**Rigid Cohomology**

Dating back to work of Berthelot, rigid cohomology appeared as a common generalization of Monsky–Washnitzer cohomology and crystalline cohomology. It is a $p$-adic Weil cohomology suitable for computing Zeta and $L$-functions for algebraic varieties on finite fields. Moreover, it is effective, in the sense that it gives algorithms to compute the number of rational points of such varieties.

This is the first book to give a complete treatment of the theory, from full discussion of all the basics to descriptions of the very latest developments. Results and proofs are included that are not available elsewhere, local computations are explained, and many worked examples are given. This accessible tract will be of interest to researchers working in arithmetic geometry, $p$-adic cohomology theory, and related cryptographic areas.

# Rigid Cohomology

BERNARD LE STUM
*Université de Rennes I, France*

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521875240



First published 2007

Printed in the United Kingdom at the University Press, Cambridge

*A catalog record for this publication is available from the British Library*

ISBN 978-0-521-87524-0 hardback

*À Pierre Berthelot.*

Une mathématique bleue,
Sur cette mer jamais étale
D'où me remonte peu à peu
Cette mémoire des étoiles

(LÉO FERRÉ)

# Contents

# Preface

In 2004, I was asked by Professor King Fai Lai to come to Peking University in order to give a course on rigid cohomology. We agreed on the last two weeks of January 2005. I want to thank here Professor Zhao Chunlai for the organization of my visit as well as Professor Zhou Jian and his wife for showing us the city. My family and I will always remember it.

While preparing this course, I realized that there was no introductory book on rigid cohomology. Actually, there was no available material in English and only an old document in French, *Cohomologie rigide et cohomologie rigide à support propre*, by Pierre Berthelot. A revised version of the first part of this document appeared as an official preprint in 1996 but the second part is not fully written yet and, therefore, not really available to the mathematical community. Fortunately, Berthelot was kind enough to answer my questions on this second part and point out some articles where I could find some more information.

Rigid cohomology was introduced by Berthelot as a $p$-adic analogue of $l$-adic cohomology for lisse sheaves, generalizing Monsky–Washnitzer theory as well as crystalline cohomology (up to torsion). Recently, it appeared that this theory may be used in order to derive new algorithms for cryptography. The first result in this direction is due to Kiran Kedlaya who has also done incredible work on the theoretical aspect of the theory.

I knew that it was impossible to cover the full story in twenty one-hour lectures. I decided to first introduce the theory from the cryptography point of view (Introduction), then describe the basics of the theory with complete proofs (heart of the the course), and conclude with an overview of the development of the theory in the last 20 years (Conclusion). In particular, the main part of this book is quite close to Berthelot's original document. I hope that this will be useful to the students who want to learn rigid cohomology and, eventually, improve on our results.

I insist on the fact that there is no original matter here and that almost everything is due to Pierre Berthelot, apart from the mistakes arising from my misunderstanding.

One can split the so-far short life of rigid cohomology into three periods: (1) foundations, (2) cohomology of varieties and (3) cohomology of $F$-isocrystals.

It is a wonderful idea of Berthelot's to generalize crystalline and Monsky–Washnitzer cohomology into one theory. The principle is to compactify the variety $X$, embed it into some smooth formal scheme $P$ and compute the limit de Rham cohomology of "strict" neighborhoods $V$ of $X$ in the generic fiber $P_K$ of $P$. The astonishing fact is that the result does not depend on the choices. Better: there exists a category of coefficients for this theory. This is simply the limit category of differential modules on strict neighborhoods $V$ that have the good idea to be "overconvergent". What makes this possible is a deep geometrical result, the strong fibration theorem. It tells us that even the geometry of $V$ is essentially independent of the choices.

Some time after Berthelot had laid the foundations, the theory got a kick thanks to Johan de Jong's alterations theorem. He made it possible to use rigid cohomology as a bridge between crystalline cohomology and Monsky–Washnitzer cohomology in order to show that the latter is finite dimensional. More generally, de Jong's theorem, which states that one can solve singularity in characteristic $p$ if one is ready to work with étale topology, can be used to show that rigid cohomology of varieties satisfies the formalism of Bloch–Ogus (finiteness, Poincaré duality, Künneth formula, cycle class, etc.).

The third period started with three almost simultaneous proofs of the so called conjecture of Crew. Even if he never stated this as an explicit conjecture, Crew raised the following question: is a differential module with a strong Frobenius structure on a Robba ring automatically quasi-unipotent? The first proof, due to Yves André is an application of representation theory using results of Christol and Mebkhout on "slopes". The second one, due to Zogman Mebkhout, is derived from his previous work. More important for us, the third one, due to Kiran Kedlaya, is a direct construction and can be generalized to higher dimension. From this, he can derive all the standard properties of rigid cohomology of overconvergent $F$-isocrystals, and in particular finite dimensionality.

In this course, we will focus on foundations. The main results of the second period will be mentioned in the Introduction where we will try to give a historical introduction based on Weil's conjectures and recent results in cryptography. In the Conclusion, we will try to evaluate the state of the art and, at the same time, review the main results of the third period.

## Acknowledgments

During the writing of this course, I had useful conversations with Gweltaz Chatel, Michel Gros, Ke-Zheng Li, David Lubicz, Laurent Moret-Bailly, Richard Crew and, of course, Pierre Berthelot. I want to thank them here. Thank you also to Marion Angibaud for helping with the drawings.

Also, this work was partially supported by the European Network Arithmetic Algebraic Geometry.

Of course, I thank PKU University for its invitation and hospitality.

Finally, I want to thank Roger Astley from Cambridge University Press who made it possible to turn this course into a real book.

## Outline

Chapter 1 is an introduction to the theory from the cryptography viewpoint. More precisely, several decades after André Weil had stated his conjectures, it appeared that an effective proof of these results would be useful to cryptography. Actually, the $p$-adic approach gives better results in some cases. We recall the discrete logarithm problem and explain why it is useful to explicitly compute the number of points of algebraic varieties. This is the purpose of Weil conjectures which predict the existence of arithmetic cohomology theories that will compute Zeta functions. Rigid cohomology is such a theory and we give its properties. Finally, in Weil's proof of the diagonal hypersurface case, the necessity of introducing coefficients for the theory already appears. This leads to the notion of $L$-functions.

Chapter 2 is devoted to the study of non-archimedean tubes. After fixing the setting (we assume that the reader is familiar with rigid analytic geometry), we introduce successively the notion of open tube of radius one and then the notion of tube of smaller radius. If we are given a subvariety of the special fiber of a formal scheme, the tube is simply the set of points in the generic fiber that specialize into the given subvariety. The idea is to see this tube as a lift of the algebraic variety (of positive characteristic) to an analytic variety (of characteristic zero). Such a tube is not quasi-compact in general, and it is therefore necessary to introduce smaller tubes which are quasi-compact and whose increasing union is the original tube. The main result of this chapter is the Weak Fibration Theorem (Corollary 2.3.16) which says that a smooth morphism of formal scheme around an algebraic variety induces locally a fibration by open balls (or better said, polydiscs). This will imply that the de

Rham cohomology of the tube does not depend on the embedding into the formal scheme.

Chapter 3 is rather technical but fundamental. The de Rham cohomology of a closed ball is infinite dimensional and it is therefore necessary to work in its neighborhood if one is looking for something interesting. Unfortunately, rigid topology is not a usual topology and it is necessary to refine the notion of neighborhood into that of strict neighborhood. We introduce the notion of frame which is a sequence of an open immersion of algebraic varieties and a closed immersion into a formal scheme. We then study the strict neighborhoods of the tube of the first variety into the tube or the second one. For future computations, it is essential to have a deep understanding of these strict neighborhoods. The first idea is to remove a small tube of the complement (the locus at infinity). But this does not give sufficiently general strict neighborhoods. We really need to play around a little more with tubes in order to define the so-called standard strict neighborhoods. It is then possible to extend the Weak Fibration Theorem to strict neighborhoods and obtain the Strong Fibration Theorem (Corollary 3.4.13). This is it for the geometrical part.

Chapter 4 is supposed to be a break. After recalling the basics about modules with integrable connections and their cohomology, we study them in the context of strict neighborhoods and show that this is closely related to the notion of radius of convergence. Actually, modules with integrable connections, $\mathcal{D}$-modules, stratified modules and crystals are simply different ways of seeing the same objects. We try to make this clear in the rigid geometric setting. We work out some examples, Dwork, Kummer, superelliptic curves, Legendre family, hypergeometric equations, etc. Then, we introduce the overconvergence condition for an integrable connection. It means that the Taylor series is actually defined on some strict neighborhood of the diagonal. Locally, there is an explicit description of this condition and this is the main result of the chapter (Theorem 4.3.9). We then introduce the notion of radius of convergence of an integrable connection with respect to a given set of étale coordinates and use it to rewrite the overconvergence condition when the geometry is not too bad. Finally, we also do the case of weakly complete algebras and Robba ring. They will appear to be very important in the future when we try to compute rigid cohomology.

Chapter 5 introduces the notion of overconvergent sheaf. The idea is to work systematically with sections defined on a strict neighborhood. This notion of overconvergence is actually very general and works in any topos but we quickly specialize to the case of frames. In the case of abelian sheaves, we introduce also the notion of sections with overconvergent support as well as the more classical notion of sections with support. Next, we consider the sheaf of overconvergent

sections of the structural sheaf on the tube of a frame, and modules on this ring, which we call dagger modules. The main result of this chapter (Theorem 5.4.4) shows that the category of coherent dagger modules is equivalent to the limit category of coherent modules on strict neighborhoods. In this chapter, we also give a geometric meaning to weakly complete algebras and Robba rings and prove Serre Duality in this context.

Chapter 6 studies dagger modules with integrable connections and their cohomology. We simply apply usual calculus as explained earlier to dagger modules. Rigid cohomology is just usual de Rham cohomology and it can be extended to rigid cohomology with support in a closed subset by using sections with overconvergent support. There is also the alternative theory of cohomology with compact support which is made out of usual sections with support. We give comparison theorems with Monsky-Washnitzer cohomology and de Rham cohomology of Robba rings. The main result of the chapter (Theorem 6.5.2) shows that rigid cohomology of coherent dagger modules with integrable connection is invariant under a morphism of frames which is the identity at the first level, proper at the second and smooth at the third. We also prove the analogous results for cohomology with support. These theorems will prove fundamental later.

Chapter 7 gives a crystalline interpretation of the theory. We define a (finitely presented) overconvergent isocrystal on a frame as a family of coherent dagger modules on all frames above it with some compatibility conditions. We prove in Proposition 7.1.8 that the category of overconvergent isocrystals is invariant under a morphism of frames which is the identity at the first level, proper at the second and smooth at the third. We also show in Proposition 7.2.13 that one recovers exactly the notion of overconvergent integrable connection introduced earlier. In particular, we can define the rigid cohomology of an overconvergent isocrystal as the rigid cohomology of the corresponding module with connection. Next, we consider what we call a virtual frame. This is simply an open immersion of algebraic varieties but we want to see it as an incomplete frame. One defines overconvergent isocrystals on a virtual frame exactly as above and shows that we do get the same category when the virtual frame extends to a smooth frame. Moreover, rigid cohomology is then independent of the chosen extension thanks to our comparison theorems. Thus, it makes sense to talk about the rigid cohomology of an overconvergent isocrystal on a virtual frame.

Chapter 8 rewards us because we may now define overconvergent isocrystals on an algebraic variety and their cohomology in a functorial way. First of all, exactly as above, an overconvergent isocrystal on an algebraic variety is just a family of dagger modules on each frame above the variety with some

compatibility conditions. If we embed successively our variety as an open subset of a proper variety and then as a closed subset of a smooth formal scheme, we get an equivalence with the category of overconvergent isocrystals on the frame as shown in Corollary 8.1.9. Moreover, the cohomology is independent of the choice of the embeddings as showed in Proposition 8.2.3. We finish with the study of Frobenius action. We show that rigid cohomology is fully compatible with Monsky–Washnitzer theory and, in particular, prove that overconvergent $F$-isocrystals correspond exactly to coherent modules with an integrable connection and a strong Frobenius.

Chapter 9 gives some informal complements. We recall what crystalline cohomology is and how it may be used to compute rigid cohomology. This comparison theorem could have been included with a complete proof in the main part of the course but it did not seem reasonable to assume that the reader was familiar with crystalline cohomology. Then, we explain how alterations can be used to derive finiteness of rigid cohomology without coefficient from this comparison theorem. Again, finiteness of rigid cohomology with compact support could have been included with full proof. Unfortunately, the proof for cohomology without support relies on a Gysin isomorphism that requires the theory of arithmetic $\mathcal{D}$-modules. We also explain the Crew conjecture and Kedlaya's methods to solve it. We end with Shiho's theory of convergent log site and his monodromy conjecture which may be seen as a generalization of the conjecture of Crew.

## Conventions and notations

When there is no risk of confusion, we will use standard multi-index notations, namely

$$\underline{i} := i_1, \dots, i_n, \quad |\underline{i}| := i_1 + \cdots + i_n, \quad \underline{i} \leq \underline{j} \Leftrightarrow \forall k, i_k \leq j_k$$

$$\underline{i}! = i_1! \cdots i_n!, \quad \binom{\underline{i}!}{\underline{j}!} := \frac{\underline{i}!}{\underline{j}!(\underline{i-j})!},$$

$$\underline{t}^{\underline{i}} := t_1^{i_1} \cdots t_n^{i_n}, \quad \underline{t}^{[\underline{i}]} := \frac{\underline{t}^{\underline{i}}}{\underline{i}!},$$

and so on. Also, if $X = \cup_{i \in I} X_i$ and $J \subset I$, then $X_J := \cap_{i \in J} X_i$ and if $\lambda_i : X_i \hookrightarrow X$ denotes the inclusion map, we will write $\lambda_J : X_J \hookrightarrow X$ for the inclusion of the intersection.

Throughout this book, we will work on a complete ultrametric field $K$ with a non trivial absolute value. We will denote by $\mathcal{V}$ its valuation ring, $\mathfrak{m}$ its maximal ideal and by $k$ its residue field. Also, $\pi$ will be a non zero element of $\mathfrak{m}$. There is no harm in assuming that $K$ has characteristic zero even if this is almost never used in the theory. The reader will get a better intuition if he also assumes that $k$ has positive characteristic $p$.

Positive real numbers are always assumed to live in $|K^*| \otimes \mathbf{Q} \subset \mathbf{R}_{>0}$.

As usual, if $S$ is any scheme (or ring), then $\mathbf{A}^N_S$ (resp. $\mathbf{P}^N_S$) will denote the affine (resp. projective) space of dimension $N$ over $S$. Also, when $K$ is a complete ultrametric field and $\rho > 0$, then $\mathbf{B}^N(0, \rho^+)$ (resp. $\mathbf{B}^N(0, \rho^-)$) will denote the closed (resp. open) ball (or polydisc) of radius $\rho$. It is the rigid analytic open subset of $\mathbf{A}^{N,\mathrm{rig}}_K$ defined by $|t_i| \leq \rho$ (resp. $|t_i| < \rho$). We may allow $\rho = 0$ in the $+$ case and $\rho = \infty$ in the $-$ case. When $N = 1$, we drop it from the notations. Finally, if $0 < \epsilon < \rho$, then

$$\mathbf{A}_K(0, \epsilon^{\pm}, \rho^{\pm}) := \mathbf{B}(0, \rho^{\pm}) \setminus \mathbf{B}(0, \epsilon^{\mp})$$

will denote the annulus off radii $\epsilon$ and $\rho$. Again, we may allow $\epsilon = \rho$ in the $++$ case as well as $\epsilon = 0$ in the $+\pm$ case and $\rho = \infty$ in the $\pm-$ case.

Since it is sometimes needed in applications, we choose not to assume that varieties or formal schemes are quasi-compact. Many results and definitions are however invalid without this assumption. It is therefore necessary to add this hypothesis form time to time. The reader who so wishes may assume that all (formal) schemes are quasi-compact in order to turn many assertions into a simpler form.