structures, namely a rigidification of a polarization, the endomorphism ring, the Hodge filtration, and the Betti realization ([39], [12]). For the general definition of Shimura varieties we refer to [12] and [13].

Let L be a semi-simple algebra over $\mathbb{Q}$ with a positive involution $*$, V a finite dimensional $\mathbb{Q}$-vector space that is a faithful L-module with a non-degenerate $\mathbb{Q}$-valued skew symmetric form φ that satisfies the equality

$$\varphi(\ell x, y) = \varphi(x, \ell^* y), \qquad \text{for } x, y \in V,\ \ell \in L.$$

The reductive group G over $\mathbb{Q}$ is the group of L-linear symplectic similitudes of V.

Let X be the set of all homomorphisms $h\colon \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m \to G_{\mathbb{R}}$ such that the $\mathbb{R}$-Hodge structure defined by h on $V_{\mathbb{R}}$ has the type $\{(-1,0),\ (0,-1)\}$ and polarized by φ. The involution of L is required to be positive for this structure.

Then X carries a natural complex structure. Each connected component of X is a hermitian symmetric domain. To simplify the situation, we assume that all $\mathbb{R}$-simple factors of the derived group G^{der} are of type A or C, and hence G^{der} is simply connected. The corresponding (non-connected) Shimura variety for (G,X) over $\mathbb{C}$ is defined by

$$Sh_K(G,X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K,$$

where K is a compact open subgroup of $G(\mathbb{A}_f)$. The space $Sh_K(G,X)$ is a quasi-projective variety defined over an explicit number field E called the *reflex field* of (G,X).

To get an arithmetic moduli, except for several successful cases, only the case of good reduction has been considered in general. Zink [47], Langlands-Rappoport [30] have defined a *smooth* arithmetic model $Sh_K(G,X)$ over $\mathscr{S}_K$ of $Sh_K(G,X)(\mathbb{C})$ as the solution of a moduli problem involving abelian schemes (under the restriction on G). Here $\mathscr{O}_E$ is the ring of integers of E, and $\mathscr{S}_K \subset \operatorname{Spec}\mathscr{O}_E$ is an open set explicitly described by K (for general K, we regard $Sh_K(G,X)$ as a Deligne-Mumford stack).

By definition, there exists the universal abelian scheme

$$f^{\mathrm{univ}}\colon \mathscr{A}^{\mathrm{univ}} \to Sh_K(G,X)$$

that gives Shimura's family over $\mathbb{C}$ in [39].

Proposition 7.4. *Choose an admissible cone decomposition Σ that is compatible with K. Then the toroidal compactification $Sh_K(G,X)(\Sigma)$ of $Sh_K(G,X)$ over $\mathscr{S}_K$ for this cone decomposition that satisfies the following properties exists*:

1. (1) *$Sh_K(G,X)(\Sigma)$ is a proper Deligne-Mumford stack over $\mathscr{S}_K$ whose local structure near the boundary is described by the toroidal embeddings that correspond to cones in Σ;*
2. (2) *the geometric fiber over* $\operatorname{Spec}\mathbb{C}$ *is the one constructed in* [1];

(3) *the universal abelian scheme* $f^{\mathrm{univ}}\colon \mathscr{A}^{\mathrm{univ}} \to Sh_K(G,X)$ *extends uniquely to a semiabelian scheme*

$$\overline{f}^{\mathrm{univ}}\colon \overline{\mathscr{A}}^{\mathrm{univ}} \longrightarrow Sh_K(G,X)(\Sigma).$$

Note that our compactification $Sh_K(G,X)(\Sigma)$ is, a priori, an algebraic stack or, for K small enough, an algebraic space. Here we suggest how the construction will be done along the line described in subsection 7.3.

The role of W in §7.1 is played by (the arithmetic model of) the mixed Shimura varieties $Sh(P,X_P)$ associated to $\mathbb{Q}$-maximal parabolic subgroup P of G [8]. These mixed Shimura varieties are seen as a moduli space of 1-motives with PEL-structure, and admit a fibration $Sh(P,X_P) \to B_P$ by a split torus T_P. Our choice of the cone decomposition determines a torus embedding $T_P \hookrightarrow T_{P,\sigma}$ for a cone σ, and a partial compactification $Sh(P,X_P)_\sigma$ of $Sh(P,X_P)$ is obtained by the contracted product $Sh(P,X_P) \wedge^{T_P} T_{P,\sigma}$ (that is, the fiber bundle with the fibers $T_{P,\sigma}$ associated to the torus bundle $Sh(P,X_P) \to B_P$).

The partial compactification $Sh(P,X_P)_\sigma$ plays the role of $\overline{W}$ in §7.1. Then one uses the Mumford construction of semi-abelian schemes, which is a generalization of Tate construction to higher dimensional abelian schemes, to get a semi-abelian scheme $\mathscr{A}_\sigma$ from the universal 1-motive on $Sh(P,X_P)$ after completion along the closed T_P-orbit D_σ. Then we algebraize $(\mathscr{A}_\sigma, Sh(P,X_P)_{\sigma/D_\sigma})$ by using Artin's approximation theorem as in §7.3, Step 1.

The difficulty to construct $Sh_K(G,X)(\Sigma)$, compared to the elliptic curve case, lies in the fact that the openness of versality is much harder to show. For example, the types of degenerations of abelian varieties of fixed PEL-type is much more complicated, so we must somehow control the various types of degenerations and different partial compactifications at the same time to show the openness of versality.

To check the versality, we use rigid geometry and the Raynaud-Tate theory for semiabelian schemes over one-dimensional complete rings (the argument is similar to that in §7.3, Step 2 and Step 3, but more complicated), with a closer analysis of degenerations using the uniformization theory in [16].

Recall that the use of rigid geometry in compactification problem goes back to Rapoport's fundamental and important work on Hilbert-Blumental varieties [36]. The use of Artin's approximation theorem goes back to Faltings work, and discussed in [16] for Siegel modular varieties. For Siegel modular varieties, there is also a method of Chai [9]. He constructs the arithmetic toroidal compactification (corresponding to projective cone decomposition) by blowing up the minimal compactification, using the theory of algebraic theta functions[26].

Remark 7.5. In [16], Kodaira-Spencer mappings are used to verify the openness of versality, so one needs to assume the smoothness of arithmetic models in principle. The method here has the advantage that it is singularity free: if a good theory of canonical arithmetic models of Shimura varieties over the ring of integers were available, our method in §7.3 also gives the arithmetic compactifications including the bad reduction cases, as long as arithmetic models of mixed Shimura varieties corresponding to parabolics are normal.

7.5. Applications of arithmetic compactifications

The existence of arithmetic compactification has important consequences on modular forms. Fix an admissible cone decomposition Σ and consider the arithmetic toroidal compactification. The line bundle

$$\omega = \det(\mathrm{Lie}(\overline{\mathscr{A}}^{\mathrm{univ}}/Sh_K(G,X)(\Sigma))^{\vee})$$

on $Sh_K(G,X)(\Sigma)$ is semi-ample by a theorem of Moret-Bailly [33]. The space of sections

$$M_k = \Gamma(Sh_K(G,K)(\Sigma), \omega^{\otimes k})$$

is independent of Σ and regarded as a space of geometric modular forms of weight k. By the properness of $Sh_K(G,X)(\Sigma)$, M_k is finitely generated $\Gamma(C_K, \mathscr{O}_{C_K})$-module, and the graded ring $\bigoplus_{k\geq 1} M_k$ is finitely generated over $\Gamma(X_K, \mathscr{O}_{C_K})$ by Moret-Bailly's theorem. This is already an important finiteness statement on geometric modular forms, which is hard to prove by other methods. The geometric modular forms in our sense is identified with holomorphic modular forms with integral coefficients (q-expansion principle). Summing up, we have the following statement.

Proposition 7.6 (cf. [16, Chap. V, §1] in the Siegel modular case). *The following properties hold if bad primes are invertible in the coefficients*:

(a) *Koecher principle,*
(b) *q-expansion principle,*
(c) *the finiteness theorem for the space of geometric modular forms of given weight (including the vector valued case).*

---

[26]This method works over $\mathbb{Z}[\frac{1}{2}]$. One must exclude prime 2 since it is a bad prime for the theory of algebraic theta functions.

One can also show that $\overline{Sh_K(G,X)}_{\min} = \operatorname{Proj} \bigoplus_{k\geq 1} M_k$ gives another compactification of $Sh_K(G,X)$, which is in fact the arithmetic minimal (= Satake, Baily-Borel) compactification:

Proposition 7.7 (cf. [16, Chap. V, §2] in the Siegel modular case). *The compactification $\overline{Sh_K(G,X)}_{\min}$ of $Sh_K(G,X)$ has the following property: for a Noetherian normal scheme S, an open dense subscheme U, and a morphism $f\colon U \to Sh_K(G,X)$ such that the pull-back of the universal abelian scheme $f^*(\mathscr{A}^{\text{univ}})$ admits a semi-stable reduction to S, f has a unique extension $\overline{f}\colon S \to \overline{Sh_K(G,X)}_{\min}$.*

These integrality results have very important consequence in number theory. For example, one can use q-expansion principle to produce congruence between two modular forms. Deligne and Ribet [14] constructed p-adic L-functions for finite order characters over a totally real field by using Hilbert-Blumenthal varieties, and recently Urban and Skinner use similar method for unitary Shimura varieties in their study of Iwasawa main conjecture of elliptic curves over $\mathbb{Q}$.

§8. Rigid spaces and Frobenius

The main subject to be dealt with in this section, as the second application, is an application of rigid geometry to theory of schemes. The category of rigid spaces is, as pointed out before, much broader than that of schemes. Hence, what we like to show is, so to speak, one of the "non-scheme-theoretic" methods for treating schemes. In fact, such methods that derail from scheme theory often reveal hidden and important features in scheme theory, which would be quite invisible only from the scheme-theoretic point of view.

In this section, we particularly focus on Frobenius. To do this, we first show the general technique to bridge between scheme theory and rigid geometry in the next subsection.

8.1. From schemes to rigid spaces; constant deformation technique

This is the general technique that is important for applying rigid geometry to geometry of schemes. The general picture is as follows (Figure 10).

First, start from a variety X over a field k. From X we are going to construct canonically a rigid space. Consider the ring of formal power series $k[[t]]$ endowed with the t-adic topology, and put $\widehat{X}_{k[[t]]} = X \times_k \operatorname{Spf} k[[t]]$ (constant deformation). Then one takes its associated rigid space $(\widehat{X}_{k[[t]]})^{\text{rig}}$ over $(\operatorname{Spf} k[[t]])^{\text{rig}}$.

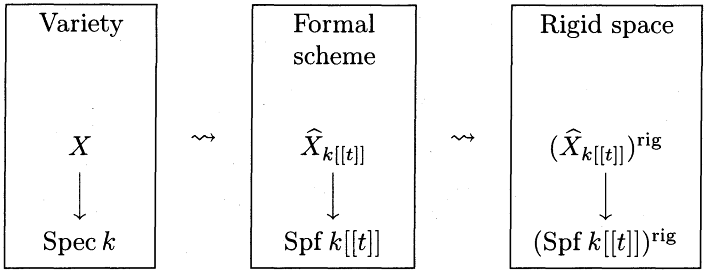

Fig. 10. Constant deformation technique

In spite of its entirely trivial looking, this construction opens the way for several effective applications of rigid geometry to algebraic geometry.

8.2. Frobenius

Rigid geometry reveals a new feature of Frobenius morphisms in positive characteristic. This feature will be given in Claim 8.1.

Consider the following situation:

- S: an $\mathbb{F}_q$-scheme,
- $\mathrm{Fr}_q\colon S \to S$: Frobenius over $\mathbb{F}_q$, that is, the q-th power map,
- $\mathscr{C}_S$: a category of geometric objects over S.

One of the properties of Frobenius morphisms that are already known to be very important in classical algebraic and arithmetic geometry is that, most of the time, the Frobenius induces a self-functor

$$\mathrm{Fr}_q^*\colon \mathscr{C}_S \longrightarrow \mathscr{C}_S.$$

In other words, one has the "dynamical system" with the "phase space" $\mathscr{C}_S$ acted on by the "self-similarity map" Fr_q^*. As usual in the theory of dynamical system, one is particularly interested in the "Fr_q^*-fixed point", that is, the Fr_q-structure

$$\mathrm{Fr}_q^* A \cong A.$$

Once one has such a structure, one is interested in the following question.

Question: *what happens* near *the "Fr_q^*-fixed point"?*

Constant deformation and Frobenius. In this context, the constant deformation technique proves to be useful. First observe that, in complex situation with $\mathbb{A}^1_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[X]$, the selfmap

$$\mathbb{A}^1_{\mathbb{C}} \longrightarrow \mathbb{A}^1_{\mathbb{C}}, \qquad X \mapsto X^q$$

is a contracting map near the origin for the analytic topology. The rigid geometric counterpart of this is the following:

$$\mathbb{D}^1 \longrightarrow \mathbb{D}^1, \qquad X \mapsto X^q$$

is contracting near 0 in the sense that $(\mathrm{Fr}_q(\mathbb{D}^1(r)) \subset \mathbb{D}^1(r^q)$.

More precisely, for an $\mathbb{F}_q$-variety X, consider the rigid space

$$\mathscr{X} = (\widehat{X}_{\mathbb{F}_q[[t]]})^{\mathrm{rig}}$$

obtained by constant deformation. Let $Y \subseteq X$ be an Fr_q-invariant subspace, and set

$$\mathscr{Y} = (\widehat{Y}_{\mathbb{F}_q[[t]]})^{\mathrm{rig}}.$$

We think of $\mathscr{X}$ as the phase space equipped with the dynamical system

$$\mathscr{X} \xrightarrow{\mathrm{Fr}_q} \mathscr{X} \xrightarrow{\mathrm{Fr}_q} \mathscr{X} \xrightarrow{\mathrm{Fr}_q} \cdots .$$

This can be more concretely done by means of the associated Zariski-Riemann space $\langle\mathscr{X}\rangle$; by this, we have a topological space (in the usual sense) as the phase space.

Claim 8.1. "The Frobenius mapping is contracting near Y," *i.e., Fr_q is contracting near $\langle\mathscr{Y}\rangle$ in $\langle\mathscr{X}\rangle$.*

The claim can be shown by the reasoning similar to that in the case of the unit disk as above. The property of Frobenius that the claim shows is so essential in general that it actually simplifies arguments in many situations. The Lefschetz trace formula, which we are to discuss in the next subsection, is one of them.

8.3. Trace formula in characteristic p

The "dynamical system" approach to Frobenius as in §8.2 has already appeared and applied in the study of Lefschetz trace formula in characteristic p by the first-named author (solution of Deligne's conjecture [19]). Let us briefly outline the argument therein.

Deligne's conjecture. Let X be an algebraic variety over a field k. Consider a correspondence

$$a\colon Y \longrightarrow X \times_k X$$

such that $a_1 = \mathrm{pr}_1 \circ a$ is proper and that $a_2 = \mathrm{pr}_2 \circ a$ is quasi-finite. Let K be a $\overline{\mathbb{Q}}_\ell$-complex (where $\frac{1}{\ell} \in k$) with a cohomological correspondence compatible with a. In this situation the *Lefschetz number*, an element of $\overline{\mathbb{Q}}_\ell$, is defined by

$$\mathrm{Lef}(a, \mathrm{R}\Gamma_{\mathrm{c}}(X, K)) = \mathrm{Trace}(a^*, \mathrm{R}\Gamma_{\mathrm{c}}(X, K)).$$

Theorem 8.2 (Deligne's conjecture; [19]). *Let* $k = \overline{\mathbb{F}}_q$. *If the above data admit* Fr_q*-structure, then there exists* $N \in \mathbb{N}$ *such that the following conditions are satisfied:*

(1) $\dim \mathrm{Fix}(\mathrm{Fr}_q^n \circ a) = 0$ *for* $q^n > N$;
(2) *for* $q^n > N$,

$$\mathrm{Lef}(\mathrm{Fr}_q^n \circ a, \mathrm{R}\Gamma_{\mathrm{c}}(X, K)) = \sum_{D \in \mathrm{Fix}(\mathrm{Fr}_q^n \circ a)} \mathrm{naive.loc}_D(\mathrm{Fr}_q^n \circ a, K).$$

Here $\mathrm{naive.loc}_D(a, K)$ *vanishes if* $K|_{a_2(D)} = 0$.

The proof is given by establishing the trace formula for certain rigid analytic correspondences; note that this argument is not completely scheme-theoretical.

Another way of proof was given by Shpiz and Pink in their work around 1990, in which they assume that X is smooth and K is a smooth sheaf, that there exists a good compactification, and that K is tame. Recently, T. Saito and Y. Varshavsky [43] independently gave scheme-theoretic proofs.

Remark 8.3. In [19] it was assumed that X and Y are schemes. But, by Equivalence Theorem (Theorem 6.12) and Nagata's Embedding Theorem for algebraic spaces, we may weaken the assumption to that X and Y are separated algebraic spaces of finite type over k. This generalization actually eliminates the use of an argument in [29] to show that the moduli space of Shtuka is a scheme.[27]

Applications of Deligne's conjecture. Finally, let us list some of the applications of Deligne's conjecture, which provides a very strong counting argument in arithmetic and many other areas in mathematics:

- Non-abelian class field theory (Shtuka moduli (L. Lafforgue [29]), Shimura varieties (Harris-Taylor...)),
- Representation theory of Chevalley groups (Digne-Rouquier, ...),
- Model theory (Hrushovski-Macintyre).

References

[1] A. Ash, D. Mumford, M. Rapoport and Y. Tai, Smooth compactifications of locally symmetric varieties, Math. Sci. Press, 1975.

[27] In fact, here, one need more general version of the trace formula, which has been proven [19].

[2] V. G. Berkovich, Spectral theory and analytic geometry over non-Archimedean fields, Mathematical Surveys and Monographs, **33**, American Mathematical Society, Providence, RI, 1990.

[3] V. G. Berkovich, Étale cohomology for non-Archimedean analytic spaces, Inst. Hautes Études Sci. Publ. Math., **78** (1993), 5–161.

[4] P. Berthelot, Géométrie rigide et cohomologie des variétés algébriques de caractéristique p, Introductions aux cohomologies p-adiques (Luminy, 1984), Mém. Soc. Math. France (N.S.), **23** (1986), 7–32.

[5] S. Bosch, U. Güntzer and R. Remmert, Non-Archimedean analysis. A systematic approach to rigid analytic geometry, Grundlehren der Mathematischen Wissenschaften, **261**, Springer-Verlag, Berlin, 1984.

[6] S. Bosch and W. Lütkebohmert, Formal and rigid geometry. II. Flattening techniques, Math. Ann., **296** (1993), 403–429.

[7] N. Bourbaki, Elements of mathematics. Commutative algebra, Translated from the French. Hermann, Paris, Addison-Wesley Publishing Co., Reading, Mass., 1972.

[8] J. L. Brylinski, 1-motifs et formes automorphes, Journées automorphes 15, Publications Mathématique de l'université Paris VII (1983).

[9] C. L. Chai, Compactification of Siegel moduli schemes, Lecture Note Series, **107**, London Mathematical Society, Cambridge University Press, 1985.

[10] G. Cornelissen, F. Kato and A. Kontogeorgis, Discontinuous groups in positive characteristic and automorphisms of Mumford curves, Math. Ann., **320** (2001), 55–85.

[11] G. Cornelissen and F. Kato, Equivariant deformation of Mumford curves and of ordinary curves in positive characteristic, Duke Math. J., **116** (2003), 431–470.

[12] P. Deligne, Travaux de Shimura, Séminaire Bourbaki, 1970/1971, Exposé 389, Lecture Notes in Mathematics, **244**, Springer-Verlag, 1971.

[13] P. Deligne, Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques, Proc. Symp. Pure. Math., A. M. S., **33** (1979), pp. 247–290.

[14] P. Deligne and K. A. Ribet, Values of abelian L-functions at negative integers over totally real fields, Invent. Math., **59** (1980), 227–286.

[15] G. Faltings, Arithmetische Kompaktifizierung des Modulraums der abelschen Varietäten, Workshop Bonn 1984 (Bonn, 1984), Lecture Notes in Mathematics, **1111**, Springer-Verlag, Berlin, 1985, 321–383.

[16] G. Faltings and Ching-L. Chai, Degeneration of abelian varieties. With an appendix by David Mumford, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), **22**, Springer-Verlag, Berlin, 1990.

[17] J. Fresnel and M. van der Put, Rigid analytic geometry and its applications, Progress in Mathematics, **218**, Birkhäuser Boston, Inc., Boston, MA, 2004.

[18] K. Fujiwara, Theory of tubular neighborhood in étale topology, Duke Math. J., **80** (1995), 15–57.

[19] K. Fujiwara, Rigid geometry, Lefschetz-Verdier trace formula and Deligne's conjecture, Invent. Math., **127** (1997), 489–533.

[20] K. Fujiwara, Arithmetic compactifications of Shimura varieties, preprint.

[21] K. Fujiwara and F. Kato, book in preparation.

[22] O. Gabber and L. Ramero, Almost ring theory, Lecture Notes in Mathematics, **1800**, Springer-Verlag, Berlin, 2003.

[23] R. Huber, Continuous valuations, Math. Z., **212** (1993), 455–477.

[24] P. T. Johnstone, Stone spaces, Cambridge Studies in Advanced Mathematics, **3**, Cambridge University Press, Cambridge, 1982.

[25] F. Kato, Non-archimedean orbifolds covered by Mumford curves, J. Algebraic Geom., **14** (2005), 1–34.

[26] R. Kiehl, Der Endlichkeitssatz für eigentliche Abbildungen in der nichtarchimedischen Funktionentheorie, Invent. Math., **2** (1967), 191–214.

[27] R. Kiehl, Theorem A und Theorem B in der nichtarchimedischen Funktionentheorie, Invent. Math., **2** (1967), 256–273.

[28] M. Kontsevich and Y. Soibelman, Affine structures and non-archimedean analytic spaces, preprint.

[29] L. Lafforgue, Chtoucas de Drinfeld et correspondance de Langlands, Invent. Math., **147** (2002), 1–241.

[30] R. P. Langlands and M. Rapoport, Shimuravarietäten und Gerben, J. Reine Angew. Math., **378** (1987), 113–220.

[31] J. Lipman, Rational singularities, with applications to algebraic surfaces and unique factorization, Inst. Hautes Études Sci. Publ. Math., **36** (1969), 195–279.

[32] W. Lütkebohmert, Formal-algebraic and rigid-analytic geometry, Math. Ann., **286** (1990), 341–371.

[33] L. Moret-Bailly, Pinceaux de variétés abélienne, Astérisque, **129** (1985).

[34] A. Mumford, An analytic construction of degenerating curves over complete local rings, Compositio Math., **24** (1972), 129–174.

[35] G. A. Mustafin, Nonarchimedean uniformization, Math. USSR Sbornik, **34** (1978), 187–214.

[36] M. Rapoport, Compactifications de l'espace de modules de Hilbert-Blumenthal, Compositio Math., **36** (1978), 255–335.

[37] M. Raynaud, Géométrie analytique rigide d'après Tate, Kiehl,$\cdots$, Table Ronde d'Analyse non archimédienne (Paris, 1972), Bull. Soc. Math. France, Mem., **39–40**, Soc. Math. France, Paris, 1974, pp. 319–327.

[38] M. Raynaud and L. Gruson, Critères de platitude et de projectivité. Techniques de "platification" d'un module, Invent. Math., **13** (1971), 1–89.

[39] G. Shimura, Moduli and fibre systems of abelian varieties, Ann. of Math., **83** (1966), 294–338.

[40] J. Tate, Rigid analytic spaces, Invent. Math., **12** (1971), 257–289.

[41] J. Tate, A review of non-Archimedean elliptic functions, Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993), Ser. Number Theory, I, Internat. Press, Cambridge, MA, 1995, 162–184.

[42] M. van der Put and P. Schneider, Points and topologies in rigid geometry, Math. Ann., **302** (1995), 81–103.

[43] Y. Varshavsky, Lefschetz-Verdier trace formula and a generalization of a theorem of Fujiwara, preprint (2005).

[44] O. Zariski, The reduction of the singularities of an algebraic surface, Ann. of Math. (2), **40** (1939), 639–689.

[45] O. Zariski, The compactness of the Riemann manifold of an abstract field of algebraic functions, Bull. Amer. Math. Soc., **50** (1944), 683–691.

[46] O. Zariski and P. Samuel, Commutative algebra, Volume II. The University Series in Higher Mathematics. D. Van Nostrand Co., Inc., Princeton, N. J.-Toronto-London-New York, 1960.

[47] Th. Zink, Isogenieklassen von Punkten von Shimuramannigfaltigkeiten mit Werten in einem endlichen Körper, Math. Nacher., **112** (1983), 103–124.

[EGA] A. Grothendieck and J. Dieudonné, Éléments de géométrie algébrique, Inst. Hautes Études Sci. Publ. Math., **4**, **8**, **11**, **17**, **20**, **24**, **28**, **32** (1961-1967).

[EGA, $\mathrm{I}_{\mathrm{new}}$] A. Grothendieck and J. Dieudonné, Éléments de géométrie algébrique I, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, **166**, Springer-Verlag, Berlin, Heidelberg, New York, 1971.

[SGA4-2] Théorie des topos et cohomologie étale des schémas. Tome 2. Séminaire de Géométrie Algébrique du Bois-Marie 1963–1964 (SGA 4). Dirigé par M. Artin, A. Grothendieck et J. L. Verdier. Avec la collaboration de N. Bourbaki, P. Deligne et B. Saint-Donat, Lecture Notes in Mathematics, **270**, Springer-Verlag, Berlin-New York, 1972.

Kazuhiro Fujiwara
Graduate School of Mathematics
Nagoya University
Nagoya 464-8502
Japan

Fumiharu Kato
Department of Mathematics
Faculty of Science
Kyoto University
Kyoto 606-8502
Japan